Martingales, chaînes de Markov, systèmes dynamiques

Jean-Pierre Conze et Albert Raugi

Université de Rennes I

2 septembre 2007

Table des matières

Chapitre 1

Modes de convergence utilisés en Probabilités

Nous rappelons ici quelques définitions et résultats usuels en théorie des probabilités. Le lecteur déjà familier avec les notions de base des Probabilités et de la théorie de l'intégration pourra passer directement au chapitre suivant.

Dans toute la suite $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé.

1.1 Convergence presque sûre

Rappels : On rappelle qu'une v.a.r. (variable aléatoire réelle) est une application mesurable de $(\Omega, \mathcal{F})$ dans $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$; où $\overline{\mathbb{R}} = [-\infty, +\infty]$ et $\mathcal{B}(\overline{\mathbb{R}})$ est la tribu des parties boréliennes de $\overline{\mathbb{R}}$ (i.e. la tribu engendrée par les intervalles (ouverts, semi-ouverts, fermés) de $\overline{\mathbb{R}}$).

Une v.a.r. est dite **presque sûrement finie** si $\mathbb{P}[X \in \mathbb{R}] = 1$ (i.e. $\mathbb{P}[X = \pm\infty] = 0$).

Deux v.a.r. X et X' sont égales presque sûrement (ou presque partout) si $\mathbb{P}[X \neq X'] = 0$. Cette relation, notée $X \overset{\text{p.s.}}{=} X'$ (ou $X = X'$ $\mathbb{P}$-p.s., quand il peut y avoir ambiguité), est manifestement une relation d'équivalence compatible avec la structure d'algèbre de l'ensemble des v.a.r. (i.e. $X \overset{\text{p.s.}}{=} X'$ et $Y \overset{\text{p.s.}}{=} Y'$ entraîne $\lambda X + \mu X' \overset{\text{p.s.}}{=} \lambda Y + \mu Y'$, pour tout $\lambda, \mu \in \mathbb{R}$ et $XY \overset{\text{p.s.}}{=} X'Y'$). En outre si $X_i \overset{\text{p.s.}}{=} X'_i$, pour tout $i \in I$, ensemble d'indices dénombrable, on a $\sup\limits_I X'_i \overset{\text{p.s.}}{=} \sup\limits_I X'_i$ et $\inf\limits_I X'_i \overset{\text{p.s.}}{=} \inf\limits_I X'_i$.

Comme on le verra par la suite, la plupart des problèmes de la théorie des probabilités font intervenir les classes d'équivalence de v.a.r. plutôt que ces v.a.r. elles-mêmes. Assez souvent, on ne distinguera pas une v.a.r. de sa classe d'équivalence pour l'égalité presque sûre.

Soit $(X_n)_{n\geq 0}$ une suite de v.a.r., p.s. finies. L'**ensemble de convergence** de la suite $(X_n)_{n\geq 0}$ est défini par :

$$\{\omega \in \Omega : \text{ la suite réelle } (X_n(\omega))_{n\geq 0} \text{ converge } \}.$$

En introduisant les v.a.r.

$$\liminf_{n\to+\infty} X_n = \lim_{n\to+\infty} \uparrow \inf_{k\geq n} X_k \quad \text{et} \quad \limsup_{n\to+\infty} X_n = \lim_{n\to+\infty} \downarrow \sup_{k\geq n} X_k,$$

le domaine de convergence de la suite s'écrit :

$$\{\omega \in \Omega : -\infty < \liminf_{n \to +\infty} X_n(\omega) = \limsup_{n \to +\infty} X_n(\omega) < +\infty\}.$$

Ce qui montre que cet ensemble est mesurable.

Plus généralement, en se plaçant dans la droite achevée $\overline{\mathbb{R}}$, le domaine de convergence dans $\overline{\mathbb{R}}$ d'une suite de v.a.r. $(X_n)_{n\geq 0}$ est défini par :

$$\{\omega \in \Omega : \text{ la suite réelle } (X_n(\omega))_{n\geq 0} \text{ admet une limite finie ou non }\}.$$

Nous notons $\{X_n \to \}$ cet ensemble. Nous avons

$$\{X_n \to \} = \{\liminf_{n \to +\infty} X_n = \limsup_{n \to +\infty} X_n\} \tag{1.1}$$

et $\{X_n \to \}$ appartient à $\mathcal{F}$.

Définition 1.1.1 Nous disons qu'une suite de v.a.r. $(X_n)_{n\geq 0}$ **converge presque sûrement** si $\mathbb{P}[\{X_n \to \}] = 1$ (i.e. d'après (1.1) si $\liminf_{n \to +\infty} X_n \overset{\text{p.s.}}{=} \limsup_{n \to +\infty} X_n$).

La limite de la suite $(X_n)_{n\geq 0}$ est alors par définition l'une quelconque des v.a.r. de la classe d'équivalence de $\liminf_{n \to +\infty} X_n$ (ou $\limsup_{n \to +\infty} X_n$). On note $X = \lim_{n \to +\infty} X_n$ cette classe d'équivalence ou l'un quelconque de ses éléments et on écrit $X_n \overset{\text{p.s.}}{\to} X$.

• **Critères de convergence**

Du critère de Cauchy pour les suites réelles, il résulte que l'on a :

Critère de Cauchy : Pour qu'une suite de v.a.r. $(X_n)_{n\geq 0}$, p.s. finies converge vers une v.a.r. p.s. finie, il faut et il suffit qu'elle soit une suite de Cauchy pour la convergence p.s. ; c'est-à-dire qu'il existe un sous-ensemble $\mathcal{F}$-mesurable Ω_0 de Ω, de $\mathbb{P}$-mesure 1, tel que, pour tout $\omega \in \Omega_0$, la suite réelle $(X_n(\omega))_{n\geq 0}$ soit de Cauchy :

$$\forall \varepsilon > 0,\ \exists N,\ \forall p \geq 0,\ \forall n \geq N,\ |X_{n+p}(\omega) - X_n(\omega)| < \varepsilon.$$

Voici quelques critères pratiques dont les démonstrations sont laissées en exercice au lecteur.

Une suite de v.a.r., p.s. finies, $(X_n)_{n\geq 0}$ converge p.s. vers une v.a.r. p.s. finie si et seulement si

$$\forall \varepsilon > 0,\ \lim_{n \to +\infty} \mathbb{P}[\bigcup_{p\geq 0} \{|X_{n+p} - X_n| > \varepsilon\}] \ = \ 0. \tag{1.2}$$

(Cette équivalence ne fait que traduire le critère de Cauchy !).

On en déduit la condition suffisante de convergence p.s. : *il existe une série positive convergente $\sum a_n$ telle que la série $\sum \mathbb{P}[\{|X_{n+1} - X_n| > a_n\}]$ soit convergente.*

Une suite de v.a.r. p.s. finies $(X_n)_{n\geq 0}$ converge p.s. vers une v.a.r. p.s. finie X si, et seulement si,

$$\forall \varepsilon > 0,\ \mathbb{P}[\limsup_{n} \{|X_n - X| > \varepsilon\}] = 0.$$

On rappelle que, pour une suite $(A_n)_{n\geq 0}$ de parties de Ω on définit les parties $\limsup_n A_n$ et $\liminf_n A_n$ par

$$\limsup_n A_n = \bigcap_n \bigcup_{k\geq n} A_k \quad \text{et} \quad \liminf_n A_n = \bigcup_n \bigcap_{k\geq n} A_k.$$

Une condition suffisante de convergence p.s vers X est donc

$$\forall \varepsilon > 0,\ \sum_{k\geq 0} \mathbb{P}[\{|X_k - X| > \varepsilon\}] < +\infty \ ;$$

ou encore : *il existe une suite positive $(\varepsilon_n)_{n\geq 0}$ décroissant vers zéro telle que la série $\sum \mathbb{P}[\{|X_n - X| > \varepsilon_n]$ soit convergente.*

Citons aussi le résultat suivant :

Théorème 1.1.2 *(Egoroff) : Soit $(X_n)_{n\geq 0}$ une suite de v.a.r. p.s. finies convergeant vers une v.a.r. X p.s. finie. Alors, pour tout $\varepsilon > 0$, il existe un sous-ensemble mesurable A_ε de Ω de probabilité $\geq 1-\varepsilon$ tel que la convergence de $(X_n)_{n\geq 0}$ vers X soit uniforme sur A_ε (i.e. $\lim_{n\to+\infty} \sup_{\omega\in A_\varepsilon} |X_n(\omega) - X(\omega)| = 0$).*

Preuve : Puisque $X_n \overset{\text{p.s.}}{\to} X$, nous avons d'après ce qui précède, pour tout entier $p \geq 1$,

$$\lim_n \downarrow \mathbb{P}[\bigcup_{k\geq n} \{|X_k - X| > \frac{1}{p}\}] = \mathbb{P}[\limsup_n \{|X_n - X| > \frac{1}{p}\}] = 0.$$

Nous pouvons donc trouver un entier n_p tel que

$$\mathbb{P}[\bigcup_{k\geq n_p} \{|X_k - X| > \frac{1}{p}\}] \leq \frac{1}{2^p}.$$

Pour tout réel $\varepsilon > 0$, choisissons alors un entier $p_0 \geq 1$ tel que $\frac{1}{2^{p_0-1}} < \varepsilon$ et posons

$$A_\varepsilon = [\bigcup_{p\geq p_0} \bigcup_{k\geq n_p} \{|X_k - X| > \frac{1}{p}\}]^c = \bigcap_{p\geq p_0} \bigcap_{k\geq n_p} \{|X_k - X| \leq \frac{1}{p}\}.$$

Nous avons, d'une part :

$$\mathbb{P}[A_\varepsilon^c] \leq \sum_{p\geq p_0} \mathbb{P}[\bigcup_{k\geq n_p} \{|X_k - X| > \frac{1}{p}\}] \leq \sum_{p\geq p_0} \frac{1}{2^p} = \frac{1}{2^{p_0-1}} < \varepsilon,$$

et d'autre part :

$$\forall p \geq p_0,\ \forall k \geq n_p,\ \sup_{\omega\in A_\varepsilon} |X_k(\omega) - X(\omega)| \leq \frac{1}{p}.$$

□

1.2 Convergence en probabilité

Définition 1.2.1 Une suite $(X_n)_{n\geq 0}$ de v.a.r., p.s. finies, **converge en probabilité** vers une v.a.r., p.s. finie X, si

$$\forall \alpha > 0, \lim_{n\to+\infty} \mathbb{P}[\{|X_n - X| > \alpha\}] \; = \; 0.$$

En posant $d(X,Y) = \mathbb{E}[|X-Y| \wedge 1]$, on définit une distance sur les classes de v.a.r. qui donne la convergence en probabilité (exercice).

Le résultat suivant précise le lien existant entre la convergence p.s. et la convergence en probabilité.

Proposition 1.2.2 1) *Toute suite $(X_n)_{n\geq 0}$ de v.a.r., p.s. finies, qui converge p.s. vers une v.a.r., p.s. finie, converge en probabilité vers la même limite.*

2) *De toute suite $(X_n)_{n\geq 0}$ de v.a.r., p. s. finies, qui converge en probabilité, on peut extraire une sous-suite qui converge p.s. vers la même limite.*

Preuve : Exercice. Pour la seconde assertion, utiliser la condition suffisante de (1.2). □

Lemme 1.2.3 *(Critère de Cauchy) Une suite $(X_n)_{n\geq 0}$ de v.a.r., p.s. finies, converge en probabilité si, et seulement si, elle est une suite de Cauchy pour la convergence en probabilité ; c'est-à-dire, si $X_m - X_n \xrightarrow{\mathbb{P}} 0$, lorsque m et n tendent vers $+\infty$, ou encore*

$$\forall \alpha > 0, \forall \varepsilon > 0, \; \exists N \in \mathbb{N}, \; \forall n \geq N, \; \forall p \geq 0, \; \mathbb{P}[\{|X_{n+p} - X_n| > \alpha\}] < \varepsilon.$$

Preuve : 1) Supposons que $X_n \xrightarrow{\mathbb{P}} X$. Pour tout $\varepsilon > 0$, nous avons :

$$\{|X_m - X_n| > \varepsilon\} \subset \{|X_m - X| > \frac{\varepsilon}{2}\} \cup \{|X_n - X| > \frac{\varepsilon}{2}\}$$

et par suite

$$\mathbb{P}[\{|X_m - X_n| > \varepsilon\}] \leq \mathbb{P}[\{|X_m - X| > \frac{\varepsilon}{2}\}] + \mathbb{P}[\{|X_n - X| > \frac{\varepsilon}{2}\}] \to 0,$$

lorsque m et n tendent vers $+\infty$.

2) Supposons que $X_m - X_n \xrightarrow{\mathbb{P}} 0$, lorsque m et n tendent vers $+\infty$. Il existe alors une suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$ telle que

$$\mathbb{P}[\{|X_{\varphi(n+1)} - X_{\varphi(n)}| > \frac{1}{2^n}\}] \; \leq \frac{1}{3^n}.$$

D'après (1.2), la suite $(X_{\varphi(n)})_{n\geq 0}$ converge p.s. (et donc en probabilité) vers une v.a.r. X. Pour tout $\varepsilon > 0$, nous avons alors

$$\mathbb{P}[\{|X_n - X| > \varepsilon\}] \leq \mathbb{P}[\{|X_n - X_{\varphi(n)}| > \frac{\varepsilon}{2}\}] + \mathbb{P}[\{|X_{\varphi(n)} - X| > \frac{\varepsilon}{2}\}],$$

majorant qui tend vers zéro quand n tend vers l'infini. □

Exemple 1.2.4

Exemple de suite de v.a.r., p.s. finies, convergeant en probabilité mais ne convergeant pas presque sûrement

Considérons l'espace probabilisé $([0,1[,\mathcal{B}([0,1[),\lambda)$, où λ désigne la mesure de Lebesgue sur les boréliens $\mathcal{B}([0,1[)$ de l'intervalle $[0,1[$.

Considérons les intervalles dyadiques de $[0,1[$ ordonnés de la façon suivante :

$$A_1=[0,1[,\ A_2=[0,\frac{1}{2}[,\ A_3=[\frac{1}{2},1[,\ A_4=[0,\frac{1}{4}[,\ A_5=[\frac{1}{4},\frac{1}{2}[,\ A_6=[\frac{1}{2},\frac{3}{4}[,\dots$$

Autrement dit, si $n=2^k+p$ avec $k\geq 0$ et $0\leq p<2^k$, alors $A_n=[\frac{p}{2^k},\frac{p+1}{2^k}[$. (On notera que k est la partie entière $\mathrm{E}(\log_2 n)$ de $\log_2 n$).

Pour tout entier $n\geq 1$, posons $X_n=1_{A_n}$. Pour tout réel x de $[0,1[$, il existe une infinité d'intervalles dyadiques qui contiennent x et une infinité qui ne le contiennent pas ; par suite

$$\liminf_{n\to+\infty}\ X_n(x)=0\ \text{ et }\ \limsup_{n\to+\infty}\ X_n(x)=1.$$

D'autre part, pour tout $0<\varepsilon<1$, nous avons $\mathbb{P}[\{|X_n|>\varepsilon\}]=\lambda(A_n)=\frac{1}{\mathrm{E}(\log_2 n)}$, qui tend vers zéro, quand n tend vers $+\infty$. On a donc : $X_n\overset{\mathbb{P}}{\to}0$.

1.3 Convergence en moyenne et équi-intégrabilité

Définition 1.3.1 Nous disons qu'une suite $(X_n)_{n\geq 0}$ de (classes d'équivalence de) v.a.r. intégrables **converge en moyenne (d'ordre 1)** vers la (classe d'équivalence de la) v.a.r. intégrable X si

$$\lim_{n\to+\infty}\int_\Omega |X_n-X|\ d\mathbb{P}\ =\ 0.$$

Nous écrivons alors $X_n\overset{L^1}{\to}X$.

L'exemple 1.2.4 prouve que **la convergence en moyenne n'entraîne pas la convergence p.s.**.

Des inégalités élémentaires

$$\forall\alpha>0,\ \mathbb{P}[\{|X|>\alpha\}]\leq\frac{1}{\alpha}\int_\Omega |X|\ d\mathbb{P}, \tag{1.3}$$

il résulte que *la convergence en moyenne implique la convergence en probabilité*. Mais la réciproque est fausse, comme le prouve le contre-exemple suivant.

Considérons l'espace probabilisé $(]0,1],\mathcal{B}(]0,1]),\lambda)$ et la suite de v.a.r. $(X_n)_{n\geq 1}$ définie par :

$$X_n(x)=\{\,n\quad si\quad x\in]0,\tfrac{1}{n}]\tfrac{1}{nx^2}\quad si\quad x\in]\tfrac{1}{n},1]\ .$$

Alors $X_n\overset{\text{p.s.}}{\to}0$ et $\mathbb{E}[X_n]\to 2$, lorsque n tend vers $+\infty$.

Cependant nous savons (théorème de convergence dominée) que si $X_n\overset{\text{p.s.}}{\to}X$ et si la suite $(|X_n|)_{n\geq 0}$ est majorée par une v.a.r. intégrable, alors $X_n\overset{L^1}{\to}X$. Mais ces conditions suffisantes ne sont pas nécessaires comme le prouve l'exemple suivant.

Considérons l'espace probabilisé $(]0,1], \mathcal{B}(]0,1]), \lambda)$ et la suite de v.a.r. $(X_n)_{n\geq 1}$ définie par :

$$X_n(x) = \{ \frac{1}{\sqrt{x}} \quad si \quad x \in]0, \frac{1}{n+1}[\cup]\frac{1}{n}, 1] \frac{1}{x} \quad si \quad x \in]\frac{1}{n+1}, \frac{1}{n}] \ .$$

Soit X la v.a.r. définie par $X(x) = \frac{1}{\sqrt{x}}$, $x \in]0,1]$. Nous avons : $X_n \overset{\text{p.s.}}{\to} X$; $X_n \overset{L^1}{\to} X$ et $\sup_{n\geq 1} X_n(x) = \frac{1}{x}$ n'est pas λ-intégrable.

Soit $(X_n)_{n\geq 0}$ une suite de v.a.r. intégrables. Si la série $\sum \mathbb{E}[|X_{n+1} - X_n|]$ converge, alors la suite de v.a.r. $(X_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. et en moyenne.

En particulier, si une suite de v.a.r. intégrables $(X_n)_{n\geq 0}$ est une suite de Cauchy dans $L^1(\Omega, \mathcal{F}, \mathbb{P})$, alors il existe une sous-suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$ telle que $\mathbb{E}[|X_{\varphi(n+1)} - X_{\varphi(n)}|] \leq \frac{1}{2^n}$; la suite de v.a.r. $(X_{\varphi(n)})_{n\geq 0}$ converge alors $\mathbb{P}$-p.s. et en moyenne.

On en déduit que toute suite de Cauchy pour la convergence en moyenne converge en moyenne.

A ce stade, nous pouvons donc dire que la convergence en moyenne se situe "entre" la convergence en probabilité et les conditions "convergence p.s." plus "domination par une fonction intégrable". Pour trouver des conditions équivalentes à la convergence en moyenne, nous sommes amenés à introduire la notion **d'équi-intégrabilité** (ou **d'uniforme intégrabilité** chez certains auteurs).

Définition 1.3.2 Nous disons qu'une famille $\mathcal{H}$ de v.a.r. intégrables est **équi-intégrable** si

$$\lim_{C\to+\infty} \downarrow \sup_{X\in\mathcal{H}} \int_{\{|X|>C\}} |X| \, d\mathbb{P} \ = \ 0.$$

Exemples et exercices 1.3.3

1. Soient $\mathcal{H}$ et $\mathcal{H}'$ deux familles de v.a.r. intégrables telles que : $\forall X \in \mathcal{H},\ \exists Y \in \mathcal{H}',\ |X| \leq |Y|$. Alors si la famille $\mathcal{H}'$ est équi-intégrable, la famille $\mathcal{H}$ l'est aussi. En particulier, toute famille finie de v.a.r. intégrables est équi-intégrable et toute famille $\mathcal{H}$ de v.a.r. majorées, en valeur absolue, par une v.a.r. Y intégrable (i.e. $\forall X \in \mathcal{H},\ |X| < Y$), est équi-intégrable.

2. Si $\mathcal{H}$ et $\mathcal{H}'$ sont deux familles équi-intégrables de v.a.r. intégrables, alors les familles $\mathcal{H} + \mathcal{H}' = \{X + Y : X \in \mathcal{H}, Y \in \mathcal{H}'\}$ et $\mathcal{H} \cup \mathcal{H}'$ sont aussi équi-intégrables.

3. Soit $(X_n)_{n\geq 0}$ une suite de v.a.r. convergeant vers X en moyenne. Alors la famille $\mathcal{H} = \{X_n : n \geq 0\} \cup \{X\}$ est équi-intégrable.

[D'après 1. et 2. en écrivant $|X_n - X| \leq |X| + |X_n|$, on se ramène au cas où la suite converge en moyenne vers la fonction nulle.]

4. Soit $\mathcal{H}$ une famille de v.a.r. intégrables. S'il existe une fonction f de $\mathbb{R}_+$ dans $\mathbb{R}_+$ telle que

$$\lim_{x\to+\infty} \frac{f(x)}{x} = +\infty \quad \text{et} \quad \sup_{X\in\mathcal{H}} \int_\Omega f(|X|) \, d\mathbb{P} \ < \ +\infty,$$

alors la famille $\mathcal{H}$ est équi-intégrable.

En particulier toute famille $\mathcal{H}$ de v.a.r. bornée dans L^p (i.e. $\sup_{X\in\mathcal{H}} \mathbb{E}[|X|^p] < +\infty$), pour $p > 1$, est équi-intégrable.

5. Soit X une v.a.r. intégrable. La famille de v.a.r. intégrables

$$\{\mathbb{E}[X|\mathcal{A}] : \mathcal{A} \text{ sous-tribu de } \mathcal{F}\}$$

est équi-intégrable.

6. Soient $\mathcal{A}$ une sous-tribu de $\mathcal{F}$ et $(X_n)_{n\geq 0}$ une suite de v.a.r. intégrables telle que la famille $\{X_n : n \geq 0\}$ soit équi-intégrable. Alors la famille $\{\mathbb{E}[X_n|\mathcal{A}] : n \geq 0\}$ est équi-intégrable.

7. Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. indépendantes et de même loi. On suppose que cette loi possède un moment d'ordre 1 (i.e. $\mathbb{E}[|X_1|] < +\infty$). Montrer que la famille $\{\dfrac{X_1 + \ldots + X_n}{n},\ n \geq 1\}$ est équi-intégrable.

8. Sans utiliser le fait que l'espace $L^1(\Omega, \mathcal{F}, \mathbb{P})$ est complet, montrer que toute suite de Cauchy de cet espace constitue une famille équi-intégrable.

Proposition 1.3.4 *Une famille $\mathcal{H}$ de v.a.r. intégrables est équi-intégrable, si, et seulement si, elle vérifie les deux conditions :*

i) (équi-continuité) Pour tout $\varepsilon > 0$, il existe un réel $\eta(\varepsilon) > 0$ tel que $\sup\limits_{X\in\mathcal{H}} \displaystyle\int_A |X|\, d\mathbb{P} \;\leq\; \varepsilon$, pour tout $A \in \mathcal{F}$ vérifiant $\mathbb{P}[A] < \eta(\varepsilon)$.

ii) La famille $\mathcal{H}$ est bornée dans L^1 (i.e. $\sup\limits_{X\in\mathcal{H}} \mathbb{E}[|X|] \;<\; +\infty$).

<u>Preuve</u> : Exercice.

Le théorème suivant nous donne des conditions nécessaires et suffisantes de convergence en moyenne. Il généralise le théorème de convergence dominée.

Théorème 1.3.5 *Soient $(X_n)_{n\geq 0}$ une suite de v.a. r. intégrables. Pour toute v.a.r X, p.s. finie, les deux conditions suivantes sont équivalentes :*

i) la famille $\{X_n : n \geq 0\}$ est équi-intégrable et $X_n \xrightarrow{\mathbb{P}} X$, quand n tend vers $+\infty$;

ii) la v.a.r. X est intégrable et $X_n \xrightarrow{L^1} X$, quand n tend vers $+\infty$.

<u>Preuve</u> : $i) \Rightarrow ii)$ D'après la proposition 1.2.2, il existe une suite strictement croissante d'entiers naturels telle que la suite de v.a.r. $(X_{\varphi(n)})_{n\geq 0}$ converge p.s. vers une v.a.r. X. Du lemme de Fatou il résulte alors que

$$\begin{aligned}\int_\Omega |X|\, d\mathbb{P} \;&=\; \int_\Omega \liminf_n |X_{\varphi(n)}|\, d\mathbb{P} \;\leq\; \liminf_n \int_\Omega |X_{\varphi(n)}|\, d\mathbb{P} \\ &\leq\; \sup_{n\geq 0} \int_\Omega |X_n|\, d\mathbb{P} \;<\; +\infty\end{aligned}$$

car la famille $\{X_n : n \geq 0\}$ est équi-intégrable et donc bornée dans L^1 (Proposition 1.3.4).

D'autre part, pour tous réels strictement positifs ε et C, nous avons :

$$\begin{aligned}\mathbb{E}[|X_n - X|] \;&\leq\; \mathbb{E}[|X_n - X|1_{\{|X_n - X|\leq\varepsilon\}}] + \mathbb{E}[|X_n - X|1_{\{\varepsilon<|X_n - X|\leq C\}}] \\ &\qquad + \mathbb{E}[|X_n - X|1_{\{|X_n - X|>C\}}] \\ &\leq\; \varepsilon + C\ \mathbb{P}[\{|X_n - X| > \varepsilon\}] + \mathbb{E}[|X_n - X|1_{\{|X_n - X|>C\}}] \\ &\leq\; \varepsilon + C\ \mathbb{P}[\{|X_n - X| > \varepsilon\}] + \sup_{p\geq 0}\ \mathbb{E}[|X_p - X|1_{\{|X_p - X|>C\}}],\end{aligned}$$

et par suite :

$$\limsup_n\ \mathbb{E}[|X_n - X|] \;\leq\; \varepsilon + \sup_{p\geq 0}\ \mathbb{E}[|X_p - X|1_{\{|X_p - X|>C\}}].$$

Des inégalités $\forall n \geq 0,\ |X_n - X| \leq |X_n| + |X|$ et de l'équi-intégrabilité de la famille $\{X_n : n \geq 0\}$, il résulte (cf. exercice 2 de 1.3.3) que la famille $\{X_n - X : n \geq 0\}$ est équi-intégrable. On obtient alors la convergence en moyenne en faisant tendre ε vers zéro et C vers $+\infty$.

ii) $\Rightarrow$ *i)* La convergence en moyenne implique la convergence en probabilité (cf. (1.3)). D'autre part nous avons vu (cf. exercice 3 de (1.3.3)) que la convergence en moyenne de la suite de v.a.r. $(X_n)_{n\geq 0}$ entraîne l'équi-intégrabilité de la famille $\{X_n : n \geq 0\}$. $\square$

1.4 Espaces L^p, $p \geq 1$

Pour toute (classe de) v.a.r. X, introduisons les nombres :

$$\|X\|_p = (\int_\Omega |X|^p \, d\mathbb{P})^{\frac{1}{p}} \in [0, +\infty],\ 1 \leq p < +\infty,$$
$$\|X\|_\infty = \sup \{M \in [0, +\infty] : \mathbb{P}[\{|X| > M\}] > 0\}.$$

Pour tout $p \in [1, +\infty]$, nous notons $L^p = L^p(\Omega, \mathcal{F}, \mathbb{P})$ l'ensemble des classes d'équivalence de v.a.r. X telles que $\|X\|_p < +\infty$.

Rappelons (sans démonstration) les résultats suivants.

Inégalité de convexité : *Si h est une fonction réelle, continue et concave, définie sur un domaine convexe D de $\mathbb{R}^d$, nous avons, pour toutes v.a.r. $X_1, \ldots, X_d$ telles que $(E[X_1], \ldots, \mathbb{E}[X_d]) \in D$ et $(X_1, \ldots, X_d) \in D$ p.s. :*

$$\mathbb{E}[h(X_1, \ldots, X_d)] \ \leq \ h(E[X_1], \ldots, \mathbb{E}[X_d]).$$

En particulier, on en déduit que, pour toute v.a.r. X,

$$\|X\|_1 \leq \|X\|_p \leq \|X\|_q \leq \|X\|_\infty, \forall 1 \leq p \leq q \leq +\infty.$$

Inégalité de Hölder : *Soient p, q et r des éléments de $[1, +\infty]$ tels que $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Alors, pour toutes v.a.r. X et Y, nous avons $\|X\ Y\|_r \leq \|X\|_p\ \|Y\|_q$.*

En particulier $\|X\ Y\|_1 \leq \|X\|_p\ \|Y\|_q$ si $p, q \in [1, +\infty[$ et $\frac{1}{p} + \frac{1}{q} = 1$. Pour $p = q = 2$, on obtient l'inégalité de Cauchy-Schwarz : $\|X\ Y\|_1 \leq \|X\|_2\ \|Y\|_2$.

Inégalité de Minkowski : *Pour toutes v.a.r. X et Y et pour tout $p \in [1, +\infty]$, $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$.*

Pour tout $p \in [1, +\infty]$, $(L^p, \|\ \|_p)$ est un espace vectoriel normé. Pour $p = 2$, L^2 est un espace préhilbertien pour le produit scalaire $\langle X, Y\rangle = \int_\Omega X\ Y\ d\mathbb{P}$.

Définition 1.4.1 Soit $p \in [1, +\infty]$. Soient $(X_n)_{n\geq 0}$ et X des (classes de) v.a.r appartenant à L^p. Nous disons que la suite de v.a.r. $(X_n)_{n\geq 0}$ **converge en norme** L^p (ou converge en moyenne d'ordre p) vers une v.a.r. X, si la suite réelle positive $(\|X_n - X\|_p)_{n\geq 0}$ converge vers zéro quand n tend vers $+\infty$. On écrit alors $X_n \xrightarrow{L^p} X$.

Du théorème 1.3.5, il résulte (exercice) :

Théorème 1.4.2 *Soit $p \in [1, +\infty[$. Soit $(X_n)_{n\geq 0}$ une suite de (classes de) v.a.r. appartenant à L^p. Pour toute v.a.r. X, les assertions suivantes sont équivalentes :*

i) la famille $\{|X_n|^p : n \geq 0\}$ est équi-intégrable et $X_n \stackrel{\mathbb{P}}{\to} X$, $n \to +\infty$;

ii) $X \in L^p$ et $X_n \stackrel{L^p}{\to} X$, quand n tend vers $+\infty$.

Lemme 1.4.3 *(Critère de Cauchy dans L^p) Soit $p \in [1, +\infty]$. Une suite $(X_n)_{n\geq 0}$ de (classes de) v.a.r. appartenant à L^p converge en moyenne d'ordre p si, et seulement si, c'est une suite de Cauchy pour la convergence en moyenne d'ordre p (i.e. $\|X_m - X_n\|_p \to 0$, lorsque m et n tendent vers $+\infty$).*

Preuve : Exercice.

Corollaire 1.4.4 *Pour tout $p \in [1, +\infty]$, l'espace L^p est un espace de Banach. L'espace L^2 est un espace de Hilbert.*

Théorème 1.4.5 *Soit $p \in [1, +\infty[$. Soient $\{X_n : n \geq 0\}$ et X des (classes de) v.a.r. appartenant à L^p. Nous avons :*

$$X_n \stackrel{L^p}{\to} X \Longleftrightarrow \{ X_n \stackrel{\mathbb{P}}{\to} X \|X_n\|_p \to \|X\|_p .$$

Preuve : L'implication $\Longrightarrow$ est claire.

Pour prouver l'implication inverse, nous commençons par montrer que

$$\{ X_n \stackrel{\text{p.s.}}{\to} X \|X_n\|_p \to \|X\|_p \Longrightarrow X_n \stackrel{L^p}{\to} X.$$

Soit $A \in \mathcal{F}$. D'après le lemme de Fatou, nous avons

$$\int_A |X|^p \, d\mathbb{P} \leq \liminf_n \int_A |X_n|^p \, d\mathbb{P}.$$

Cela dit, nous avons :

$$\begin{aligned}
\limsup_n \int_A |X_n|^p \, d\mathbb{P} &= \limsup_n (\mathbb{E}[|X_n|^p] - \int_{A^c} |X_n|^p \, d\mathbb{P}) \\
&= \mathbb{E}[|X|^p] - \liminf_n \int_{A^c} |X_n|^p \, d\mathbb{P} \\
&\leq \mathbb{E}[|X|^p] - \int_{A^c} |X|^p \, d\mathbb{P} = \int_A |X|^p \, d\mathbb{P}.
\end{aligned}$$

On en déduit donc que :

$$\forall A \in \mathcal{F}, \int_A |X_n|^p \, d\mathbb{P} \longrightarrow \int_A |X|^p \, d\mathbb{P}, \quad \text{quand } n \to +\infty.$$

Soit $\varepsilon > 0$. D'après le théorème d'Egoroff (cf. (1.1.2)), nous pouvons trouver un sous-ensemble mesurable Ω_ε de Ω, vérifiant $\mathbb{P}[\Omega_\varepsilon] \geq 1 - \varepsilon$ et sur lequel la suite de v.a.r. $(X_n)_{n\geq 0}$ converge uniformément vers X. Nous avons alors :

$$\begin{aligned}
\limsup_n \mathbb{E}[|X_n - X|^p] &= \limsup_n \int_{\Omega_\varepsilon^c} |X_n - X|^p \, d\mathbb{P} \\
&\leq 2^{p-1} \limsup_n (\int_{\Omega_\varepsilon^c} |X_n|^p \, d\mathbb{P} + \int_{\Omega_\varepsilon^c} |X|^p \, d\mathbb{P}) \\
&\leq 2^p \int_{\Omega_\varepsilon^c} |X|^p \, d\mathbb{P}.
\end{aligned}$$

On obtient alors le résultat cherché en faisant tendre ε vers zéro.

Plaçons nous à présent sous les hypothèses

$$\{ \, X_n \xrightarrow{\mathbb{P}} X \|X_n\|_p \to \|X\|_p$$

Si la suite de v.a.r. $(X_n)_{n\geq 0}$ ne convergeait pas vers X en norme L^p, il existerait un réel $\varepsilon_0 > 0$ et une suite strictement croissante d'entiers naturels $(\varphi(n))_{n\geq 0}$ tels que

$$\forall n \geq 0, \;\; \|X_{\varphi(n)} - X\|_p > \varepsilon_0. \tag{1.4}$$

Comme $X_{\varphi(n)} \xrightarrow{\mathbb{P}} X$, on peut extraire une sous-suite $(\varphi \circ \psi(n))_{n\geq 0}$ de la suite $(\varphi(n))_{n\geq 0}$ telle que $X_{\varphi\circ\psi(n)} \xrightarrow{\text{p.s.}} X$. D'après ce qui précède, on aurait alors $X_{\varphi\circ\psi(n)} \xrightarrow{L^p} X$; ce qui contredit $(*)$.

$\square$

Exercice : Soient $\{X_n \, : \, n \geq 0\}$ et X des v.a.r. intégrables *positives*. En utilisant l'égalité $|X_n - X| = X_n + X - 2 \; X_n \wedge X$, montrer (sans faire appel au théorème précédent) l'équivalence suivante :

$$X_n \xrightarrow{L^1} X \Longleftrightarrow \{ \, X_n \xrightarrow{\mathbb{P}} X \mathbb{E}[X_n] \to \mathbb{E}[X].$$

Chapitre 2

Martingales

Nous rappelons ici quelques éléments de la théorie des martingales à temps discret qui joue un rôle important dans l'étude des propriétés stochastiques des systèmes dynamiques et des processus markoviens.

Dans toute la suite $(\Omega, \mathcal{F}, \mathbb{P})$ désignera un espace probabilisé. On notera $\mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$, l'espérance d'une variable aléatoire X à valeurs réelles (en abrégé v.a.r.) intégrable.

2.1 Processus, filtrations et temps d'arrêt

Définition 2.1.1 Un **processus** sur $(\Omega, \mathcal{F}, \mathbb{P})$ est une suite $X = (X_n)_{n\geq 0}$ de variables aléatoires définies sur cet espace, à valeurs dans un espace mesurable.

On appelle **filtration** (croissante) sur $(\Omega, \mathcal{F}, \mathbb{P})$ toute suite croissante $(\mathcal{F}_n)_{n\geq 0}$ de sous-tribus de F.

La réunion des tribus $\mathcal{F}_n$, $n \geq 0$, est alors une algèbre de Boole. On note $\mathcal{F}_\infty$ la tribu $\bigvee_{n\geq 0} \mathcal{F}_n$ engendrée par cette algèbre.

Un processus $(X_n)_{n\geq 0}$ est dit **adapté** à une filtration $(\mathcal{F}_n)_{n\geq 0}$ si, pour tout $n \geq 0$, X_n est $\mathcal{F}_n$-mesurable.

Etant donné un processus $X = (X_n)_{n\geq 0}$, on note $\mathcal{F}_n^X = \sigma(X_0, \cdots, X_n)$ la tribu engendrée par les v.a. X_0,...,X_n. La suite $(\mathcal{F}_n^X)_{n\geq 0}$ forme une filtration appelée **filtration naturelle** associée au processus X.

Tout processus est adapté à sa filtration naturelle. Si X est adapté à une filtration $(\mathcal{G}_n)_{n\geq 0}$, on a $\mathcal{F}_n^X \subset \mathcal{G}_n$, $\forall n \geq 0$.

Définition 2.1.2 Soit $(\mathcal{F}_n)_{n\geq 0}$ une filtration sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Une variable aléatoire T définie sur $(\Omega, \mathcal{F})$, à valeurs dans $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$ est un **temps d'arrêt** pour la filtration $(\mathcal{F}_n)_{n\geq 0}$ si, pour tout entier $n \geq 0$, l'ensemble $\{T = n\}$ est $\mathcal{F}_n$-mesurable.

Si T est à valeurs dans $\mathbb{N}$, on dit que T est **fini**. Si $\mathbb{P}[\{T = +\infty\}] = 0$, on dit que T est **$\mathbb{P}$-p.s. fini**.

Exercices 2.1.3

1. Soit T un temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Montrer que :

a) pour tout entier $n \geq 0$, les ensembles $\{T \leq n\}$ et $\{T > n\}$ sont $\mathcal{F}_n$-mesurables ;

b) l'ensemble $\{T = +\infty\}$ est $\mathcal{F}_\infty$-mesurable ;

c) $\mathcal{F}_T = \{A \in \mathcal{F} : \forall n \in \overline{\mathbb{N}},\ A \bigcap \{T = n\} \in \mathcal{F}_n\}$ est une sous-tribu de $\mathcal{F}_\infty$.

2. Soient S et T deux temps d'arrêt relativement à la même filtration $(\mathcal{F}_n)_{n\geq 0}$. Montrer que les v.a.r. $S \wedge T$, $S \vee T$ et $S + T$ sont des temps d'arrêt relativement à $(\mathcal{F}_n)_{n\geq 0}$.

3. Soient $(X_n)_{n\geq 0}$ un processus à valeurs dans un espace mesurable $(E, \mathcal{E})$ et $A \in \mathcal{E}$. On définit le **temps d'entrée T_A dans A** par :

$$T_A = \inf \{n \geq 0 : X_n \in A\} \quad \text{(avec la convention } \inf \emptyset = +\infty),$$

et les **temps successifs $S_A^{(p)}$, $p \geq 1$, de retour en A** en posant :

$$S_A = S_A^{(1)} = \inf \{n \geq 1 : X_n \in A\}$$

$$\text{et pour } p \geq 1,\ S_A^{(p+1)} = \inf \{n > S_A^{(p)} : X_n \in A\}.$$

Montrer que $\{T_A = n\} = \{X_0 \in A^c, \cdots, X_{n-1} \in A^c, X_n \in A\} \in \mathcal{F}_n^X$, et que les variables aléatoires T_A et $S_A^{(p)}$, $p \geq 1$, sont des temps d'arrêt relativement à la filtration naturelle associée au processus $(X_n)_{n\geq 0}$.

Définition 2.1.4 Soit T un temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. On appelle **tribu antérieure à** T la tribu $\mathcal{F}_T$ (cf. exercice 1 de 2.1.3).

Exercices 2.1.5

1. Soit T un temps d'arrêt relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Montrer que :

a) pour tous $n \in \overline{\mathbb{N}}$ et $A \in \mathcal{F}_T$, $A \cap \{T \leq n\} \in \mathcal{F}_n$;

b) une v.a.r. Z est $\mathcal{F}_T$-mesurable si et seulement si $Z\ 1_{\{T=n\}}$ est $\mathcal{F}_n$-mesurable, pour tout $n \in \overline{\mathbb{N}}$;

c) pour toute v.a.r. intégrable X, $\mathbb{E}[X|\mathcal{F}_T] = \sum_{n\in\overline{\mathbb{N}}} 1_{\{T=n\}}\ \mathbb{E}[X|\mathcal{F}_n]$.

2. Soient S et T deux temps d'arrêt relativement à la même filtration $(\mathcal{F}_n)_{n\geq 0}$. Montrer que :

a) $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}$;

b) $S \leq T$ implique $\mathcal{F}_S \subset \mathcal{F}_T$;

c) les ensembles $\{S \leq T\}$ et $\{S > T\}$ appartiennent à $\mathcal{F}_{S\wedge T}$;

d) pour toute v.a.r. $\mathcal{F}_S$-mesurable Z, les v.a.r. $Z\ 1_{\{S<T\}}$ et $Z\ 1_{\{S\leq T\}}$ sont $\mathcal{F}_{S\wedge T}$-mesurables.

e) pour toute v.a.r. X,

$$\mathbb{E}[\mathbb{E}[X|\mathcal{F}_S]|\mathcal{F}_T] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_T]|\mathcal{F}_S] = \mathbb{E}[X|\mathcal{F}_{S\wedge T}].$$

3. Soient $(X_n)_{n\geq 0}$ un processus adapté à une filtration $(\mathcal{F}_n)_{n\geq 0}$ et T un temps d'arrêt (p.s. fini) relativement à la même filtration. On pose $X_T(\omega) = X_{T(\omega)}(\omega)$. Montrer qu'on définit ainsi une v.a.r. $\mathcal{F}_T$-mesurable.

Définition 2.1.6 Soient $X = (X_n)_{n\geq 0}$ un processus adapté à une filtration $(\mathcal{F}_n)_{n\geq 0}$ et T un temps d'arrêt relativement à la même filtration. On appelle **processus arrêté** au temps T, le processus défini par $X^T = (X_{T\wedge n})_{n\geq 0}$. D'après l'exercice précédent, ce processus est adapté à la filtration $(\mathcal{F}_{T\wedge n})_{n\geq 0}$ et donc à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

2.2 Martingales à temps discret

Définition 2.2.1 Considérons une filtration $(\mathcal{F}_n)_{n\geq 0}$ et un processus $(X_n)_{n\geq 0}$ formé de variables aléatoires réelles.

On dit que $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une **martingale**, ou encore que le processus $(X_n)_{n\geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$, si les conditions suivantes sont satisfaites :

i) $(X_n)_{n\geq 0}$ est adapté à la filtration $(\mathcal{F}_n)_{n\geq 0}$;

ii) $\mathbb{E}[|X_n|] < +\infty$, pour tout $n \geq 0$,

iii) $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$, pour tout $n \geq 0$.

La suite $(\mathbb{E}(X_n))_{n\geq 0}$ est alors constante : $\mathbb{E}[X_n] = \mathbb{E}[X_0], \ \forall n \geq 0$, et nous avons

$$\mathbb{E}[X_{n+p}|\mathcal{F}_n] = X_n, \ \forall n, p \geq 0.$$

Dans la définition précédente, lorsque la filtration $(\mathcal{F}_n)_{n\geq 0}$ est la filtration naturelle associée au processus $(X_n)_{n\geq 0}$, on dit simplement que le processus $(X_n)_{n\geq 0}$ est une martingale.

On dit que $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une **surmartingale** s'il vérifie *i)*, *ii)* et

iii') $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n, \ \forall n \geq 0$,

On dit que $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une **sous-martingale** si $(-X_n, \mathcal{F}_n)_{n\geq 0}$ est une surmartingale, c'est-à -dire si l'on a *i)*, *ii)* et

iii'') $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n, \ \forall n \geq 0$,

La suite $(\mathbb{E}(X_n))_{n\geq 0}$ est décroissante pour une surmartingale et croissante pour une sous-martingale.

Définition 2.2.2 A une martingale $(X_n, \mathcal{F}_n)_{n\geq 0}$, on associe le processus $(Y_n)_{n\geq 1}$ formé des **accroissements** : $Y_n = X_n - X_{n-1}$, $n \geq 1$. On a donc $X_n = X_0 + Y_1 + \ldots + Y_n$, $\forall n \geq 0$.

La condition *iii)* de martingale se traduit par

$$\mathbb{E}[Y_n|\mathcal{F}_{n-1}] = 0, \forall n \geq 1.$$

Inversement, étant donnés un processus $(Y_n)_{n\geq 1}$ à valeurs réelles tel que $\mathbb{E}[|Y_n|] < +\infty$, $\forall n \geq 1$, et une filtration $(\mathcal{F}_n)_{n\geq 0}$, on dit que $((Y_n)_{n\geq 1}, (\mathcal{F}_n)_{n\geq 0})$ est une **suite d'accroissements de martingale** si :

i) le processus $(Y_n)_{n\geq 1}$ est adapté à la filtration $(\mathcal{F}_n)_{n\geq 1}$;

ii) $\forall n \geq 0, \ \mathbb{E}[Y_{n+1}|\mathcal{F}_n] = 0$.

Pour toute v.a.r. $\mathcal{F}_0$-mesurable X_0, le processus (X_n) défini, pour $n \geq 1$, par $X_n = X_0 + Y_1 + \cdots + Y_n$ est alors une martingale relativement à $(\mathcal{F}_n)_{n\geq 0}$.

On montrera en exercice le lemme suivant :

Lemme 2.2.3 *Soit $(X_n, \mathcal{F}_n)_{n\geq 0}$ une martingale [resp. une surmartingale]. Soit f une fonction concave [resp. concave et croissante] sur $\mathbb{R}$, telle que les v.a.r. $f(X_n)$, $n \geq 0$, soient intégrables. Le processus $(f(X_n))_{n\geq 0}$ est alors une surmartingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.*

On obtient en particulier :

- si $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une surmartingale, $(X_n^-, \mathcal{F}_n)_{n\geq 0}$ est une sous-martingale (appliquer le lemme à $f(x) = -x^-$) ;

- si $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une sous-martingale positive ou une martingale quelconque, alors $(|X_n|, \mathcal{F}_n)_{n\geq 0}$ est une sous-martingale.

Exercices 2.2.4

1. Soient $(X_n, \mathcal{F}_n)_{n\geq 0}$ et $(Y_n, \mathcal{F}_n)_{n\geq 0}$ deux martingales [resp. deux surmartingales]. Alors $(X_n \wedge Y_n, \mathcal{F}_n)_{n\geq 0}$ est une surmartingale et, pour tous réels [resp. pour tous réels positifs] a et b, $(aX_n + bY_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale [resp. une surmartingale].

2. Soient Z une v.a.r. intégrable et $(\mathcal{F}_n)_{n\geq 0}$ une filtration. On pose

$$\forall n \geq 0, X_n = \mathbb{E}[Z|\mathcal{F}_n].$$

Montrer que $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale.

3. Soit $(Y_n)_{n\geq 1}$ une suite de v.a.r. normales centrées réduites indépendantes. On pose $X_0 = 1$, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et

$$\forall n \geq 1,\ X_n = \exp\,(Y_1 + \ldots + Y_n - \frac{n}{2}),\ \mathcal{F}_n = \sigma(Y_1, \ldots, Y_n).$$

a) Montrer que $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale bornée dans L^1.

b) Montrer que $X_n \overset{\text{p.s.}}{\to} 0$ et qu'il n'existe pas de v.a.r. intégrable X telle que $\forall n \geq 0, X_n = \mathbb{E}[X|\mathcal{F}_n]$.

c) Montrer que la famille $\{X_n : n \geq 0\}$ n'est pas équi-intégrable.

d) Retrouver ce résultat par un calcul explicite.

On montrera que, pour tout $c > 0$, on a :

$$\sup_{n\geq 0} \int_{\{X_n \geq c\}} X_n \, d\mathbb{P} = 1.$$

4. Un joueur affronte une banque au jeu de pile ou face. On modélise le jeu à l'aide d'une suite de v.a.r. $(U_n)_{n\geq 1}$ indépendantes et prenant les valeurs -1 et 1 avec probabilité $\frac{1}{2}$. La v.a.r. U_n vaut 1 si le n-ième lancer est favorable au joueur et -1 sinon.

A chaque lancer, le joueur choisit le montant de la mise en tenant compte des résultats précédents (stratégie). On peut donc modéliser la mise du n-ième lancer par une v.a.r. M_n bornée (ou plus généralement intégrable) et $\sigma(U_1, \ldots, U_{n-1})$-mesurable. Le gain du joueur G_n après le n-ième lancer s'écrit donc $G_n = \sum_{k=1}^n U_k\ M_k$. On pose $\mathcal{F}_0 = \{\emptyset, \Omega\}$ et, $\forall n \geq 1,\ \mathcal{F}_n = \sigma(U_1, \ldots, U_n)$.

a) Montrer que $(G_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale.

Le joueur adopte la stratégie suivante :
- il commence par miser a francs puis à chaque nouveau lancer, il double la mise précédente ;
- il arrête le jeu au premier lancer favorable.

On pose $T = \inf\,\{n \geq 1 : U_n = 1\}$.

b) Montrer que T est un temps d'arrêt p.s. fini.

c) Calculer le gain G_T du joueur à l'arrêt du jeu.

d) Soit F la somme totale misée par le joueur au cours de la partie. Calculer F et $\mathbb{E}[F]$ et interpréter le résultat.

• Martingales et temps d'arrêt

Lemme 2.2.5 *Si $(X_n)_{n\geq 0}$ est une martingale (resp. une surmartingale) relativement à une filtration $(\mathcal{F}_n)_{n\geq 0}$ et T un temps d'arrêt relativement à cette filtration, le processus arrêté $X^T = (X_{T\wedge n})_{n\geq 0}$ est encore une martingale (resp. une surmartingale) relativement à $(\mathcal{F}_n)_{n\geq 0}$.*

<u>Preuve</u> : Il suffit de prouver le résultat dans le cas d'une surmartingale. On part des relations

$$\forall n \geq 0,\ X_{T\wedge n} = \sum_{p=0}^{n-1} X_p 1_{\{T=p\}} + 1_{\{T\geq n\}} X_n, \tag{2.1}$$

qui montrent que, pour tout $n \geq 0$, $\mathbb{E}[|X_{T\wedge n}|] \leq \sum_{k=0}^{n} \mathbb{E}[|X_k|] < +\infty$.

En notant que les v.a.r. $1_{\{T\geq n+1\}} = 1_{\{T\leq n\}^c}$ et $\sum_{p=0}^{n} X_p 1_{\{T=p\}}$ sont $\mathcal{F}_n$-mesurables, on obtient :

$$\begin{aligned} \mathbb{E}[X_{T\wedge(n+1)}|\mathcal{F}_n] &= \sum_{p=0}^{n} X_p 1_{\{T=p\}} + 1_{\{T\geq n+1\}} \mathbb{E}[X_{n+1}|\mathcal{F}_n] \\ &\leq \sum_{p=0}^{n} X_p 1_{\{T=p\}} + 1_{\{T\geq n+1\}} X_n = X_{T\wedge n}. \end{aligned}$$

□

Théorème 2.2.6 *(d'arrêt) Soient k un entier > 0 et S et T deux temps d'arrêt relativement à une filtration $(\mathcal{F}_n)_{n\geq 0}$ tels que $S \leq T \leq k$. Si $(X_n)_{n\geq 0}$ est une surmartingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$, nous avons :*

$$\mathbb{E}[X_T|\mathcal{F}_S] \leq X_S. \tag{2.2}$$

Pour une martingale, nous avons l'égalité : $\mathbb{E}[X_T|\mathcal{F}_S] = X_S$.

<u>Preuve</u> Il suffit de prouver le résultat pour une surmartingale.

La variable aléatoire X_T est intégrable, d'après l'inégalité $\mathbb{E}[|X_T|] \leq \sum_{j=0}^{k} \mathbb{E}[|X_j|] < +\infty$. D'après le lemme 2.2.5 nous avons, pour $0 \leq p \leq k$,

$$\mathbb{E}[X_T|\mathcal{F}_p] = \mathbb{E}[X_{T\wedge k}|\mathcal{F}_p] \leq X_{T\wedge p}$$

et donc, si $B \in \mathcal{F}_p$, $\int_B X_T\, d\mathbb{P} \leq \int_B X_{T\wedge p}\, d\mathbb{P}$.

Pour tout $A \in \mathcal{F}_S$, nous avons alors, en notant que l'inégalité $S \leq T$ implique $T \wedge p = p = S$ sur $\{S = p\}$,

$$\begin{aligned} \int_A X_T\, d\mathbb{P} &= \sum_{p=0}^{k} \int_{A\cap\{S=p\}} X_T\, d\mathbb{P} \leq \sum_{p=0}^{k} \int_{A\cap\{S=p\}} X_{T\wedge p}\, d\mathbb{P} \\ &= \sum_{p=0}^{k} \int_{A\cap\{S=p\}} X_S\, d\mathbb{P} = \int_A X_S\, d\mathbb{P}. \end{aligned}$$

□

• **Inégalités maximales**

Théorème 2.2.7 *(Inégalités de Doob) Soit $(X_n)_{n\geq 0}$ une surmartingale et λ un réel > 0. Pour tout entier k, nous avons :*

$$\begin{aligned}
\lambda\,\mathbb{P}[\sup_{0\leq p\leq k} X_p \geq \lambda] &\leq \mathbb{E}[X_0] - \int_{\{\sup_{0\leq p\leq k} X_p<\lambda\}} X_k \; d\mathbb{P} && (2.3)\\
&\leq \mathbb{E}[X_0] + \mathbb{E}[X_k^-], && (2.4)
\end{aligned}$$

$$\lambda\,\mathbb{P}[\inf_{0\leq p\leq k} X_p \leq -\lambda] \leq \int_{\{\inf_{0\leq p\leq k} X_p\leq-\lambda\}} (-X_k)\; d\mathbb{P} \leq \mathbb{E}[X_k^-]. \qquad (2.5)$$

En particulier, si (X_n) est une martingale, pour tout λ réel > 0 et tout entier k, nous avons :

$$\lambda\,\mathbb{P}[\sup_{0\leq p\leq k} X_p \geq \lambda] \leq \mathbb{E}[X_k^+].$$

Preuve : Pour établir la première inégalité, considérons le temps d'arrêt T défini par

$$T = \begin{cases} \inf\{0\leq p\leq k : X_p \geq \lambda\}, & \text{si cet ensemble est non vide,}\\ k, & \text{sinon.}\end{cases}$$

D'après le théorème d'arrêt 2.2.6, nous avons : $\mathbb{E}[X_T|\mathcal{F}_0] \leq X_0$, et par suite

$$\begin{aligned}
\mathbb{E}[X_0] &\geq \mathbb{E}[X_T] = \mathbb{E}[X_T 1_{\{\sup_{0\leq p\leq k} X_p\geq\lambda\}}] + \mathbb{E}[X_T 1_{\{\sup_{0\leq p\leq k} X_p<\lambda\}}]\\
&\geq \lambda\mathbb{P}[\sup_{0\leq p\leq k} X_p \geq \lambda] + \mathbb{E}[X_k 1_{\{\sup_{0\leq p\leq k} X_p<\lambda\}}].
\end{aligned}$$

En considérant le temps d'arrêt

$$T = \begin{cases} \inf\{0\leq p\leq k : X_p \leq -\lambda\}, & \text{si cet ensemble est non vide,}\\ k, & \text{sinon.}\end{cases}$$

nous obtenons, de même :

$$\begin{aligned}
\mathbb{E}[X_k] \leq \mathbb{E}[X_T] &= \mathbb{E}[X_T 1_{\{\inf_{0\leq p\leq k} X_p\leq-\lambda\}}] + \mathbb{E}[X_T 1_{\{\inf_{0\leq p\leq k} X_p>-\lambda\}}]\\
&\leq -\lambda\mathbb{P}[\inf_{0\leq p\leq k} X_p \leq -\lambda] + \mathbb{E}[X_k 1_{\{\inf_{0\leq p\leq k} X_p>-\lambda\}}];
\end{aligned}$$

d'où :

$$\begin{aligned}
\lambda\,\mathbb{P}[\inf_{0\leq p\leq k} X_p \leq -\lambda] &\leq \mathbb{E}[X_k 1_{\{\inf_{0\leq p\leq k} X_p>-\lambda\}}] - \mathbb{E}[X_k]\\
&\leq \int_{\{\inf_{0\leq p\leq k} X_p\leq-\lambda\}} (-X_k)\; d\mathbb{P} \leq \mathbb{E}[X_k^-].
\end{aligned}$$

□

Corollaire 2.2.8 *Soit $(X_n)_{n\geq 0}$ une sous-martingale positive, ou une martingale, telle que, pour un entier $n_1 \geq 1$, on ait $X_{n_1} \in L^p$, $p \geq 1$. Nous avons alors, pour tout réel $\lambda > 0$,*

$$\mathbb{P}[\sup_{0\leq k\leq n_1} |X_k| \geq \lambda] \leq \frac{1}{\lambda^p}\|X_{n_1}\|_p^p. \qquad (2.6)$$

Preuve : On applique l'inégalité (2.5) à la surmartingale $(-Y_n)$ définie par $Y_n = |X_{n\wedge n_1}|^p$, $n \geq 0$.

□

Corollaire 2.2.9 *(Inégalité de Kolmogorov) Soit $(Y_i)_{i\geq 1}$ une suite de v.a.r. indépendantes, centrées et de carré intégrable. Nous avons, pour tout réel $\lambda > 0$,*

$$\mathbb{P}[\sup_{0\leq k\leq n} |Y_1 + \cdots + Y_k| \geq \lambda] \leq \frac{1}{\lambda^2}\sum_{n=1}^{n} \mathbb{E}[Y_k^2].$$

En reprenant la démonstration du théorème 2.2.6, on montre (exercice) :

Proposition 2.2.10 *(lemme maximal pour les surmartingales positives) Si $(X_n)_{n\geq 0}$ est une surmartingale positive, nous avons, pour tout réel $\lambda > 0$, l'inégalité :*

$$\mathbb{P}[\sup_n X_n \geq \lambda|\mathcal{F}_0] \leq \inf\{\frac{X_0}{\lambda}, 1\}.$$

Donnons maintenant une application des inégalités maximales.

Théorème 2.2.11 *Soient p un réel > 1 et $(X_n)_{n\geq 0}$ une martingale ou une sous-martingale positive bornée dans L^p. Alors la v.a.r. $\sup_{n\geq 0} |X_n|$ appartient à L^p et vérifie :*

$$\|\sup_{n\geq 0} |X_n|\|_p \;\leq\; \frac{p}{p-1} \sup_{n\geq 0} \|X_n\|_p.$$

<u>Preuve</u> : Soit N un entier ≥ 1. Posons $Z_N = \sup_{0\leq n\leq N} |X_n|$.

L'inégalité (2.5) du théorème 2.2.7, appliquée à la surmartingale $(-|X_{n\wedge N}|)_{n\geq 0}$, nous donne :

$$\lambda\, \mathbb{P}[Z_N \geq \lambda] \leq \mathbb{E}[|X_N| 1_{\{Z_N\geq\lambda\}}],\ \forall \lambda > 0.$$

Nous avons :

$$\begin{aligned}
\mathbb{E}[Z_N^p] &= p\,\mathbb{E}[\int_0^{Z_N} x^{p-1}\,dx] \\
&= p\int_0^{+\infty} x^{p-1}\,\mathbb{P}[Z_N \geq x]\,dx \quad \text{(par le théorème de Fubini)} \\
&\leq p\int_0^{+\infty} x^{p-2}\,(\int_{\{Z_N\geq x\}} |X_N|\,d\mathbb{P})\,dx \\
&= p\,\mathbb{E}[|X_N|\int_0^{Z_N} x^{p-2}\,dx] \text{ (Fubini)} \\
&= \frac{p}{p-1}\,\mathbb{E}[|X_N|\; Z_N^{p-1}] \\
&\leq \frac{p}{p-1}\,\|X_N\|_p\,(\|Z_N\|_p)^{p-1} \quad \text{(par l'inégalité de Hölder).}
\end{aligned}$$

Il en résulte :

$$\|Z_N\|_p \leq \frac{p}{p-1}\,\|X_N\|_p \leq \frac{p}{p-1}\,\sup_{n\geq 0}\|X_n\|_p.$$

Le résultat s'obtient en faisant tendre N vers $+\infty$. □

2.3 Théorèmes de convergence pour les martingales

• Convergence des martingales bornées dans L^1

Nous commençons par définir le nombre de montées d'une suite de réels à travers un intervalle.

Soit $x = (x_n)_{n\geq 0}$ une suite de réels. Etant donnés deux réels a et b, $a < b$, considérons les entiers $s_1 < t_1 < \cdots < s_m < t_m < \cdots$ définis par :

$$\begin{array}{rcl} s_1 & = & \inf\{n \geq 0 : x_n \leq a\} \\ t_1 & = & \inf\{n > s_1 : x_n \geq b\} \\ & \vdots & \\ s_m & = & \inf\{n > t_{m-1} : x_n \leq a\} \\ t_m & = & \inf\{n > s_m : x_n \geq b\}, \\ & \vdots & \end{array}$$

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

Définition 2.3.1 On appelle **nombre de montées**, $M_a^b(x)$, de la suite $x = (x_n)_{n\geq 0}$ à travers $[a, b]$ le plus grand entier m tel que $t_m < \infty$.

C'est le nombre de traversées montantes de l'intervalle $[a, b]$ par la suite x. Nous utiliserons le lemme élémentaire suivant :

Lemme 2.3.2 *Une suite $x = (x_n)_{n\geq 0}$ converge dans $\overline{\mathbb{R}}$ si, et seulement si, $M_a^b(x) < +\infty$, pour tout couple (a, b) de rationnels vérifiant $a < b$.*

Soit $X = (X_n)_{n\geq 0}$ un processus stochastique. Pour chaque $\omega \in \Omega$, considérons le nombre de montées $M_a^b((X_n(\omega)))$ de la suite $(X_n(\omega))_{n\geq 0}$ à travers l'intervalle $[a, b]$. On note $M_a^b(X)$ la variable aléatoire définie par $M_a^b(X) : \omega \to M_a^b((X_n(\omega)))$ (la preuve du théorème qui suit montre que c'est bien une variable aléatoire).

Théorème 2.3.3 *(Inégalité fondamentale sur les montées) Si $X = (X_n)_{n\geq 0}$ est une surmartingale, pour tous réels a et b vérifiant $a < b$, nous avons :*

$$\mathbb{E}[M_a^b(X)] = \frac{1}{b-a} \sup_{k\geq 0} \mathbb{E}[(X_k - a)^-].$$

<u>Preuve</u> : Considérons les temps d'arrêt relativement à la filtration $(\mathcal{F}_n^X)_{n\geq 0}$ définis par

$$\begin{array}{rcl} S_1 & = & \inf\{n \geq 0 : X_n \leq a\} \\ T_1 & = & \inf\{n > S_1 : X_n \geq b\} \\ & \vdots & \\ S_m & = & \inf\{n > T_{m-1} : X_n \leq a\} \\ T_m & = & \inf\{n > S_m : X_n \geq b\}, \\ & \vdots & \end{array}$$

On a, par définition,

$$M_a^b(X(\omega)) = \sup\{m \geq 0 : T_m(\omega) < +\infty\},\ \omega \in \Omega.$$

Soient k un entier ≥ 1 et X^k le processus obtenu en arrêtant le processus X en k. Nous avons :

$$\mathbb{E}[M_a^b(X^k)] = \sum_{n\geq 1} \mathbb{P}[M_a^b(X^k) \geq n] = \sum_{n\geq 1} \mathbb{P}[T_n \leq k].$$

et d'autre part :

$$\begin{array}{rcl} (b-a)\mathbb{P}[T_n \leq k] & \leq & \mathbb{E}[(X^k_{T_n} - X^k_{S_n})1_{\{T_n\leq k\}}] \\ & \leq & \mathbb{E}[(X_{T_n\wedge k} - X_{S_n\wedge k})1_{\{S_n\leq k\}})] \\ & & \quad -\mathbb{E}[(X^k_{T_n} - X^k_{S_n})1_{\{S_n\leq k\}\cap\{T_n\geq k+1\}}], \end{array}$$

car $\{S_n \leq k\} = \{T_n \leq k\} \cup (\{S_n \leq k\} \cap \{T_n \geq k+1\})$.

Comme, d'après le théorème d'arrêt,

$$\mathbb{E}[(X_{T_n\wedge k} - X_{S_n\wedge k})|\mathcal{F}_{S_n\wedge k}] \leq 0,$$

on en déduit :

$$(b-a)\mathbb{P}[T_n \leq k] \leq \mathbb{E}[(a - X_k)1_{\{S_n\leq k\}\cap\{T_n\geq k+1\}}].$$

Les évènements $A_n = \{S_n \leq k\} \cap \{T_n \geq k+1\}$ étant deux à deux disjoints, nous avons, $\forall k \geq 1$,

$$\begin{array}{rcl} (b-a)\mathbb{E}[M_a^b(X^k)] & \leq & -\mathbb{E}[(X_k - a)1_{\cup_{n\geq 1} A_n}] \\ & \leq & \mathbb{E}[(X_k - a)^-]; \end{array}$$

d'où l'on déduit aisément le résultat. □

Théorème 2.3.4 *(Convergence des surmartingales) Soit $(X_n)_{n\geq 0}$ une surmartingale telle que $\sup_n \mathbb{E}[|X_n|] < +\infty$. Alors la suite $(X_n)_{n\geq 0}$ converge p.s. vers une variable aléatoire X_∞ intégrable.*

<u>Preuve</u> : Notons d'abord que la condition $\sup_n \mathbb{E}[|X_n|] < +\infty$, pour une surmartingale, équivaut à $\sup_n \mathbb{E}[X_n^-] = \lim_n \uparrow \mathbb{E}[X_n^-] < +\infty$. En effet, en écrivant $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] + 2\mathbb{E}[X_n^-]$, on obtient :

$$\sup_n \mathbb{E}[|X_n|] \leq \mathbb{E}[X_0] + 2\sup_n \mathbb{E}[X_n^-].$$

D'après le théorème précédent, nous avons :

$$\mathbb{E}[M_a^b(X)] \leq \frac{1}{b-a}[\sup_{k\geq 1} \mathbb{E}[X_k^-] + |a|] < +\infty,$$

et, par suite, $M_a^b(X) < +\infty$, $\mathbb{P}$-p.s., pour tout couple (a,b) de rationnels vérifiant $a < b$. Il existe donc un sous-ensemble Ω_0 mesurable de Ω de $\mathbb{P}$-probabilité 1, tel que, pour tout $\omega \in \Omega_0$, $M_a^b((X_n(\omega))) < +\infty$, pour tout couple (a,b) de rationnels vérifiant $a < b$. Ceci entraîne la convergence dans $\overline{\mathbb{R}}$ de la suite $(X_n(\omega))_{n\geq 0}$, pour tout $\omega \in \Omega_0$.

On conclut par le lemme de Fatou :

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf_n |X_n|] \leq \liminf_n \mathbb{E}[|X_n|] \leq \sup_n \mathbb{E}[|X_n|] < +\infty.$$

□

Remarque 2.3.5 Le théorème précédent peut être généralisé de la façon suivante : on étend la notion de surmartingale en remplaçant la condition *ii*) de la définition 2.2.1 par l'hypothèse plus faible *ii'*) suivante :
ii') $\mathbb{E}[X_n^-] < +\infty, \ \forall n \geq 0$.

Théorème 2.3.6 *Soit* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *une surmartingale telle que* $\sup \mathbb{E}[X_n^-] < +\infty$. *Alors la suite* (X_n) *converge p.s. vers une variable aléatoire* X_∞ *telle que* X_∞^- *soit intégrable. Si de plus* X_0 *est intégrable, nous avons* $\sup_n \mathbb{E}[|X_n|] < +\infty$ *et* X_∞ *est intégrable.*

On peut se demander si, sous les hypothèses du théorème (2.3.4), on a $\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n, \ \forall n \geq 0$. Remarquons que, si cette dernière condition est vérifiée, on a :

$$\forall n, \ X_n^- \leq \mathbb{E}[X_\infty^-|\mathcal{F}_n];$$

ce qui implique que la famille $\{X_n^-, n \geq 0\}$ est équi-intégrable. Cette condition nécessaire est en fait suffisante :

Théorème 2.3.7 *Soit* $(X_n, \mathcal{F}_n)_{n\geq 0}$ *une surmartingale bornée dans* L^1. *Notons* X_∞ *la limite* $\mathbb{P}$*-p.s. de la suite* $(X_n)_{n\geq 0}$ *(cf. le théorème (2.3.4). Nous avons*

$$\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n, \ \forall n \geq 0, \tag{2.7}$$

si et seulement si la famille $\{X_n^-, n \geq 0\}$ *est équi-intégrable.*

<u>Preuve</u> : Nous avons vu que (2.7) entraîne l'équi-intégrabilité de la famille $\{X_n^-, n \geq 0\}$.

Inversement, nous avons, pour tous entiers positifs p et n :

$$X_n \geq \mathbb{E}[X_{n+p}|\mathcal{F}_n] = \mathbb{E}[X_{n+p}^+|\mathcal{F}_n] - \mathbb{E}[X_{n+p}^-|\mathcal{F}_n]. \tag{2.8}$$

Sous l'hypothèse d'équi-intégrabilité la suite de v.a. $(X_n^-)_{n\geq 0}$ converge dans L^1 vers X_∞^-. Il en résulte la convergence en norme L^1 :

$$\lim_{p\to+\infty} \mathbb{E}[X_{n+p}^-|\mathcal{F}_n] = \mathbb{E}[X_\infty^-|\mathcal{F}_n],$$

et nous pouvons extraire une sous-suite $(\varphi(p))_{p\geq 1}$, telle qu'il y ait convergence presque-sŨre de la sous-suite $(\mathbb{E}[X_{n+\varphi(p)}^-|\mathcal{F}_n])_{p\geq 1}$ vers $\mathbb{E}[X_\infty^-|\mathcal{F}_n]$.

D'autre part, d'après le lemme de Fatou, nous avons :

$$\begin{aligned} \mathbb{E}[X_\infty^+|\mathcal{F}_n] &= \mathbb{E}[\liminf_p X_{n+\varphi(p)}^+|\mathcal{F}_n] \\ &\leq \liminf_p \mathbb{E}[X_{n+\varphi(p)}^+|\mathcal{F}_n]. \end{aligned}$$

Le résultat s'en déduit en passant à la limite inférieure dans (2.8) sur la sous-suite $(\varphi(p))_{p\geq 1}$. □

Corollaire 2.3.8 *Une martingale* $(X_n)_{n\geq 0}$ *est équi-intégrable si et seulement s'il existe une variable aléatoire* Z *intégrable telle que* $X_n = \mathbb{E}[Z|\mathcal{F}_n], \ \forall n \geq 0$.

Dans ce cas, on a, pour la convergence p.s. et en norme L^1 *:*

$$\lim_n \mathbb{E}[Z|\mathcal{F}_n] = \mathbb{E}[Z|\mathcal{F}_\infty].$$

Corollaire 2.3.9 *(cas des surmartingales positives) Toute surmartingale positive* (X_n) *converge p.s. vers une v.a.* X_∞ *et, pour tout* $n \geq 0$*, nous avons :*

$$\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n.$$

Preuve : La convergence résulte du théorème 2.3.7 (car $X_n^- = 0$). En appliquant le lemme de Fatou, on obtient :

$$\mathbb{E}[X_\infty|\mathcal{F}_n] = \mathbb{E}[\liminf_p X_p|\mathcal{F}_n] \leq \liminf_p \mathbb{E}[X_p|\mathcal{F}_n] \leq X_n.$$

□

Plus généralement, le théorème 2.3.9 est vrai pour toute surmartingale $(X_n)_{n\geq 0}$ telle que $X_n^- \leq U, \forall n \geq 0$, avec U intégrable.

2.4 Martingales renversées

Nous considérons dans ce paragraphe une **surmartingale** $(X_n, \mathcal{F}_n)_{n\in\mathbb{Z}}$, indexée par $\mathbb{Z}$. La définition est celle de 2.2.1 où l'on remplace l'ensemble d'indice $\mathbb{N}$ par $\mathbb{Z}$.

On suppose donc X_n intégrable et $\mathcal{F}_n$-mesurable, $\forall n \in \mathbb{Z}$, et

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n, \ \forall n \in \mathbb{Z}, \tag{2.9}$$

On a donc :

$$\begin{aligned} \mathbb{E}[X_m|\mathcal{F}_n] &\leq X_n, \ \forall n \leq m, & (2.10)\\ X_n^- &\leq \mathbb{E}[X_m^-|\mathcal{F}_n], \ \forall n \leq m. & (2.11) \end{aligned}$$

Ces inégalités montrent que les suites $(\mathbb{E}[X_n])_{n\in\mathbb{Z}}$ et $(\mathbb{E}[X_n^-]_{n\in\mathbb{Z}})$ sont respectivement décroissante et croissante et que $(X_n^-)_{n\in\mathbb{Z}}$ est une sous-martingale.

Une surmartingale indexée par un intervalle fini d'entiers ou par $\mathbb{N}$ peut toujours être étendue en une surmartingale indexée par $\mathbb{Z}$.

Si $(\mathcal{F}_n)_{n\in\mathbb{Z}}$ est une filtration indexée par $\mathbb{Z}$ on notera

$$\mathcal{F}_{-\infty} = \bigcap_{n\geq 0} \mathcal{F}_{-n}.$$

Notion de martingale renversée

Si nous notons $M_n = X_{-n}$, $\mathcal{G}_n = \mathcal{F}_{-n}$, alors $(\mathcal{G}_n)_{n\geq 0}$ est une filtration décroissante et le processus $(M_n)_{n\geq 0}$ vérifie les conditions suivantes :

i) $(M_n)_{n\geq 0}$ est adapté à la filtration décroissante $(\mathcal{G}_n)_{n\geq 0}$;

ii) $\mathbb{E}[|M_n|] < +\infty$, pour tout $n \geq 0$,

iii) $\mathbb{E}[M_n|\mathcal{G}_{n+1}] \leq M_{n+1}$, pour tout $n \geq 0$.

Un processus $(M_n)_{n\geq 0}$ vérifiant ces conditions est appelé **surmartingale renversée** relativement à la filtration décroissante $(\mathcal{G}_n)_{n\geq 0}$. S'il y a égalité dans *iii*), on dit qu'on a une **martingale renversée**.

Exemple : Soit $(\mathcal{B}_n)_{n\geq 0}$ une suite décroissante de sous-tribus de F et Y une v.a.r. intégrable. on obtient une martingale $(X_n, \mathcal{F}_n)_{n\in\mathbb{Z}}$ indexée par $\mathbb{Z}$, en posant

$$\begin{aligned} X_n &= \mathbb{E}[Y|\mathcal{B}_{-n}],\ \mathcal{F}_n = \mathcal{B}_{-n},\ \text{pour } n < 0, \\ X_n &= \mathbb{E}[Y|\mathcal{B}_0],\ \mathcal{F}_n = \mathcal{B}_0,\ \text{pour } n \geq 0. \end{aligned}$$

Nous avons étudié le comportement des surmartingales lorsque n tend vers $+\infty$. Nous nous intéressons maintenant au comportement quand n tend vers $-\infty$.

Théorème 2.4.1 *Si $(X_n, \mathcal{F}_n)_{n\in\mathbb{Z}}$ est une surmartingale, alors*
a) pour tout entier relatif p, la famille $\{X_n^-, n \leq p\}$ est équi-intégrable ;
b) quand n tend vers $-\infty$ la suite $(X_n)_{n\in\mathbb{Z}}$ converge p.s. vers une v.a.r. $X_{-\infty}$ telle que $\mathbb{E}[X_n|\mathcal{F}_{-\infty}] \leq X_{-\infty}$;
c) si on suppose en outre que $\sup_n \mathbb{E}[X_n] = \lim_{n\to-\infty} \uparrow \mathbb{E}[X_n] < +\infty$, pour tout entier relatif p, la famille $\{X_n, n \leq p\}$ est équi-intégrable et la convergence $\lim_{n\to-\infty} X_n = X_{-\infty}$ a lieu aussi en norme L^1.

<u>Preuve</u> : a) Les inégalités (2.11), avec $m = p$, montrent que la famille $\{X_n^-, n \leq p\}$ est équi-intégrable.

b) Soit k un entier ≥ 0. Considérons la surmartingale (dépendant de k) $(Y_n, \mathcal{G}_n)_{n\geq 0}$ définie par :

$$\begin{aligned} &Y_0 = X_{-k}, \qquad Y_1 = X_{-k+1}, \cdots,\ Y_k = X_0 \text{ et } Y_m = X_0, \forall m > k, \\ &\mathcal{G}_0 = \mathcal{F}_{-k}, \qquad \mathcal{G}_1 = \mathcal{F}_{-k+1}, \cdots,\ \mathcal{G}_k = \mathcal{F}_0 \text{ et } \mathcal{G}_m = \mathcal{F}_0, \forall m > k. \end{aligned}$$

Soit $[a, b]$ un intervalle réel quelconque. D'après le théorème sur les montées d'une surmartingale, nous avons :

$$\begin{aligned} \mathbb{E}[M_a^b(Y)] &\leq \frac{1}{b-a}[\sup_{m\geq 0} \mathbb{E}[Y_m^-] + |a|] \\ &\leq \frac{1}{b-a}[\mathbb{E}[X_0^-] + |a|]. \end{aligned}$$

En observant que le nombre de montées de la suite $(X_{-k}, X_{-k+1}, \cdots, X_0, X_0, \cdots)$ et celui de la suite $(X_0, X_{-1}, \cdots, X_{-k}, X_{-k}, \cdots)$ diffèrent d'au plus une unité, on obtient, en faisant croître k vers $+\infty$,

$$\mathbb{E}[M_a^b((X_m)_{m\leq 0})] < +\infty,$$

pour tout intervalle $[a, b]$. Ceci montre la convergence p.s. de la suite (X_n), quand n tend vers $-\infty$, vers une v.a.r. que nous noterons $X_{-\infty}$.

c) L'égalité $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] + 2\mathbb{E}[X_n^-]$ montre que la suite $(\mathbb{E}[|X_n|])$ converge, quand n tend vers $-\infty$ vers

$$\lim_{n\to-\infty} \uparrow \mathbb{E}[X_n] + 2 \lim_{n\to-\infty} \downarrow \mathbb{E}[X_n^-] \leq \sup_n \mathbb{E}[X_n] + 2\mathbb{E}[X_0^-].$$

Lorsque $\sup_n \mathbb{E}[X_n] = \lim_{n\to-\infty} \uparrow \mathbb{E}[X_n]$ est fini, la suite $(\mathbb{E}[|X_n|])$ converge vers une limite finie, quand n tend vers $-\infty$.

Soient $\varepsilon > 0$ et $p \in \mathbb{Z}$. Choisissons un entier $k = k(\varepsilon) \leq p$ tel que, $\forall n \leq k$, on ait :

$$0 \leq \mathbb{E}[X_n] - \mathbb{E}[X_k] < \varepsilon. \tag{2.12}$$

Pour $c \geq 0$, l'intégrale $\int_{\{|X_n|>c\}} |X_n| \, d\mathbb{P}$ s'écrit

$$-\int_{\{X_n<-c\}} X_n \, d\mathbb{P} + \mathbb{E}[X_n] - \int_{\{X_n\leq c\}} X_n \, d\mathbb{P}.$$

D'après l'inégalité de surmartingales et (2.12), cette expression est majorée par

$$-\int_{\{X_n<-c\}} X_k \; d\mathbb{P} + (\mathbb{E}[X_k]+\varepsilon) - \int_{\{X_n\le c\}} X_k \; d\mathbb{P}.$$

Nous obtenons donc, pour $n \le k = k(\varepsilon)$,

$$\int_{\{|X_n|>c\}} |X_n| \; d\mathbb{P} \le \int_{\{|X_n|>c\}} |X_k| \; d\mathbb{P} + \varepsilon. \tag{2.13}$$

La suite $(\mathbb{E}[|X_n|])$ convergeant vers une limite finie, quand n tend vers $-\infty$, la majoration

$$\sup_{n\le p} \mathbb{P}[\{|X_n|>c\}] \le \frac{1}{c} \sup_{n\le p} \mathbb{E}[|X_n|],$$

implique que

$$\lim_{c\to+\infty} \downarrow \sup_{n\le p} \int_{\{|X_n|>c|\}} |X_k| \; d\mathbb{P} = 0.$$

D'après la majoration (2.13), on a donc,

$$\begin{aligned}\sup_{n\le p} \int_{\{|X_n|>c|\}} |X_n| \; d\mathbb{P} \le \quad & \sup_{n\le k(\varepsilon)} \int_{\{|X_n|>c|\}} |X_k| \; d\mathbb{P} \\ & +\varepsilon + \sup_{k(\varepsilon)\le n\le p} \int_{\{|X_n|>c|\}} |X_n| \; d\mathbb{P},\end{aligned}$$

d'où, pour tout $\varepsilon > 0$:

$$\lim_{c\to+\infty} \sup_{n\le p} \int_{\{|X_n|>c|\}} |X_n| \; d\mathbb{P} \le \varepsilon,$$

ce qui établit l'équi-intégrabilité de la famille $(X_n, n \le p)$.

Nous avons alors la convergence de la suite de v.a.r. $(X_n)_{n\ge 0}$ vers $X_{-\infty}$ p.s. et en norme L^1, quand n tend vers $-\infty$.

Supposons toujours qu'on ait $\sup_n \mathbb{E}[X_n] < +\infty$. Pour $A \in \mathcal{F}_{-\infty}$, nous avons, d'après l'inégalité de surmartingale :

$$\int_A X_n \; d\mathbb{P} \le \int_A X_m \; d\mathbb{P}, \;\; \forall m \le n, m, n \in \mathbb{Z},$$

ce qui entraîne, en faisant tendre m vers $-\infty$,

$$\int_A X_n \; d\mathbb{P} \le \int_A X_{-\infty} \; d\mathbb{P}, \;\; \forall n \in \mathbb{Z};$$

d'où l'inégalité

$$\mathbb{E}[X_n|\mathcal{F}_{-\infty}] \le X_{-\infty}, \;\; \forall n \in \mathbb{Z}.$$

d) Revenons maintenant à la seconde assertion du théorème. Le processus $(X_n \wedge a, n \le p)$ est, pour tout réel $a > 0$ une surmartingale relativement à la filtration $(\mathcal{F}_n)_{n\ge 0}$. En outre, cette surmartingale est équi-intégrable, car nous avons vu que $\{X_n^-, n \le p\}$ est équi-intégrable. D'après la deuxième assertion du théorème, on a alors :

$$\mathbb{E}[X_n \wedge a|\mathcal{F}_{-\infty}] \le X_{-\infty} \wedge a, \;\; \forall n \in \mathbb{Z}$$

Lorsque a tend vers $+\infty$, $X_n \wedge a$ converge en croisant vers X_n. Comme $\mathbb{E}[|X_n|] < +\infty$, pour tout $n \in \mathbb{Z}$, on a :

$$\mathbb{E}[X_n \wedge a|\mathcal{F}_{-\infty}] \uparrow \mathbb{E}[X_n|\mathcal{F}_{-\infty}], \;\; \forall n \in \mathbb{Z}.$$

Par suite, on a :

$$\mathbb{E}[X_n|\mathcal{F}_{-\infty}] \leq X_{-\infty},\ \forall n \in \mathbb{Z}.$$

□

(4.4) Corollaire : *Si $(X_n, \mathcal{F}_n)_{n\in\mathbb{Z}}$ est une martingale, pour tout entier relatif p, la famille de v.a.r. $\{X_n, n \leq p\}$ est équi-intégrable, la suite $(X_n)_{n\in\mathbb{Z}}$ converge p.s. et dans L^1 vers une v.a.r. $X_{-\infty}$, quand n tend vers $-\infty$, et on a :*

$$\mathbb{E}[X_n|\mathcal{F}_{-\infty}] = X_{-\infty},\ \forall n \in \mathbb{Z}.$$

Dans les corollaires qui suivent, on considère une suite **décroissante** de tribus $(\mathcal{B}_n)_{n\geq 0}$. On note $\mathcal{B}_\infty = \bigcap_n \mathcal{B}_n$.

Corollaire 2.4.2 *Si $(Z_n)_{n\geq 0}$ est une surmartingale renversée relativement à la filtration décroissante $(\mathcal{B}_n)_{n\geq 0}$, la famille $\{Z_n^-, n \geq 0\}$ est équi-intégrable et la suite $(Z_n)_{n\geq 0}$ converge p.s. vers une v.a.r. Z_∞ vérifiant*

$$\mathbb{E}[Z_n|\mathcal{B}_\infty] \leq Z_\infty,\ \forall n \geq 0.$$

En particulier, la suite $(Z_n^-)_{n\geq 0}$ converge vers Z_∞^-, p.s. et dans L^1.

Si, en outre, $\sup_{n\geq 0} \mathbb{E}[Z_n] = \lim_n \uparrow \mathbb{E}[Z_n] < +\infty$, alors la convergence de $(Z_n)_{n\geq 0}$ vers Z_∞ a lieu p.s. et dans L^1.

Corollaire 2.4.3 *Si $(Z_n)_{n\geq 0}$ est une martingale renversée relativement à la filtration décroissante $(\mathcal{B}_n)_{n\geq 0}$,la famille $\{Z_n, n \geq 0\}$ est équi-intégrable et la suite (Z_n) converge p.s. et dans L^1 vers une v.a.r. Z_∞ vérifiant*

$$\mathbb{E}[Z_n|\mathcal{B}_\infty] = Z_\infty,\ \forall n \geq 0.$$

Corollaire 2.4.4 *Soit $(\mathcal{B}_n)_{n\geq 0}$ une suite décroissante de sous-tribus de F et Y une v.a.r. intégrable. Alors la suite $(\mathbb{E}[Y|\mathcal{B}_n])_{n\geq 0}$ converge p.s. et en norme L^1, vers $\mathbb{E}[Y|\mathcal{B}_\infty]$.*

Preuve : Le corollaire précédent montre la convergence p.s. et dans L^1 de la suite $(Z_n = \mathbb{E}[Y|\mathcal{B}_n])_{n\geq 0}$ vers une v.a.r. Z_∞ vérifiant

$$\mathbb{E}[Z_n|\mathcal{B}_\infty] = Z_\infty,\ \forall n \geq 0.$$

Comme $\mathcal{B}_\infty \subset \mathcal{B}_n$, on en déduit :

$$Z_\infty = \mathbb{E}[\mathbb{E}[Y|\mathcal{B}_n]|\mathcal{B}_\infty] = \mathbb{E}[Y|\mathcal{B}_\infty].$$

□

2.5 Loi des grands nombres et accroissements de martingale

Considérons une suite d'accroissements de martingale $((Y_n)_{n\geq 1}, (\mathcal{F}_n)_{n\geq 0})$ (cf. définition 2.2.2).

Pour toute suite de nombres réels non nuls $(u_n)_{n\geq 1}$, le processus $(X_n)_{n\geq 0}$, défini par

$$X_0 = 0 \quad \text{et} \quad \forall n \geq 1,\ X_n = \frac{Y_1}{u_1} + \cdots + \frac{Y_n}{u_n},$$

est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

Rappelons le lemme de Kronecker :

Lemme 2.5.1 *Soit $(u_n)_{n\geq 1}$ une suite croissante de nombres réels strictement positifs telle que $\lim_{n\to+\infty} u_n = +\infty$. Soit $(v_n)_{n\geq 1}$ une suite réelle.*

Si la série $\sum \frac{v_n}{u_n}$ converge, alors la suite $(\frac{1}{u_n} \sum_{k=1}^{n} v_k)_{n\geq 1}$ converge vers zéro.

Théorème 2.5.2 *Soit $((Y_n)_{n\geq 1}, (\mathcal{F}_n)_{n\geq 0})$ une suite d'accroissements de martingale. On suppose que les v.a.r. Y_n, $n \geq 1$, sont identiquement distribuées et que $\mathbb{E}[|Y_1| \operatorname{Log}^+|Y_1|] < +\infty$. Alors la suite de v.a.r.*

$$(\frac{1}{n} (Y_1 + \cdots + Y_n))_{n\geq 1}$$

converge, $\mathbb{P}$-p.s., vers zéro.

<u>Preuve</u> : Pour tout entier $n \geq 1$, nous écrivons : $Y_n = Z_n + T_n$ avec

$$\begin{aligned} Z_n &= Y_n 1_{\{|Y_n|\leq n\}} - \mathbb{E}[Y_n 1_{\{|Y_n|\leq n\}} | \mathcal{F}_{n-1}] \\ T_n &= Y_n 1_{\{|Y_n|> n\}} - \mathbb{E}[Y_n 1_{\{|Y_n|> n\}} | \mathcal{F}_{n-1}]. \end{aligned}$$

Les suites $(Z_n)_{n\geq 1}$ et $(T_n)_{n\geq 1}$ sont des suites d'accroissements de martingale, relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Nous posons :

$$X_0 = 0, \quad M_0 = 0, \quad \forall n \geq 1, \quad X_n = \sum_{k=1}^{n} \frac{Z_k}{k} \ \text{ et } M_n = \sum_{k=1}^{n} \frac{T_k}{k}.$$

Les processus $(X_n)_{n\geq 0}$ et $(M_n)_{n\geq 0}$ sont des martingales relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Il nous suffit de montrer que la martingale $(X_n)_{n\geq 0}$ est bornée dans L^2 et que la martingale $(M_n)_{n\geq 0}$ est bornée dans L^1. D'après le théorème fondamental de convergence, les martingales $(X_n)_{n\geq 0}$ et $(M_n)_{n\geq 0}$ seront alors convergentes $\mathbb{P}$-p.s. et le théorème résultera du lemme de Kronecker.

Nous avons :

$$\begin{aligned} \sum_{k=1}^{n} \frac{1}{k^2} \mathbb{E}[Z_k^2] &= \sum_{k=1}^{n} \frac{1}{k^2} \left(\mathbb{E}[Y_k^2 1_{\{|Y_k|\leq k\}}] - \mathbb{E}[(\mathbb{E}[Y_k 1_{\{|Y_k|\leq k\}} | \mathcal{F}_{k-1}])^2] \right) \\ &\leq \sum_{k=1}^{n} \frac{1}{k^2} \mathbb{E}[Y_1^2 1_{\{|Y_1|\leq k\}}] \\ &\leq \mathbb{E}[Y_1^2 \sum_{k=1}^{n} \frac{1}{k^2} 1_{[|Y_1|,+\infty[}(k)\ ] \\ &\leq \mathbb{E}[Y_1^2 \sum_{k\geq \max\ \{|Y_1|,1\}} \frac{1}{k^2}\ ] \\ &\leq \mathbb{E}[Y_1^2 \frac{2}{\max\ \{|Y_1|,1\}}] \\ &\leq 2\ \mathbb{P}[\{|Y_1| < 1\}] + 2\ \mathbb{E}[|Y1|]. \end{aligned}$$

La martingale $(X_n)_{n\geq 0}$ est donc bornée dans L^2.

D'autre part, nous avons :

$$\begin{aligned} \mathbb{E}[|\sum_{k=1}^{n} \frac{1}{k} T_k|] &\leq 2 \sum_{k=1}^{n} \frac{1}{k} \mathbb{E}[|Y_1| 1_{\{|Y_1|>k\}}] \leq 2 \sum_{k\geq 1} \frac{1}{k} \mathbb{E}[|Y_1| 1_{\{|Y_1|>k\}}] \\ &\leq 2\mathbb{E}[\ |Y_1| 1_{\{|Y_1|>1\}} \sum_{k=1}^{|Y_1|} \frac{1}{k}\ ] \leq 2\mathbb{E}[|Y_1|(1 + \operatorname{Log}^+|Y_1|)], \end{aligned}$$

ce qui montre que la martingale $(M_n)_{n\geq 0}$ est bornée dans L^1. $\square$

Exercices 2.5.3

Exercice 1

On considère l'espace probabilisé $([0,1],\mathcal{B}([0,1]),\lambda)$, $\mathcal{B}([0,1])$ étant la tribu borélienne de $[0,1]$ et λ la mesure de Lebesgue.

Soit f une fonction borélienne **croissante** sur $[0,1]$ à valeurs réelles. On note

$$\Phi_n(y) = 2^n \sum_{k=0}^{2^n-1} (f(\frac{k+1}{2^n}) - f(\frac{k}{2^n}))\, 1_{[\frac{k}{2^n},\frac{k+1}{2^n}[}(y),\ \ n \in \mathbb{N},$$

et on appelle $\mathcal{F}_n,\ n \geq 0$, la tribu engendrée par la partition $\{[\frac{k}{2^n},\frac{k+1}{2^n}[, k = 0,1,\cdots,2^n-1\}$ de l'intervalle $[0,1[$.

1) Montrer que $(\Phi_n,\mathcal{F}_n)_{n\geq 0}$ est une martingale positive. Qu'en déduit-on ?

2) On suppose qu'il existe h borélienne positive telle que f soit de la forme $f(x) = \int_0^x h(y)\ dy$.

Montrer que $X_n = \mathbb{E}[h|\mathcal{F}_n]$, $n \geq 0$. Qu'en déduit-on ?

3) On suppose que f est de la forme $f(x) = \mu([0,x[),\ \ x \in [0,1]$, μ étant une mesure de Radon positive sur $[0,1]$.

a) Montrer que $\mu(A) = \int_A \Phi_n(y)\ dy$, $\forall A \in \mathcal{F}_n$.

Soient μ_n (resp. λ_n) la restriction de μ (resp. λ) à la tribu $\mathcal{F}_n$.

Que représente la v.a.r. Φ_n pour les mesures μ_n et λ_n ?

b) On note $\mu = Z\,\lambda + \nu$ la décomposition de Radon-Nikodym-Lebesgue de μ par rapport à λ. On se propose de prouver que,

$$\lim_n \Phi_n = Z,\ \ \lambda-\ \text{p.s.}. \tag{2.14}$$

b_1) Si Z' désigne la limite λ-p.s. de la suite $(\Phi_n)_{n\geq 0}$, montrer que l'on a :

$$\mu(A) \geq Z'\,\lambda(A),\ \ \forall A \in \mathcal{F}_n, \forall n \geq 0.$$

b_2) En déduire que $\nu' = \mu - Z'\,\lambda$ est une mesure de Radon positive.

b_2) En considérant la décomposition de Radon-Nikodym-Lebesgue de ν' par rapport à λ , montrer que ν' est étrangère à μ. En déduire (2.14).

Exercice 2

Si x est un réel positif, nous notons $[x]$ sa partie entière. Soit Z une v.a.r. définie sur un espace probabilisé $(\Omega ega,\mathcal{F},\mathbb{P})$ distribuée suivant la loi uniforme sur $[0,1[$. Pour tout entier $n \geq 0$, posons :

$$Y_n = \frac{[10^n Z]}{10^n} \ \text{ et } A_{n+1} = 10^{n+1}(Y_{n+1} - Y_n) \in \{0,\ldots,9\}.$$

On notera qu'en écriture décimale, on a :

$$Y_n = 0, A_1A_2\ldots A_n \ \text{ et } \ Z = 0, A_1A_2\ldots A_n\ldots.$$

Notons $(\mathcal{F}_n)_{n\geq 0}$ la filtration naturelle associée au processus $(Y_n)_{n\geq 0}$ (pour tout $n \geq 0$, $\mathcal{F}_n$ est la tribu engendrée par les v.a.r. $Y_0, \dots, Y_n$).

1) Montrer que les v.a.r. A_n, $n \geq 1$, sont indépendantes et de loi uniforme sur $\{0, \dots, 9\}$.

2) Soit f une fonction borélienne bornée sur [0,1]. On pose :

$$\forall n \geq 0,\ X_n = 10^n[f(Y_n + 10^{-n}) - f(Y_n)].$$

Montrer que le processus $(X_n)_{n\geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

3) On suppose que f est Lipschitzienne, c'est-à -dire qu'il existe un réel $C > 0$ tel que :

$$\forall x, y \in [0,1],\ |f(x) - f(y)| \leq C|x - y|.$$

a) Montrer que la martingale $(X_n)_{n\geq 0}$ converge, $\mathbb{P}$-p.s. et en norme L^1, vers une v.a.r. X_∞.

b) Pour tout entier $n \geq 0$ et tous réels a, b vérifiant $0 \leq a < b \leq 1$, calculer $\mathbb{E}[1_{[a,b]}(Z)\ X_n]$.

c) Montrer qu'il existe une fonction borélienne bornée g sur $[0,1]$ telle que

$$\forall 0 \leq a < b \leq 1,\ f(b) - f(a) = \int_a^b g(t)\ dt.$$

Exercice 3

Soient $(\Omega, \mathcal{F})$ un espace mesurable et $(\mathcal{F}_n)_{n\geq 0}$ une suite croissante de sous-tribus de $\mathcal{F}$ telle que $\vee_n \mathcal{F}_n = \mathcal{F}$.

On se donne deux probabilités $\mathbb{P}$ et $\mathbb{Q}$ sur $(\Omega, \mathcal{F})$. Pour tout entier $n \geq 0$, on désigne respectivement par $\mathbb{P}_n$ et $\mathbb{Q}_n$ les restrictions des probabilités $\mathbb{P}$ et $\mathbb{Q}$ à la tribu $\mathcal{F}_n$.

1) On suppose que, pour tout $n \geq 0$, la probabilité $\mathbb{Q}_n$ est absolument continue par rapport à la probabilité $\mathbb{P}_n$ et on note Z_n la dérivée de Radon-Nikodym de $\mathbb{Q}_n$ par rapport à $\mathbb{P}_n$.

a) Montrer que, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, le processus $(Z_n)_{n\geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

b) Montrer que le processus $(Z_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. vers une v.a.r. positive ou nulle Z_∞ telle que $\int_\Omega\ Z_\infty\ d\mathbb{P} \leq 1$.

c) Montrer que la famille de v.a.r. est équi-intégrable si, et seulement si, $\int_\Omega\ Z_\infty\ d\mathbb{P} = 1$.

d) On suppose que la probabilité $\mathbb{Q}$ est absolument continue par rapport à la probabilité $\mathbb{P}$.

Pour tout entier $n \geq 0$, exprimer Z_n à l'aide de la dérivée de Radon-Nikodym $\dfrac{d\mathbb{Q}}{d\mathbb{P}}$.

En déduire que $\int_\Omega\ Z_\infty\ d\mathbb{P} = 1$.

e) Réciproquement on suppose que $\int_\Omega\ Z_\infty\ d\mathbb{P} = 1$. Montrer que la probabilité $\mathbb{Q}$ est absolument continue par rapport à la probabilité $\mathbb{P}$.

2) On se place maintenant dans le cas général et on désigne par $\mathbb{Q} = U\mathbb{P} + \mathbb{L}$ la décomposition de Radon-Nikodym-Lebesgue de la probabilité $\mathbb{Q}$ par rapport à la probabilité $\mathbb{P}$ (U est une v.a.r positive ou nulle et $\mathbb{L}$ est une mesure positive étrangère à $\mathbb{P}$). De même pour tout entier $n \geq 0$, on note $\mathbb{Q}_n = U_n\ \mathbb{P}_n + I\!K_n$ la décomposition de Radon-Nikodym-Lebesgue de $\mathbb{Q}_n$ par rapport à $\mathbb{P}_n$ (U_n est une v.a.r. $\mathcal{F}_n$-mesurable et $I\!K_n$ une mesure positive sur $\mathcal{F}_n$ étrangère à $\mathbb{P}_n$).

Montrer que la suite de v.a.r. $(U_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. vers U. [Pour cela, on commencera par appliquer le résultat précédent aux couples

$$(\mathbb{Q}, \frac{\mathbb{P}+\mathbb{Q}}{2}) \quad \text{et} \quad (\mathbb{P}, \frac{\mathbb{P}+\mathbb{Q}}{2}).]$$

Solution de l'exercice

Partie 1

1) a) Pour tout $n \geq 0$, Z_n est une v.a.r. $\mathcal{F}_n$-mesurable positive telle que

$$\mathbb{E}_{\mathbb{P}}\ [Z_n] = \mathbb{E}_{\mathbb{P}_n}\ [Z_n] = \mathbb{E}_{\mathbb{Q}_n}\ [1] = 1.$$

D'autre part, pour tout $A \in \mathcal{F}_n$ nous avons :

$$\begin{aligned} \int_A Z_{n+1}\ d\mathbb{P} &= \int_A Z_{n+1}\ d\mathbb{P}_{n+1} = \mathbb{Q}_{n+1}[A] = \mathbb{Q}_n[A] \\ &= \int_A Z_n\ d\mathbb{P}_n = \int_A Z_n\ d\mathbb{P}. \end{aligned}$$

Ceci montre que

$$\mathbb{E}_{\mathbb{P}}[Z_{n+1}|\mathcal{F}_n] = Z_n,\ \mathbb{P}\text{-p.s.}.$$

Le processus $(Z_n)_{n\geq 0}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P})$ est donc une martingale positive relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

b) Nous avons $\sup_{n\geq 0} \mathbb{E}_{\mathbb{P}}\ [Z_n] = 1$, d'après le théorème de convergence des martingales, le processus $(Z_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. vers une v.a.r $\mathbb{P}$-intégrable Z_∞ telle que (lemme de Fatou)

$$\mathbb{E}_{\mathbb{P}}\ [Z_\infty] \leq \liminf_{n\to+\infty}\ \mathbb{E}_{\mathbb{P}}\ [Z_n] = 1.$$

c) Si la famille $\{Z_n : n \geq 0\}$ est équi-intégrable, alors la suite de v.a.r. $(Z_n)_{n\geq 0}$ converge dans $L^1(\mathbb{P})$ vers Z_∞ et par suite

$$\mathbb{E}_{\mathbb{P}}\ [Z_\infty] = \lim_{n\to+\infty}\ \mathbb{E}_{\mathbb{P}}\ [Z_n] = 1.$$

Réciproquement, supposons que $\mathbb{E}_{\mathbb{P}}\ [Z_\infty] = 1$. Ecrivons :

$$|Z_\infty - Z_n| = Z_\infty + Z_n - 2\ Z_\infty \wedge Z_n.$$

Le processus $(Z_\infty \wedge Z_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. vers Z_∞ et $0 \leq Z_\infty \wedge Z_n \leq Z_\infty$; d'après le théorème de convergence dominée, ce processus converge dans $L^1(\mathbb{P})$ vers Z_∞, ce qui entraîne la convergence de

$$\mathbb{E}_{\mathbb{P}}[|Z_\infty - Z_n|] = 2 - 2\mathbb{E}_{\mathbb{P}}[\ Z_\infty \wedge Z_n]$$

vers zéro. Ceci montre que le processus $(Z_n)_{n\geq 0}$ converge dans $L^1(\mathbb{P})$ vers Z_∞ ; la famille $\{Z_n : n \geq 0\}$ est donc équi-intégrable.

d) Supposons que $\mathbb{Q} = U\ \mathbb{P}$, pour une v.a.r. positive U vérifiant $E_{\mathbb{P}}[U] = E_{\mathbb{Q}}[1] = 1$. Pour tout $A \in \mathcal{F}_n$, nous avons

$$\mathbb{Q}_n[A] = \int_A d\mathbb{Q} = \int_A U\ d\mathbb{P} = \int_A \mathbb{E}_{\mathbb{P}}[U|\mathcal{F}_n]\ d\mathbb{P}_n$$

et par suite $Z_n = \mathbb{E}_{\mathbb{P}}[U|\mathcal{F}_n]$, $\mathbb{P}$-p.s.. Le processus $(Z_n)_{n\geq 0}$ est donc une martingale équi-intégrable convergeant $\mathbb{P}$-p.s. vers U.

e) Supposons que $\int_\Omega Z_\infty \, d\mathbb{P} = 1$. D'après *b)* et *c)*, le processus $(Z_n)_{n\geq 0}$ converge $\mathbb{P}$-p.s. et dans $L^1(\mathbb{P})$ vers Z_∞.

Pour tout entier $p \geq 0$ et tout $A \in \mathcal{F}_p$, nous avons donc :

$$\forall n \geq p, \; \mathbb{Q}[A] = \int_A Z_p \, d\mathbb{P} = \int_A Z_n \, d\mathbb{P} = \int_A Z_\infty \, d\mathbb{P}.$$

Ce qui montre que les probabilités $\mathbb{Q}$ et $Z_\infty \, \mathbb{P}$ coïncident sur l'algèbre de Boole $\bigcup_{p\geq 0} \mathcal{F}_p$ et par suite sur $\mathcal{F} = \vee_{p\geq 0} \mathcal{F}_p$.

2) Appelons respectivement V et W les dérivées de Radon-Nikodym des probabilités $\mathbb{Q}$ et $\mathbb{P}$ par rapport à $\dfrac{\mathbb{P}+\mathbb{Q}}{2}$. On peut choisir des versions de V et W à valeurs dans [0,2]. D'après 1), les probabilités $\mathbb{Q}_n$ et $\mathbb{P}_n$ admettent respectivement, par rapport à $\dfrac{\mathbb{P}_n+\mathbb{Q}_n}{2}$, les densités $V_n = \mathbb{E}_{\frac{\mathbb{P}+\mathbb{Q}}{2}} [V|\mathcal{F}_n]$ et $W_n = \mathbb{E}_{\frac{\mathbb{P}+\mathbb{Q}}{2}} [W|\mathcal{F}_n]$.

Nous avons alors les décompositions de Lebesgue :

$$\mathbb{Q}_n = \frac{V_n}{W_n} \, 1_{\{W_n>0\}} \, \mathbb{P}_n + 1_{\{W_n=0\}} \, \mathbb{Q}_n \quad \text{et} \quad \mathbb{Q} = \frac{V}{W} \, 1_{\{W>0\}} \, \mathbb{P} + 1_{\{W=0\}} \, \mathbb{Q}.$$

Puisque $\mathbb{P}[W=0]=0$, la suite de v.a.r.

$$U_n = \frac{V_n}{W_n} \, 1_{\{W_n>0\}}$$

converge bien $\mathbb{P}$-p.s. vers la v.a.r.

$$U = \frac{V}{W} \, 1_{\{W>0\}}.$$

D'où le résultat.

Exercices 4 : On considère une filtration $(\mathcal{F}_n)_{n\geq 0}$, un processus (X_n) adapté à $(\mathcal{F}_n)$ et un temps d'arrêt fini T pour $(\mathcal{F}_n)$.

1. Montrer que la variable aléatoire X_T définie par $\omega \to X_{T(\omega)}(\omega)$ est $\mathcal{F}_T$-mesurable.

2. On définit le processus (X_n^T) "arrêté" par le temps d'arrêt T par $X_n^T = X_{T\wedge n}$.

a) Montrer que le processus $(X_n^T)_{n\geq 0}$ est encore adapté à $(\mathcal{F}_n)_{n\geq 0}$.

b) Montrer que, si $(X_n)_{n\geq 0}$ est une surmartingale [resp. une martingale] relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$, le processus $(X_n^T)_{n\geq 0}$ arrêté est encore une surmartingales [resp. une martingale] relativement aux filtrations $(\mathcal{F}_n)_{n\geq 0}$ et $(\mathcal{F}_{T\wedge n})_{n\geq 0}$.

Exemples et exercices 2.5.4

1) Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(\mathcal{F}_n)_{n\geq 0}$ une filtration sur cet espace. Pour toute variable aléatoire réelle X intégrable, la suite des espérances conditionnelles $(X_n)_{n\geq 0} = (\mathbb{E}[X|\mathcal{F}_n])_{n\geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

2) Si $(Y_n)_{n\geq 0}$ est une suite de variables aléatoires indépendantes intégrables, la suite (S_n) des sommes partielles, $S_n = Y_0 + \cdots + Y_n$ est une martingale si les Y_n sont centrées, une surmartingale si $\mathbb{E}[Y_n] \leq 0$, $\forall n \geq 1$.

3)

a) Si $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une sous-martingale, $(-X_n, \mathcal{F}_n)_{n\geq 0}$ et $(X_n^-, \mathcal{F}_n)_{n\geq 0}$ sont des surmartingales.

b) Si $(X_n)_{n\geq 0}$ et $(Y_n)_{n\geq 0}$ sont deux martingales relativement à la même filtration $(\mathcal{F}_n)_{n\geq 0}$, alors le processus $(X_n \wedge Y_n)_{n\geq 0}$ est une surmartingale relativement à $(\mathcal{F}_n)_{n\geq 0}$.

c) Si $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale et f une fonction convexe sur $\mathbb{R}$ telle que, pour tout $n \geq 0$, $f(X_n)$ soit intégrable, alors $(f(X_n), \mathcal{F}_n)_{n\geq 0})$ est une surmartingale.

Ce résultat subsiste pour une sous-martingale $(X_n, \mathcal{F}_n)_{n\geq 0}$, si f est de plus croissante.

4) Si $(X_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale (resp. surmartingale, resp. sous-martingale) et $(\mathcal{G}_n)$ une filtration telle que, pour tout $n \geq 0$, X_n soit $\mathcal{G}_n$-mesurable et $\mathcal{G}_n \subset \mathcal{F}_n$, alors $(X_n, \mathcal{G}_n)_{n\geq 0}$ est une martingale (resp. surmartingale, resp. sous-martingale).

Chapitre 3

Chaînes de Markov

3.1 Noyaux positifs et noyaux markoviens

Nous introduisons la notion de noyau positif et en particulier de noyau markovien conduisant à la construction des processus markoviens. Nous appliquerons ensuite la théorie des martingales à l'étude de ces processus.

Définition : Etant donnés deux espaces mesurables $(E, \mathcal{E})$ et $(F, \mathcal{F})$, on appelle **noyau positif** de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$ toute application N de $E \times \mathcal{F}$ dans $\mathbb{R}^+$ telle que :
$i)$ pour tout $x \in E$, $N(x, .)$ est une mesure positive (σ-finie) sur la tribu $\mathcal{F}$,
$ii)$ pour tout $B \in \mathcal{F}$, l'application $x \to N(x, B)$ est $\mathcal{E}$-mesurable.

• **Action d'un noyau sur les fonctions et les mesures**

Lemme 3.1.1 *Soit N un noyau positif de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$. Soit f une fonction sur l'espace produit $E \times F$, $\mathcal{E} \otimes \mathcal{F}$-mesurable positive. La formule*

$$g(x) = \int_F f(x, y)\ N(x, dy),\ \ x \in E,$$

définit une fonction sur E, $\mathcal{E}$-mesurable à valeurs dans $[0, +\infty]$.

Preuve : Si f est l'indicatrice d'un pavé $A \times B$ de $\mathcal{E} \otimes \mathcal{F}$, nous avons $g(x) = 1_A(x)\, N(x, B)$, et la fonction g est bien $\mathcal{E}$-mesurable.

Posons $\mathcal{M} = \{C \in \mathcal{E} \otimes \mathcal{F} : \int 1_C(., y)\ N(., dy)$ soit $\mathcal{E}$-mesurable$\}$. La famille $\mathcal{M}$ contient l'algèbre de Boole engendrée par les pavés et est stable par limites croissantes et décroissantes. D'après le théorème des classes monotones, on a donc $\mathcal{M} = \mathcal{E} \otimes \mathcal{F}$.

Nous avons ainsi montré le résultat pour les fonctions f indicatrices d'éléments de $\mathcal{E} \otimes \mathcal{F}$. Par linéarité, le résultat s'étend aux fonctions étagées, puis par limite croissante à toutes les fonctions $\mathcal{E} \otimes \mathcal{F}$-mesurables positives. $\square$

Pour toute fonction f $\mathcal{F}$-mesurable positive, on définit une fonction Nf $\mathcal{E}$-mesurable positive, en posant

$$Nf(x) = \int_E f(y)\ N(x, dy).$$

Pour toute mesure positive ν sur $(E,\mathcal{E})$, on définit une mesure positive νN sur $(F,\mathcal{F})$ par

$$\int f\ d(\nu N) = \int_E Nf\ d\nu,\ f \text{ positive } \mathcal{F}-\text{mesurable.}$$

La construction précédente contient le cas particulier où $(E,\mathcal{E})$ et $(F,\mathcal{F})$ coïncident. Nous disons alors que N est un noyau positif sur $(E,\mathcal{E})$. Dans ce cas, on obtient un opérateur N agissant sur les fonctions $\mathcal{E}$-mesurables positives, ainsi que sur les mesures positives sur $(E,\mathcal{E})$.

Composition de noyaux

Soient $(E,\mathcal{E})$, $(F,\mathcal{F})$ et $(G,\mathcal{G})$ trois espaces mesurables. Etant donnés des noyaux positifs N et M respectivement de $(E,\mathcal{E})$ dans $(F,\mathcal{F})$ et de $(F,\mathcal{F})$ dans $(G,\mathcal{G})$, nous définissons un noyau positif NM, composé de N et de M, de $(E,\mathcal{E})$ dans $(G,\mathcal{G})$ en posant, pour toute fonction f $\mathcal{G}$-mesurable positive :

$$NMf(x) = \int_F Mf(y)\ N(x,dy).$$

On montre que cette composition est associative.

En particulier, si N est un noyau positif sur $(E,\mathcal{E})$, par compositions successives, on obtient les noyaux itérés $N^k, k \geq 1$, de N. On convient que N^0 est l'opérateur identité, ou encore que $N^0(x,dy)$ est la mesure de Dirac $\delta_x(dy)$ en x.

• Une classe de noyaux positifs

Notations : Considérons un ensemble S de transformations sur E. Nous noterons $x \to s.x$, ou simplement $x \to sx$, l'action sur E d'une transformation s appartenant à la famille S.

Supposons que S soit muni d'une tribu $\mathcal{S}$ telle que l'application $(s,x) \to s.x$ soit mesurable sur l'espace $S \times E$ muni de la tribu produit.

Soit Q un noyau positif de $(E,\mathcal{E})$ dans $(S,\mathcal{S})$.

On définit un noyau positif N sur $(E,\mathcal{E})$ par

$$N(x,A) = \int_S 1_A(s.x)\ Q(x,ds),\ \forall A \in \mathcal{E}.$$

L'opérateur associé s'écrit alors :

$$Nf(x) = \int_S f(s.x)\ Q(x,ds),\ \forall f\ \mathcal{E}-\text{mesurable positive.}$$

Dans la plupart des exemples, l'espace E sera un espace métrique réunion dénombrable de compacts, muni de sa tribu borélienne $\mathcal{B}(E)$. On appellera **noyau positif sur** E tout noyau positif sur $(E,\mathcal{B}(E))$.

Lorsque S est une famille finie ou dénombrable de transformations continues sur E, le noyau positif s'écrit :

$$Nf(x) = \sum_{s\in S} q(s,x)f(s.x), \tag{3.1}$$

pour une famille $q(s,.)$, $s \in S$, de fonctions boréliennes positives.

• **Noyaux markoviens**

Définition 3.1.2 Un noyau positif P de $(E,\mathcal{E})$ dans $(F,\mathcal{F})$ est dit **markovien** si, pour tout $x \in E$, $P(x,F) = 1$. Dans ce cas, le noyau transforme toute fonction $\mathcal{F}$-mesurable bornée en une fonction $\mathcal{E}$-mesurable bornée.

Quand les espaces $(E,\mathcal{E})$ et $(F,\mathcal{F})$ sont confondus, un noyau markovien est appelé aussi **probabilité de transition**. Etant donné un espace mesurable $(E,\mathcal{E})$, une probabilité de transition sur $(E,\mathcal{E})$ est donc une application N de $E \times \mathcal{F}$ dans $\mathbb{R}^+$ telle que :
$i)$ pour tout $x \in E$, $N(x,.)$ est une mesure de probabilité sur la tribu $\mathcal{E}$,
$ii)$ pour tout $B \in \mathcal{E}$, l'application $x \to N(x,E)$ est $\mathcal{E}$-mesurable.

Elle définit un opérateur sur l'espace vectoriel des fonctions $\mathcal{E}$-mesurables bornées et opère sur le convexe des mesures de probabilités sur $\mathcal{E}$.

Reprenons le noyau défini par (3.1), en notant $p(s,.)$, $s \in S$, une famille de fonctions boréliennes positives vérifiant

$$\sum_{s\in S} p(s,x) = 1, \forall x \in E. \tag{3.2}$$

Le noyau P sur E défini par

$$Pf(x) = \sum_{s\in S} p(s,x)f(s.x), \tag{3.3}$$

est alors markovien.

La donnée des poids $p(s,\cdot)$, $s \in S$, définit un processus aléatoire (à temps discret), que l'on peut se représenter comme la trajectoire d'une particule se déplaçant sur E et passant de x à $s.x$ avec la probabilité $p(s,x)$.

Nous allons présenter plusieurs exemples de noyaux de ce type.

3.2 Exemples

Opérateurs de transfert

a) Soit p un entier > 1. On considère la transformation τ définie sur le cercle E, identifié à $[0,1[$, par

$$\tau : x \to px \bmod 1.$$

On note $S = \{s_i, i = 0,\dots,p-1\}$ la famille des inverses de τ.

Pour $i = 0,\dots,p-1$, on a $s_i x = \dfrac{x+i}{p}$. Etant donnée une fonction u continue positive sur E, on définit un noyau positif N_u par :

$$N_u f(x) = \sum_{i=0}^{p-1} u(s_i x)f(s_i x).$$

Lorsque u est constant et égal à $1/p$, on obtient un noyau markovien qui est le dual de la composition par τ relativement à la mesure de Lebesgue sur le cercle.

b) Plus généralement, considérons une matrice carrée A d'ordre $d \geq 1$ inversible, à coefficients réels, **dilatante** (c'est-à-dire dont toutes les valeurs propres sont de module strictement supérieur à 1) et une famille $T = \{\sigma_0, \sigma_1, \dots, \sigma_{q-1}\}$, de $q \geq 1$ éléments de $\mathbb{R}^d$.

Choisissons sur $\mathbb{R}^d$ une norme $\| \quad \|$ pour laquelle $\|A^{-1}\| < 1$. Pour $\sigma \in T$, notons $\breve{\sigma}$ l'affinité de $\mathbb{R}^d$ définie par $\breve{\sigma}x = A^{-1}(x+\sigma), x \in \mathbb{R}^d$. Nous désignons par S la famille de ces affinités.

La boule fermée K de centre 0 et de rayon $R = \sup_{\tau \in T} \|\tau\| \dfrac{\|A^{-1}\|}{(1 - \|A^{-1}\|)}$ est invariante par les affinités de la famille S et constitue un "attracteur" pour cette famille.

On a en effet, pour $x \in \mathbb{R}^d$ et $(\tau_k)_{k \geq 0}$ dans T, $\breve{\tau}_n \cdots \breve{\tau}_1 x = \breve{\tau}_n \cdots \breve{\tau}_1 0 + A^{-n}x$, et donc :

$$\lim_{n \to +\infty} d(\breve{\tau}_n \cdots \breve{\tau}_1 x, K) = 0.$$

Considérons une fonction u à valeurs positives ou nulles, appartenant à l'espace $C(\mathbb{R}^d)$ des fonctions continues sur $\mathbb{R}^d$, et formons l'opérateur N_u défini par

$$N_u f(x) = \sum_{s \in S} u(sx) f(sx), \quad x \in \mathbb{R}^d.$$

Cet opérateur opère sur $C(\mathbb{R}^d)$ et peut être restreint à l'espace $C(K)$ des fonctions continues sur K.

Prenons maintenant pour A une matrice dilatante à coefficients entiers et pour S une famille de représentants dans $\mathbb{Z}^d$ de $\mathbb{Z}^d/A\mathbb{Z}^d$ (on a alors $q = |\det A|$). Lorsque u est $\mathbb{Z}^d$-périodique, le noyau N_u opère sur les fonctions $\mathbb{Z}^d$-périodiques.

Notons τ_A l'action de A sur le tore $\mathbb{T}^d$, $x \to \tau_A x = Ax \bmod \mathbb{Z}^d$. Si u est constant et égal à $|\det A|^{-1}$, on obtient un noyau markovien qui est le dual de la composition par τ_A relativement à la mesure de Lebesgue sur $\mathbb{T}^d$.

Action sur un espace projectif

Considérons le groupe $G = \mathrm{GL}(d, \mathbb{R})$ des matrices carrées inversibles d'ordre d. Ce groupe agit sur l'espace projectif $\mathcal{P}(\mathbb{R}^d)$. Soit μ une mesure de probabilité sur G. On définit un noyau markovien P_μ sur $\mathcal{P}(\mathbb{R}^d)$ par :

$$P_\mu f(x) = \int_G f(g.x)\ \mu(dg).$$

Rotations aléatoires

Munissons le cercle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ de la mesure de Lebesgue notée λ et considérons N angles $\alpha_1, \dots, \alpha_N$, deux à deux distincts et des poids $p_1, \dots, p_N$, strictement positifs et de somme égale à 1. La formule

$$Pf(x) = \sum_{j=1}^{N} p_j f(x + \alpha_j).$$

définit un noyeau markovien sur $\mathbb{T}$.

Dans le cas particulier où l'on prend comme angles α et $-\alpha$ avec des poids égaux, on obtient l'opérateur de transition "symétrique" : $Pf(x) = \frac{1}{2}[f(x+\alpha) + f(x-\alpha)]$.

3.3 Mesures invariantes pour un opérateur markovien

Soit P une probabilité de transition sur un espace mesurable $(E, \mathcal{E})$. Nous notons P^n, $n \geq 0$, les itérées de P ; par convention, P^0 est l'opérateur identité.

Définition 3.3.1 Une fonction mesurable h sur E, à valeurs dans $[-\infty, +\infty]$, $\mathcal{E}$-mesurable, est dite P-**harmonique** si, pour tout $x \in E$, h est $P(x, \cdot)$-quasi-intégrable (i.e. de partie positive ou de partie négative $P(x, \cdot)$-intégrable) et si $Ph = h$, i.e. :

$$\forall x \in E, \ Pf(x) = \int_E f(y) \ P(x, dy) = f(x).$$

Définitions 3.3.2 Une mesure positive ν est dite P-**invariante** si elle vérifie $\nu P = \nu$. Dans ce cas, pour tout $1 \leq p \leq +\infty$, P définit une contraction sur $L^p(E, \mathcal{E}, \nu)$.

Une mesure positive ν est dite P-**invariante extrémale** si elle est P-invariante et si toute mesure positive P-invariante ν', vérifiant $\nu' \leq \nu$, est nécessairement un multiple de ν.

Soit ν une mesure positive P-invariante sur $(E, \mathcal{E})$. Une fonction mesurable h sur E est dite P-**harmonique** ν-**p.p.** si, pour ν-presque tout $x \in E$, h est $P(x, \cdot)$-quasi-intégrable et $Ph(x) = h(x)$.

Lemme 3.3.3 *Si ν est une mesure positive P-invariante (ou plus généralement telle que νP soit absolument continue par rapport à ν), toute fonction f mesurable bornée, P-harmonique ν-p.s., est ν-presque sûrement égale à une fonction h mesurable bornée, P-harmonique.*

Preuve : La fonction définie par

$$g = \liminf_{n \to +\infty} \frac{f + \ldots + P^{n-1} f}{n}$$

vérifie $Pg \leq g$ et $g = f$, ν-p.s.. On obtient le résultat voulu, en posant $h = \lim_{n \to +\infty} \downarrow \ P^n g$. □

Lemme 3.3.4 *Soit ν une mesure positive P-invariante extrémale. Alors toute fonction h de $\mathcal{L}^1(E, \mathcal{E}, \nu)$ telle que la mesure $h\nu$ soit P-invariante, est ν-p.p. constante.*

Preuve : Soit h dans $\mathcal{L}^1(E, \mathcal{E}, \nu)$ telle que la mesure $h\nu$ soit P-invariante. Nous avons, pour tout $A \in \mathcal{E}$:

$$h^+\nu \ (A) = h\nu \ (A \cap \{h > 0\}) \leq (h^+\nu)P \ (A \cap \{h > 0\}) \leq (h^+\nu)P \ (A).$$

Comme $P1 = 1$, les inégalités précédentes sont des égalités ; d'où $(h^+\nu)P = h^+\nu$.

Pour tout $C > 0$, nous avons :

$$(\inf\{h^+, C\}\nu)P \leq (h^+\nu) \wedge (C\nu) = \inf\{h^+, C\}\nu$$

et, comme $P1 = 1$, on a en fait l'égalité : $(\inf\{h^+, C\}\nu)P = \inf\{h^+, C\}\nu$.

L'hypothèse d'extrémalité de ν implique que, pour tout $C > 0$, $h^+ \wedge C$ est ν-p.s. constante. On a donc montré que h^+ est ν-p.s. constante.

De la même façon, on montre que h^- est ν-p.s. constante. □

Lemme 3.3.5 *Si T est une contraction ($\|T\| \leq 1$) dans un espace de Hilbert, T et son adjoint T^* ont les mêmes vecteurs invariants. Si $Tf = \lambda f$, avec $|\lambda| = 1$, alors $T^* f = \overline{\lambda} f$.*

Preuve : Pour une contraction dans un espace de Hilbert, la relation $Tf = f$ est équivalente à $\langle Tf, f \rangle = \|f\|^2$, (cas d'égalité dans l'inégalité de Cauchy-Schwarz).

Les opérateurs T et T^* ayant la même norme, T^* est aussi une contraction et les opérateurs ont les mêmes vecteurs invariants.

Ce résultat appliqué à $\overline{\lambda} T$, où λ est une valeur propre de module 1 de T, montre la deuxième assertion. $\square$

Lemme 3.3.6 *Soient P un noyau markovien et ν une probabilité P-invariante sur un espace mesurable $(E, \mathcal{E})$. Les assertions suivantes sont équivalentes :*
i) la mesure ν est P-invariante extrémale ;
ii) tout élément de $\mathcal{L}^2(E, \mathcal{E}, \nu)$ P-harmonique ν-p.s. est ν-presque sûrement constant ;
iii) toute fonction mesurable bornée P-harmonique ν-p.s. est ν-presque sûrement constante ;
iv) toute fonction mesurable bornée P-harmonique est ν-presque sûrement constante.

Preuve :
i) $\Rightarrow$ *ii)* Soit h un élément P-invariant de $L^2(E, \mathcal{E}, \nu)$. D'après le lemme 3.3.5, h est P^*-invariante. La mesure $h\nu$ est donc P-invariante et d'après le lemme 3.3.4, puisque ν est extrémale, h est ν-p.s. constante.

L'assertion *ii)* $\Rightarrow$ *iii)* est évidente et les assertions *iii)* et *iv)* sont équivalentes d'après le lemme 3.3.3.

iii) $\Rightarrow$ *i)* Soit ν' une mesure positive P-invariante vérifiant $\nu' \leq \nu$.

D'après le théorème de Radon-Nikodym, il existe une fonction mesurable h vérifiant $0 \leq h \leq 1$ telle que $\nu' = h\ \nu$. Cette fonction définit une classe de $L^2(E, \mathcal{E}, \nu)$ invariante par l'opérateur dual de P et donc invariante par P, d'après le lemme 3.3.5, puisque P est une contraction de $L^2(E, \mathcal{E}, \nu)$. Il en résulte que h est une fonction bornée, P-harmonique ν-p.s.. Sous la condition *iii)*, h est ν-p.s. constante et ν' est un multiple de ν. $\square$

Remarquons que le lemme 3.3.5 peut être précisé dans la situation étudiée ici de la façon suivante :

Lemme 3.3.7 *Soient P un noyau markovien et ν une mesure positive P-invariante sur un espace mesurable $(E, \mathcal{E})$. Pour toute fonction h de $\mathcal{L}^2(E, \mathcal{E}, \nu)$ vérifiant $Ph = h$, ν-p.s., nous avons, pour ν-presque tout $x \in E$ et pour tout entier $n \geq 1$,*

$$h(y) = h(x), \text{ pour } P^n(x, \cdot)\text{-presque tout } y.$$

Preuve : Soit n un entier ≥ 1. Nous avons, pour ν-presque tout $x \in E$,

$$\int_E (h(y) - h(x))^2 \ P^n(x, dy) \ = \ P^n h^2(x) - h^2(x)$$

et l'intégrale par rapport à ν du second membre est nulle. $\square$

3.4 Chaîne canonique associée à un noyau markovien

Considérons un espace mesurable $(E, \mathcal{E})$. Soient Ω l'espace produit $E^{\mathbb{N}}$, $\mathcal{F}$ la tribu produit de Ω et $(X_n)_{n\geq 0}$ les applications coordonnées sur l'espace Ω.

Nous appelons **décalage** (ou "shift") la transformation θ sur Ω définie par

$$\theta(\omega_0, ..., \omega_n, ...) = (\omega_1, ..., \omega_{n+1}, ...).$$

Pour tout entier $n \geq 0$, nous avons $X_n \circ \theta = X_{n+1}$.

Pour tout entier $n \geq 0$, nous appelons $\mathcal{F}_n = \sigma(X_0, ..., X_n)$ la sous-tribu de $\mathcal{F}$ engendrée par les variables aléatoires $\{X_k : 0 \leq k \leq n\}$.

Théorème 3.4.1 *Soit P un noyau markovien sur $(E, \mathcal{E})$. Pour tout $x \in E$, il existe une unique mesure de probabilité, notée $\mathbb{P}_x$, sur l'espace mesurable $(\Omega, \mathcal{F})$ telle que, pour tout entier $n \geq 1$ et toute fonction mesurable bornée (ou positive) f sur E^n,*

$$\int_\Omega f(X_0(\omega), \ldots, X_n(\omega))\ \mathbb{P}_x(d\omega) \tag{3.4}$$

$$= \int_E \cdots \int_E f(x, u_1, \ldots, u_n)\ P(u_{n-1}, du_n) \cdots P(x, du_1). \tag{3.5}$$

Pour tout $B \in \mathcal{F}$, l'application $x \to \mathbb{P}_x(B)$ est $\mathcal{E}$-mesurable.

Preuve : La formule (3.5) détermine de façon unique $\mathbb{P}_x$ sur l'algèbre de Boole $\mathcal{A} = \cup_{n\geq 0}\mathcal{F}_n$. Il est clair que $\mathbb{P}_x$ ainsi définie est une probabilité sur $(\Omega, \mathcal{F}_n)$, $\forall n \geq 0$, et donc que $\mathbb{P}_x$ est une fonction additive d'ensemble sur $\mathcal{A}$.

Pour appliquer le théorème de prolongement, il reste à montrer que $\mathbb{P}_x$ est σ-additive sur $\mathcal{A}$ ou encore que, pour toute suite $(A_n)_{n\geq 0}$ d'éléments de $\mathcal{A}$ décroissant vers l'ensemble vide, la suite $(\mathbb{P}_x(A_n))_{n\geq 0}$ décroit vers 0.

Pour tout entier $n \geq 0$, soit $q(n)$ un entier tel que $A_n \in \mathcal{F}_{q(n)}$. En posant $\phi(n) = \max\{q(0), \cdots, q(n)\}$, $\forall n \geq 0$, on obtient une suite croissante d'entiers telle que $A_n \in \mathcal{F}_{\phi(n)}$.

Si $\lim_n \phi(n) = m \in \mathbb{N}$, alors on applique le fait que $\mathbb{P}_x$ est une probabilité sur l'espace $(\Omega, \mathcal{F}_m, \mathbb{P}_x)$. Supposons donc que $\lim_n \phi(n) = +\infty$. Nous allons raisonner par l'absurde, en supposant que $\lim_n \downarrow \mathbb{P}_x(A_n) > 0$.

Pour tout entier $p \geq 0$, nous avons :

$$\begin{aligned} &\mathbb{P}_x(A_n) \\ &= \int_E \cdots \int_E \Big(\int_\Omega 1_{A_n}(x, u_1, \cdots, u_p, \omega)\ \mathbb{P}_{u_p}(d\omega)\Big)\ P(u_{p-1}, du_p) \cdots P(x, du_1), \end{aligned}$$

et par suite (théorème de convergence monotone)

$$\begin{aligned} &\lim_n \downarrow \mathbb{P}_x(A_n) \\ =\ &\int_E \cdots \int_E \Big[\lim_n \downarrow \int_\Omega 1_{A_n}(x, u_1, \cdots, u_p, \omega)\ \mathbb{P}_{u_p}(d\omega)\Big]\ P(u_{p-1}, du_p) \cdots P(x, du_1). \end{aligned}$$

Puisque $\lim_n \downarrow \mathbb{P}_x(A_n) > 0$, il existe $(u_1, \cdots, u_p) \in E^p$ tel que

$$\lim_n \downarrow \int_\Omega 1_{A_n}(x, u_1, \cdots, u_p, \omega)\ \mathbb{P}_{u_p}(d\omega) > 0.$$

Cette remarque permet de définir par récurrence une suite $(u_n)_{n\geq 0}$ (avec $u_0 = x$) d'éléments de E telle que $(u_0, u_1, \cdots) \in A_n$, $\forall n \geq 0$, ce qui contredit le fait que $\cap_{n\geq 0} A_n = \emptyset$.

L'application $x \to \mathbb{P}_x(B)$ est mesurable lorsque B est dans $\mathcal{A}$, donc aussi pour B dans $\mathcal{E}$, d'après le théorème des classes monotones. $\square$

Pour $x \in E$, on note $\mathbb{E}_x$ l'espérance associée à la probabilité $\mathbb{P}_x$.

Définition 3.4.2 Soit $(Y_n)_{n\geq 0}$ un processus défini sur un espace probabilisé $(W, \mathcal{W}, \mathbb{Q})$ à valeurs dans un espace mesurable $(E, \mathcal{E})$. On dit que le processus $(Y_n)_{n\geq 0}$ est une **chaîne de Markov homogène** de probabilité de transition P, relativement à une filtration $(\mathcal{W}_n)_{n\geq 0}$, si
1) $(Y_n)_{n\geq 0}$ est adapté à la filtration $(\mathcal{W}_n)_{n\geq 0}$,
2) pour toute fonction f sur E, $\mathcal{E}$-mesurable positive ou bornée, on a $\mathbb{Q}$-p.s. :

$$\mathbb{E}_{\mathbb{Q}}[f(Y_{n+1})|\mathcal{W}_n] = Pf(Y_n), \ \forall n \geq 0. \tag{3.6}$$

La loi de la v.a. Y_0 est appelée **loi initiale** de la chaîne. Lorsque cette loi est la mesure de Dirac en un point x de E, on dit que la "chaîne part de x".

On dit que $(Y_n)_{n\geq 0}$ est une chaîne de Markov, si c'est une chaîne de Markov relativement à sa filtration naturelle. Il est facile de voir que la condition 2) de la définition (3.4.2) est équivalente à la propriété suivante :
pour tout $n \geq 0$, pour toute application g mesurable de $(E^{n+1}, \otimes_{n+1}\mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

$$\begin{aligned}&\mathbb{E}_Q[g(Y_0, \cdots, Y_n)] \\ &= \int_E \cdots \int_E g(u_0, u_1, \cdots, u_n)\ P(u_{n-1}, du_n) \cdots P(u_0, du_1)\ \nu(du_0),\end{aligned}$$

où ν désigne la loi de Y_0.

Avec ces définitions, nous pouvons donc dire que, pour tout $x \in E$, le processus $(X_n)_{n\geq 0}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P}_x)$ est une **chaîne de Markov de probabilité de transition P, partant de x.**

Pour toute mesure de probabilité ν sur $(E, \mathcal{E})$, posons

$$\mathbb{P}_\nu[B] = \int_E \mathbb{P}_x[B]\ \nu(dx), \ B \in \mathcal{F}.$$

Nous obtenons ainsi une mesure de probabilité sur $\mathcal{F}$. Le processus (X_n) défini sur $(\Omega, \mathcal{F}, \mathbb{P}_\nu)$ est une chaîne de Markov de probabilité de transition P, de loi initiale ν. La probabilité $\mathbb{P}_\nu$ est l'unique probabilité sur $(\Omega, \mathcal{F})$ vérifiant $\mathbb{E}_\nu[f(X_0)] = \nu(f)$ et $\mathbb{P}_\nu$-p.s. :

$$\mathbb{E}_\nu[f(X_{n+1})|\mathcal{F}_n] = Pf(X_n), \forall n \geq 1. \tag{3.7}$$

Pour toute probabilité ν sur $\mathcal{E}$, nous avons :

$$\theta(\mathbb{P}_\nu) = \mathbb{P}_{\nu P}. \tag{3.8}$$

La probabilité $\mathbb{P}_\nu$ est donc θ-invariante si, et seulement si, ν est P-invariante. Dans ce cas, le quadruplet $(\Omega, \mathcal{F}, \mathbb{P}_\nu, \theta)$ forme un système dynamique (i.e. un espace probabilisé muni d'une transformation mesurable laissant la mesure de probabilité invariante).

Le quadruplet $(\Omega, \mathcal{F}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ est appelé "chaîne de Markov **canonique**" associée à P.

• Propriétés de Markov d'une chaîne canonique

Proposition 3.4.3 *Pour toute fonction mesurable positive (ou bornée) Z sur $(\Omega, \mathcal{F})$, on a :*

$$\forall x \in E,\ \mathbb{E}_x[Z \circ \theta^n | \mathcal{F}_n] = \mathbb{E}_{X_n}[Z],\ \mathbb{P}_x - p.s..$$

Preuve : Il suffit de prouver le résultat pour les fonctions Z mesurables positives ne dépendant que d'un nombre fini de coordonnées.

Soit $Z = \varphi(X_0, \ldots, X_p)$, avec $p \in \mathbb{N}$. On est amené à montrer que, pour toute fonction f mesurable sur E^{n+1}, on a :

$$\mathbb{E}_x[\varphi(X_n, \ldots, X_{p+n})\ f(X_0, \ldots, X_n)] = \mathbb{E}_x[\mathbb{E}_{X_n}[\varphi(X_0, \ldots, X_p)]\ f(X_0, \ldots, X_n)].$$

On obtient cette égalité en appliquant la formule ((3.5) une fois au membre de gauche et deux fois au terme de droite. □

La propriété précédente s'étend aux temps d'arrêt (propriété de Markov forte).

Proposition 3.4.4 *Pour toute fonction mesurable positive (ou bornée) Z sur $(\Omega, \mathcal{F})$ et pour tout temps d'arrêt T relativement à la filtration $(\mathcal{F}_n)_{n \geq 0}$, on a :*

$$\forall x \in E,\ \mathbb{E}_x[Z \circ \theta^T\ 1_{\{T<+\infty\}} | \mathcal{F}_T] = \mathbb{E}_{X_T}[Z]\ 1_{\{T<+\infty\}},\ \mathbb{P}_x - p.s.. \tag{3.9}$$

Preuve : Pour $A \in \mathcal{F}_T$, nous avons, en appliquant la proposition précédente :

$$\begin{aligned}
\int_A Z \circ \theta^T\ 1_{\{T<+\infty\}}\ d\mathbb{P}_x &= \sum_{n \in \mathbb{N}} \int_A Z \circ \theta^n\ 1_{\{T=n\}}\ d\mathbb{P}_x \\
&= \sum_{n \in \mathbb{N}} \int_{A \cap \{T=n\}} Z \circ \theta^n\ d\mathbb{P}_x \\
&= \sum_{n \in \mathbb{N}} \int_{A \cap \{T=n\}} \mathbb{E}_{X_n}[Z]\ d\mathbb{P}_x \\
&= \sum_{n \in \mathbb{N}} \int_A \mathbb{E}_{X_T}[Z]\ 1_{\{T=n\}}\ d\mathbb{P}_x \\
&= \int_A \mathbb{E}_{X_T}[Z]\ 1_{\{T<+\infty\}}\ d\mathbb{P}_x.
\end{aligned}$$

□

• Chaîne des "sauts"

Dans les exemples d'opérateurs de la forme (3.3) avec $S = \{s_i : i \in I\}$ pour I fini ou dénombrable, il est commode de remplacer la chaîne de Markov canonique construite précédemment par la chaîne des sauts obtenue de la façon suivante.

Considérons l'espace produit $\Omega = S^{\mathbb{N}^*}$ muni de sa tribu produit $\mathcal{F}$. Appelons $Y_n, n \geq 1$, les applications coordonnées de Ω. Pour tout $x \in E$, appelons $\mathbb{P}_x$ la probabilité définie sur $\mathcal{F}$ par :

$$\begin{aligned}
&\int f(Y_1(\omega), \ldots, Y_n(\omega)) \mathbb{P}_x(d\omega) \\
&\quad = \sum_{(s_1,\ldots,s_n) \in S^n} p(s_n, s_{n-1} \cdots s_1 \cdot x) \cdots p(s_1, x)\ f(s_1, \ldots, s_n),
\end{aligned}$$

pour tout entier $n \geq 1$ et toute fonction f définie sur S^n.

Nous désignons par $\mathcal{F}_0$ la tribu triviale et par $\mathcal{F}_n, n \geq 1$, la tribu engendrée par les variables aléatoires $\{Y_k, 1 \leq k \leq n\}$.

Pour $x \in E$, nous définissons : $Z_0(x, \omega) = x$, et, pour $n \geq 1$, $Z_n(x, \omega) = Y_n(\omega) \cdots Y_1(\omega) \cdot x$. Le processus $\{Z_n(x, .), n \geq 0\}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P}_x)$ est une chaîne de Markov sur E de probabilité de transition P, partant de x.

Pour la chaîne $(\Omega, \mathcal{F}, (Z_n(x, .)_{n \geq 0}, \mathbb{P}_x)_{x \in E})$, la propriété de Markov s'écrit :

$$\mathbb{E}_x[V(\theta^n .)|\mathcal{F}_n] = \mathbb{E}_{Z_n(x,.)}[V(.)], \mathbb{P}_x - p.s.$$

pour tout $x \in E$ et toute variable aléatoire bornée V sur Ω.

3.5 Fonctions harmoniques et probabilités P-invariantes

Soit P une probabilité de transition sur un espace mesurable $(E, \mathcal{E})$.

• **Fonctions P-harmoniques et variables aléatoires θ-invariantes**

Notations : Nous notons $\mathcal{H}$ l'espace vectoriel des fonctions $\mathcal{E}$-mesurables P-harmoniques bornées.

Nous désignons par $\mathcal{J} = \{A \in \mathcal{F} : \theta^{-1}(A) = A\}$ la tribu des boréliens θ-invariants de Ω. Une variable aléatoire Z est $\mathcal{J}$-mesurable si, et seulement si, $Z \circ \theta = Z$.

Proposition 3.5.1 *(cf. J. Neveu [Nev64]) Pour toute fonction h de $\mathcal{H}$ et tout $x \in E$, le processus $(h(X_n))_{n \geq 0}$, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}_x)$, est une martingale bornée relativement à la filtration $(\mathcal{F}_n)_{n \geq 0}$. Ce processus converge $\mathbb{P}_x$-p.s. vers la variable aléatoire $Z = \limsup_n h(X_n)$, qui est θ-invariante et telle que, pour tout entier $n \geq 0$:*

$$h(x) = \mathbb{E}_x[h(X_n)] = \mathbb{E}_x[Z] \quad \text{et} \quad h(X_n) = \mathbb{E}_x[Z \mid \mathcal{F}_n] \quad \mathbb{P}_x\text{-}p.s..$$

Réciproquement, si Z est une variable aléatoire bornée θ-invariante, alors la fonction h, définie par $h(x) = \mathbb{E}_x[Z], x \in E$, appartient à $\mathcal{H}$ et la martingale $(h(X_n))_{n \geq 0}$ converge $\mathbb{P}_x$-p.s., $\forall\, x \in E$, vers Z.

<u>Preuve</u> : Pour tout entier naturel n, nous avons :

$$\begin{aligned} \mathbb{E}_x[h(X_{n+1}) \mid \mathcal{F}_n] &= Ph(X_n) \quad \text{(propriété de chaîne de Markov)} \\ &= h(X_n) \quad \text{(invariance de } h \text{ par } P); \end{aligned}$$

Le processus $(h(X_n))_{n \geq 0}$, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}_x)$, est donc une martingale relativement à la filtration $(\mathcal{F}_n)_{n \geq 0}$, bornée par $\|h\|_\infty$. D'après la théorie des martingales, on sait alors que ce processus converge, $\mathbb{P}_x$-p.s., et sa limite Z "ferme la martingale" ; c'est-à-dire $h(X_n) = \mathbb{E}_x[Z \mid \mathcal{F}_n]$ $\mathbb{P}_x$-p.s.. Par passage à l'espérance, cette dernière relation nous donne $\mathbb{E}_x[h(X_n)] = \mathbb{E}_x[Z]$. D'autre part, nous avons $\mathbb{E}_x[h(X_n)] = P^n h(x) = h(x)$. D'où la première assertion.

Réciproquement soient Z une variable aléatoire bornée invariante par θ et h la fonction définie par $h(x) = \mathbb{E}_x[Z]$. D'après la propriété de Markov, nous avons, pour tout entier naturel n,

$$h(X_n) = \mathbb{E}_{X_n}[Z] = \mathbb{E}_x[Z \circ \theta^n \mid \mathcal{F}_n] \quad \mathbb{P}_x\text{-p.s.};$$

et donc grâce à l'invariance de Z par θ, $h(X_n) = \mathbb{E}_x[Z \mid \mathcal{F}_n]$. Par passage à l'espérance, il vient $P^n h(x) = \mathbb{E}_x[Z]$, pour tout entier naturel n ; d'où la P-harmonicité de h. Enfin, d'après le théorème de convergence des martingales (chap. I), le processus $(\mathbb{E}_x[Z \mid \mathcal{F}_n])_{n\geq 0}$ converge $\mathbb{P}_x$-p.s. vers Z. □

Cette correspondance entre fonctions P-harmoniques bornées et variables aléatoires θ-invariantes, permet de munir $\mathcal{H}$ d'une structure de C^*-algèbre. Si h_1 et h_2 sont deux fonctions de $\mathcal{H}$, on définit leur produit par :

$$h_1 * h_2(x) = \mathbb{E}_x[\lim_{n\to+\infty} h_1(X_n)\ h_2(X_n)\ ] = \lim_{n\to+\infty} P^n(h_1 h_2)\ (x).$$

Elle permet aussi d'énoncer le lemme 3.3.6 de la façon suivante :

Corollaire 3.5.2 *Une mesure de probabilité P-invariante ν est extrémale si et seulement si la tribu $\mathcal{J}$ des boréliens θ-invariants de Ω est $\mathbb{P}_\nu$-presque sûrement égale à la tribu grossière.*

Dans la situation du corollaire ,on dit que le quadruplet $(\Omega, \mathcal{F}, \theta, \mathbb{P}_\nu)$ est un système dynamique ergodique. Nous développerons plus loin (chapitre 8) cette notion.

Lemme 3.5.3 *Soit h une fonction de $\mathcal{H}$. Pour tout entier $r \geq 1$ et tout $x \in E$, nous avons :*

$$\sum_{n\geq 0} \mathbb{E}_x\Big[\int_E [h(y) - h(X_n)]^2\, P^r(X_n, dy)\Big]\ <\ +\infty.$$

<u>Preuve</u> : Le résultat est une conséquence de la relation

$$\mathbb{E}_x\Big[\int_E [h(y) - h(X_n)]^2\, P^r(X_n, dy)\Big] = \mathbb{E}_x[h^2(X_{n+r})] - \mathbb{E}_x[h^2(X_n)].$$

□

• **Mesures P-invariantes**

On peut montrer (? ?cf. partie II, Théorie ergodique) que, si ν est P-invariante extrémale, alors pour toute fonction f de $L^1(\mathbb{P}_\nu)$, on a $\mathbb{P}_\nu$-p.s. :

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} f(X_{k-1}) = \nu(f).$$

Sur un espace métrique compact, dans le cas où ν est unique, nous allons établir une "loi des grands nombres" au sens des probabilités $\mathbb{P}_x$, pour les fonctions continues.

Proposition 3.5.4 *Soient (E, d) un espace métrique compact et P un opérateur de transition sur $(E, \mathcal{B}(E))$, préservant l'espace $\mathcal{C}(E)$ des fonctions continues. Supposons qu'il existe sur $(E, \mathcal{B}(E))$ une unique mesure de probabilité P-invariante ν. Alors, pour toute fonction f de $\mathcal{C}(E)$ et pour tout $x \in E$, nous avons :*

$$\lim_n \frac{1}{n} \sum_{k=1}^{n} f(X_{k-1}) = \nu(f),\ \ \mathbb{P}_x\text{-p.s. et dans } L^1(\mathbb{P}_x).$$

Preuve : Soit $(\phi_i)_{i\geq 1}$ une suite de fonctions dans $\mathcal{C}(E)$ dense pour la norme uniforme. Soit Ω_1 l'ensemble des ω tels que, pour tout $i \geq 1$,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}(\phi_i(X_k(\omega)) - P\phi_i(X_{k-1}(\omega)) = 0.$$

Les sommes portant sur des accroissements de martingale bornés, d'après la loi des grands nombres pour les accroissements de martingale, l'ensemble Ω_1 est de $\mathbb{P}_x$-mesure égale à 1, quelque soit $x \in E$. Pour tout $\omega \in \Omega_1$ et pour tout $f \in \mathcal{C}(E)$, on a alors la convergence

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}(f(X_k(\omega)) - Pf(X_{k-1}(\omega)) = 0.$$

Soit $\omega \in \Omega_1$. Appelons ν_ω une valeur d'adhérence de la suite de probabilité $(\frac{1}{n}\sum_{k=1}^{n}\delta_{X_k(\omega)})$. Les convergences précédentes assurent que ν_ω est P-invariante et donc égale à ν. Ceci entraîne la convergence de la suite de probabilités $(\frac{1}{n}\sum_{k=1}^{n}\delta_{X_k(\omega)})$ vers ν. □

- **Fonctions harmoniques pour la chaîne espace-temps**

(cf. D. Revuz [Rev75])

Nous appelons Q la probabilité de transition sur l'espace produit $E \times \mathbb{N}$ définie par :

$$QF(x,n) = \int_E F(y,n+1)\,P(x,dy), \quad (x,n) \in E \times \mathbb{Z}.$$

Nous désignons par $\mathcal{K}$ l'espace vectoriel des fonctions $\mathcal{E}\otimes\mathcal{P}(\mathbb{Z})$-mesurables bornées Q-harmoniques sur $E \times \mathbb{Z}$. Tout élément H de $\mathcal{K}$, est une suite bornée $(h_n = H(\cdot,n))_{n\in\mathbb{Z}}$ de fonctions $\mathcal{E}$-mesurables sur E telle que

$$\forall\, n \in \mathbb{Z},\ Ph_n = h_{n-1}.$$

Comme pour les fonctions P-harmoniques, on définit le produit de deux fonctions de $\mathcal{K}$ par :

$$H_1 * H_2 = \lim_{n\to+\infty} Q^n(H_1H_2). \tag{3.10}$$

L'espace $\mathcal{K}$ est une algèbre de Banach quand on le munit de ce produit et de la norme uniforme $\|\ \|_\infty$, définie par

$$\|H\|_\infty = \sup_{n\in\mathbb{Z}} \|H(\cdot,n)\|_\infty.$$

A tout élément h de $\mathcal{H}$ est associé un élément $H = (h_k)_{k\in\mathbb{Z}}$ de $\mathcal{K}$ défini par $h_k = h$, pour tout entier relatif k. Cette correspondance permet d'identifier l'algèbre $\mathcal{H}$ à une sous-algèbre de $\mathcal{K}$.

- **Fonctions Q-harmoniques et variables aléatoires asymptotiques**

Pour tout entier $n \geq 0$, nous notons $\mathcal{G}_n$ la sous-tribu de $\mathcal{F}$ engendrée par les variables aléatoires $\{X_k : k \geq n\}$.

Nous appelons *tribu asymptotique* la tribu $\mathcal{A} = \bigcap_{n\geq 0} \mathcal{G}_n$. La tribu $\mathcal{J}$ des boréliens θ-invariants de Ω est contenue dans la tribu asymptotique.

Une variable aléatoire U est mesurable par rapport à la tribu $\mathcal{A}$ si et seulement si, pour tout entier naturel $n \geq 1$, il existe une variable aléatoire U_n telle que $U = U_n \circ \theta^n$; une telle variable sera dite *asymptotique*. Soit n un entier supérieur où égal à 1. Pour tous éléments $x_1, \ldots, x_n$ de E et tout $\omega \in \Omega$, il vient :

$$\forall\, k \geq 1, U(x_1, \ldots, x_n, \omega) = U_n(\omega) = U_{n+k} \circ \theta^k(\omega).$$

Il s'ensuit que les variables aléatoires $\{U_n \,:\, n \geq 1\}$ sont déterminées de façon unique et asymptotiques. Nous écrirons donc $\forall\, n \geq 1$, $U_n = U \circ \theta^{-n}$.

Par une démonstration analogue à celle de la proposition 3.5.1, on montre :

Proposition 3.5.5 *Pour tout $H \in \mathcal{K}$ et tout $(x, k) \in E \times \mathbb{Z}$, le processus $(H(X_n, n+k))_{n\geq 0}$, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}_x)$, est une martingale bornée relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Ce processus converge $\mathbb{P}_x$-p.s. vers une variable aléatoire asymptotique $U \circ \theta^{-k}$, vérifiant, pour tout entier $n \geq 0$,*

$$H(x, k) = \mathbb{E}_x[H(X_n, n+k)] = \mathbb{E}_x[U \circ \theta^{-k}]$$

et

$$H(X_n, n+k) = \mathbb{E}_x[U \circ \theta^{-k} \mid \mathcal{F}_n] \quad \mathbb{P}_x\text{-}p.s.$$

Réciproquement si U est une variable aléatoire asymptotique bornée, alors la fonction H, définie par $H(x, k) = \mathbb{E}_x[U \circ \theta^{-k}]$ $((x, k) \in E \times \mathbb{Z})$, appartient à $\mathcal{K}$ et la martingale $(H(X_n, k+n))_{n\geq 0}$ converge, $\mathbb{P}_x$-p.s. $\forall\, x \in E$, vers $U \circ \theta^{-k}$.

Lemme 3.5.6 *Soit H une fonction de $\mathcal{K}$. Pour tout $r \geq 1$, tout $k \in \mathbb{Z}$ et tout $x \in E$, nous avons*

$$\sum_{n\geq 0} \mathbb{E}_x\Big[\int_E [H(y, r+n+k) - H(X_n, n+k)]^2\, P^r(X_n, dy)\Big] \;<\; +\infty.$$

<u>Preuve</u> : La démonstration est identique à celle du lemme 3.5.3. □

3.6 Temps d'entrée

Soit P une probabilité de transition sur un espace mesurable $(E, \mathcal{E})$. Nous appelons G le noyau potentiel $G = \sum_{k\geq 0} P^k$.

Pour tout $x \in E$, $G(x, \cdot)$ est une mesure positive sur $(E, \mathcal{E})$, qui n'est pas nécessairement σ-finie.

Définition 3.6.1 Une fonction f sur E, à valeurs dans $[-\infty, +\infty]$, $\mathcal{E}$-mesurable, est dite P-**surharmonique**, si elle est quasi-intégrable pour toutes les probabilités $P(x, \cdot)$ et si l'on a :

$$\forall x \in E,\ Pf(x) = \int_E f(y)\ P(x, dy) \leq f(x).$$

La fonction f est dite P-**sous-harmonique** si la fonction $-f$ est surharmonique.

Enfin, nous dirons que f est un **potentiel**, s'il existe une fonction u $\mathcal{E}$-mesurable de E dans $[-\infty, +\infty]$, quasi-intégrable pour toutes les mesures $G(x, \cdot)$, $x \in E$, telle que $\forall x \in E,\ f(x) = Gu(x)$.

• **Décomposition de Riesz d'une fonction surharmonique positive**.

Lemme 3.6.2 *Soit f une fonction surharmonique positive finie, alors f s'écrit de façon unique $f = h + Gu$, où h est une fonction P-harmonique positive et u une fonction $\mathcal{E}$-mesurable positive. De plus, nous avons :*

$$u = f - Pf \quad \text{et} \quad h = \lim_{n \to +\infty} \downarrow \; P^n f.$$

Preuve : 1) La fonction f étant surharmonique, la suite de fonctions $(P^n f)_{n \geq 0}$ est décroissante. Appelons h la limite simple sur E de cette suite de fonctions. Du théorème de convergence monotone, il résulte que h est harmonique. En posant $u = f - Pf$, nous avons :

$$\forall N \geq 0, \; \sum_{k=0}^{N} P^k u = f - P^{N+1} f$$

qui converge, quand N tend vers $+\infty$ vers la fonction $f - h$. D'où $Gu = f - h$.

2) Soit $f = g + Gv$ une décomposition de f en la somme d'une fonction harmonique positive g et d'un potentiel Gv, avec v positive $\mathcal{E}$-mesurable. Pour tout entier $n \geq 0$, $P^n f = g + P^n Gv$. Mais la suite de fonctions $(P^n Gv = \sum_{k \geq n} P^k v)_{n \geq 0}$ décroît vers la fonction nulle. Par suite g est la limite décroissante de la suite $(P^n f)_{n \geq 0}$; autrement dit $g = h$. On en déduit que $Gu = Gv$ et $u = (I - P)Gu = (I - P)Gv = v$. □

• **Temps d'entrée**

Désignons par $(\Omega, \mathcal{F}, (X_n)_{n \geq 0}, (\mathbb{P}_x)_{x \in E})$ la chaîne de Markov canonique associée à la probabilité de transition P. Pour $A \in \mathcal{E}$, on définit les v.a.

$$T_A = \inf\{n \geq 0 : X_n \in A\} \quad \text{(temps d'entrée dans } A)$$

$$S_A = \inf\{n \geq 1 : X_n \in A\} \quad \text{(temps de retour dans } A)$$

$$N_A = \sum_{n \geq 0} 1_A \circ X_n \quad \text{(nombre de passages dans } A).$$

T_A et S_A sont des temps d'arrêt relativement à la filtration naturelle $(\mathcal{F}_n^X)_{n \geq 0}$ associée au processus $(X_n)_{n \geq 0}$. Pour tout $x \in A$, $T_A = 0$, $\mathbb{P}_x$-p.s. ; pour tout $x \notin A$, $T_A = S_A$, $\mathbb{P}_x$-p.s. ; $S_A \geq T_A$ et $S_A = T_A \circ \theta$. Posons, pour $A \in \mathcal{E}$ et $x \in E$,

$$f_A(x) = \mathbb{P}_x[\{T_A < +\infty\}] \quad \text{et} \quad h_A(x) = \mathbb{P}_x[\{N_A = +\infty\}].$$

Le réel positif $f_A(x)$ représente la probabilité, partant de x, que la chaîne passe au moins une fois dans A ; $f_A(x) = \mathbb{P}_x[\{N_A \geq 1\}]$.

Le réel positif $h_A(x)$ représente la probabilité, partant de x, que la chaîne passe dans A une infinité de fois.

Proposition 3.6.3 *Soit $A \in \mathcal{E}$. La fonction f_A est surharmonique positive et sa partie harmonique (dans la décomposition de Riesz) est h_A. Pour tout $x \in E$, la surmartingale $(f_A(X_n))_{n \geq 0}$ et la martingale $(h_A(X_n))_{n \geq 0}$ convergent $\mathbb{P}_x$-p.s. vers la v.a. θ-invariante $1_{\{N_A = +\infty\}}$.*

Preuve. Nous avons, pour tout entier $n \geq 1$,

$$\begin{aligned} P^n f_A(x) &= \mathbb{E}_x[f_A(X_n)] = \mathbb{E}_x[\mathbb{P}_{X_n}[\{T_A < +\infty\}]] \\ &= \mathbb{E}_x[\mathbb{E}_x[1_{\{T_A\circ\theta^n<+\infty\}}|\mathcal{F}_n^X]] \quad \text{(propriété de Markov)} \\ &= \mathbb{P}_x[\{T_A \circ \theta^n < +\infty\}]. \end{aligned}$$

Pour $n = 1$, il s'ensuit que

$$Pf_A(x) = \mathbb{P}_x[\{T_A \circ \theta < +\infty\}] \leq \mathbb{P}_x[\{T_A < +\infty\}] = f_A(x)$$

car $T_A \circ \theta = S_A \geq T_A$. Ce qui montre que la fonction f_A est bien surharmonique.

D'autre part, la suite d'ensembles $\{T_A \circ \theta^n < +\infty\}$ est décroissante et

$$\bigcap_{k\geq 0}\{T_A \circ \theta^k < +\infty\} \;=\; \{N_A = +\infty\}.$$

Il en résulte que la suite $(P^n f_A(x))$ décroît vers $h_A(x)$.

La convergence, $\mathbb{P}_x$-p.s., de la martingale $(h_A(X_n))_{n\geq 0}$ vers $1_{\{N_A=+\infty\}}$ résulte de la proposition 3.5.1.

Appelons Z la limite $\mathbb{P}_x$-p.s. de la surmartingale positive bornée $(f_A(X_n))_{n\geq 0}$. Comme $f_A \geq h_A$, nous avons

$$Z \geq 1_{\{N_A=+\infty\}} \quad \mathbb{P}_x\text{-p.s.}.$$

D'autre part, d'après le lemme de Fatou,

$$\mathbb{E}_x[Z] \leq \liminf_{n\to+\infty} \mathbb{E}_x[f_A(X_n)] = \liminf_{n\to+\infty} P^n f_A(x) = h_A(x) = \mathbb{P}_x[\{N_A = +\infty\}] \; ;$$

d'où :

$$Z = 1_{\{N_A=+\infty\}} \quad \mathbb{P}_x\text{-p.s.}.$$

□

Chapitre 4

Chaînes de Markov à espace d'états discrets

4.1 Classification des états

• Notations

Soit P une probabilité de transition sur un ensemble dénombrable E muni de la tribu $\mathcal{P}(E)$ de ses parties. Pour tous $x, y \in E$, le réel positif $P(x, \{y\})$ sera noté plus simplement $P(x, y)$; nous avons :

$$\forall x \in E, \sum_{y \in E} P(x, y) = 1.$$

La chaîne de Markov canonique associée à P (cf. ? ?II.4) est notée

$$(\Omega = E^{\mathbb{N}},\ \mathcal{F} = \otimes_{\mathbb{N}} \mathcal{P}(E),\ (X_n)_{n \geq 0},\ (\mathbb{P}_x)_{x \in E}).$$

Les éléments de E sont appelés "états" de la chaîne.

Nous notons P^n, $n \geq 0$, les itérés de P :

$$\forall n \geq 1,\ P^{n+1}(x, y) = \sum_{z \in E} P(x, z)\ P^n(z, y) = \sum_{z \in E} P^n(x, z)\ P(z, y),$$

en convenant que $P^0(x, y) = \delta_x^y$ (symbole de Kronecker), et G le noyau potentiel associé à P défini par

$$\forall x, y \in E,\ G(x, y) = \sum_{n \geq 0} P^n(x, y).$$

Nous désignons par T_x et S_x les temps d'entrée et de retour en x :

$$T_x = \inf\ \{n \geq 0 : X_n = x\}, \quad S_x = \inf\ \{n \geq 1 : X_n = x\}$$

(avec la convention $\inf \emptyset = +\infty$), et par N_x le nombre de passage de la chaîne en x :

$$N_x = \sum_{n \geq 0} 1_{\{x\}}(X_n) = \sum_{n \geq 0} 1_{\{X_n = x\}}.$$

Remarquons que, pour tous $x, y \in E$, nous avons $G(x, y) = \mathbb{E}_x[N_y]$.

Nous posons

$$\forall y \in E,\ f_y(x) = \mathbb{P}_x[\{T_y < +\infty\}] \text{ et } h_y(x) = \mathbb{P}_x[\{N_y = +\infty\}].$$

Pour tout $y \in E$, la fonction f_y est P-surharmonique, de partie harmonique h_y. Les processus $(f_y(X_n))_{n\geq 0}$ et $(h_y(X_n))_{n\geq 0}$, définis sur $(\Omega, \mathcal{F})$, convergent, $\mathbb{P}_z - p.s.$ $\forall z \in E$, vers $1_{\{N_y=+\infty\}}$ (cf. ? ?(II.6)).

• Classification des états

Définition 4.1.1 Soient $x, y \in E$ deux états de la chaîne. On dit que "x **mène à** y", et on écrit "$x \to y$", s'il existe un entier $n \geq 0$ tel que $P^n(x,y) > 0$ ou, de façon équivalente, si $G(x,y) > 0$.

Lemme 4.1.2 *Soient x et y deux états. Nous avons :*

$$"x \to y" \iff f_y(x) = \mathbb{P}_x[\{T_y < +\infty\}] > 0.$$

Preuve : L'implication "$\Rightarrow$" résulte des inclusions :

$$\forall n \geq 0,\ \{X_n = y\} \subset \{T_y \leq n\} \subset \{T_y < +\infty\}$$

et l'implication "$\Leftarrow$" de l'égalité :

$$\{T_y < +\infty\} = \bigcup_{k\geq 0} \{X_k = y\}.$$

□

Définition 4.1.3 Soient x et y deux états. On dit que x et y **communiquent**, et on écrit "$x \leftrightarrow y$", si x mène à y et y mène à x.

Lemme 4.1.4 *Sur E, la relation binaire "$\to$" est transitive et la relation binaire "$\leftrightarrow$" est une relation d'équivalence.*

Preuve : Exercice.

Définition 4.1.5 La chaîne est dite **irréductible** si tous les états de E communiquent. Nous avons alors $\forall x, y \in E,\ G(x,y) > 0$.

En général, E se décompose en classes d'équivalence deux à deux disjointes. Partant d'un état x appartenant à une classe C :
− ou bien la chaîne évolue dans C (dans ce cas $\forall y \in C,\ \sum_{z\in C} P(y,z) = 1$ et la restriction de la chaîne à C est irréductible) ;
− ou bien la chaîne quitte C et dans ce cas elle n'y revient plus.

• Période d'une classe d'équivalence

Définition 4.1.6 On appelle **période d'un état** x, l'entier $d(x)$ défini par :

$$d(x) = p.g.c.d.\{n \geq 1 : P^n(x,x) > 0\}.$$

On convient que $p.g.c.d.\ \emptyset = 0$. On notera que $d(x) = 0$ équivaut à dire que x ne communique qu'avec lui-même.

Lemme 4.1.7 *Si x et y sont deux états de E qui communiquent, alors $d(x) = d(y)$.*

Preuve : Il existe des entiers naturels p et q tels que

$$P^p(x, y) > 0 \quad \text{et} \quad P^q(y, x) > 0.$$

Pour tout entier naturel n, nous avons

$$P^{p+n+q}(x, x) \geq P^p(x, y)\ P^n(y, y)\ P^q(y, x).$$

Si $P^n(y, y) > 0$ alors $P^{2n}(y, y) \geq (P^n(y, y))^2 > 0$ et par suite

$$P^{p+n+q}(x, x) > 0 \text{ et } P^{p+2n+q}(x, x) > 0.$$

On en déduit que $d(x)$ divise les entiers $p+n+q$ et $p+2n+q$ et donc $d(x)$ divise n. Ce qui prouve que $d(x)$ divise $d(y)$. En échangeant les rôles de x et y, on obtient que $d(y)$ divise $d(x)$. □

Définition 4.1.8 On appelle **période** d'une classe d'équivalence C de la chaîne, la période commune à chacun de ses éléments.

Une classe C de la chaîne est dite **apériodique** si sa période est égale à 1.

• **Sous-classes cycliques d'une classe d'équivalence de période** > 1

Soit C une classe d'équivalence de la chaîne de période d strictement supérieure à 1.

Introduisons les relations binaires sur E, notées "$\overset{d}{\rightarrow}$" et "$\overset{d}{\leftrightarrow}$", associée à la probabilité de transition P^d :

$$x \overset{d}{\rightarrow} y \Leftrightarrow G^{(d)}(x, y) = \sum_{n \geq 0} P^{nd}(x, y) > 0,$$

et $x \overset{d}{\leftrightarrow} y$ si et seulement si $x \overset{d}{\rightarrow} y$ et $y \overset{d}{\rightarrow} x$.

Il est clair que $x \overset{d}{\leftrightarrow} y$ implique $x \leftrightarrow y$ et la relation $\overset{d}{\leftrightarrow}$ induit sur C une relation d'équivalence. Nous avons le résultat suivant :

Lemme 4.1.9 *La relation d'équivalence "$\overset{d}{\leftrightarrow}$" découpe C en d sous-classes non vides, $C_0, \dots C_{d-1}$, telles que :*

$$\forall 0 \leq \ell \leq d-1,\ \forall x \in C_\ell,\ P(x, C \setminus C_{\ell+1}) = 0\ ;$$

avec la convention $C_d = C_0$.

Preuve : Fixons un état x_0 de C. Pour tout $\ell \in \{0, \dots, d-1\}$, posons

$$C_\ell = \{y \in C : P^\ell G^{(d)}(x_0, y) = \sum_{k \geq 0} P^{kd+\ell}(x_0, y) > 0\}.$$

Il est clair que C est la réunion des C_ℓ, $0 \leq \ell \leq d-1$.

Considérons deux entiers ℓ et j de $\{0, \dots, d-1\}$. Soient $x \in C_\ell$, $y \in C_j$ et m un entier tel que $P^m(x, y) > 0$. Appelons p, q et r des entiers tels que

$$P^{pd+\ell}(x_0, x) > 0\ ; P^{qd+j}(x_0, y) > 0 \text{ et } P^r(y, x_0) > 0.$$

Nous avons :

$$P^{pd+\ell+m+r}(x_0, x_0) \geq P^{pd+\ell}(x_0, x)\ P^m(x, y)\ P^r(y, x_0)\ >\ 0$$

et de même $P^{qd+j+r}(x_0, x_0) > 0$. Il s'ensuit que les entiers $pd+\ell+m+r$ et $qd+j+r$ sont divisibles par la période d de x_0 ; ce qui implique que l'entier $m+\ell-j$ est divisible par d.

On en déduit que les C_ℓ sont des sous-classes d'équivalences de C pour la relation "$\overset{d}{\leftrightarrow}$". En effet, si $x \in C_\ell$ et $y \in C_j$ vérifient $x \overset{d}{\to} y$, alors on doit avoir $\ell = j$. On en déduit, qu'au sens de P^d, tout état d'une sous-classe C_ℓ ne communique qu'avec des états de C_ℓ ; d'où le résultat. (Rappelons qu'en général un état de C peut mener à un état n'appartenant pas à C). □

Définition 4.1.10 Les sous-classes $C_0, \ldots C_{d-1}$ sont appelées **sous-classes cycliques** de C.

Exercice 4.1.11

On suppose que la chaîne est irréductible et de période $d > 1$. On note $C_0, \ldots C_{d-1}$ les sous-classes cycliques de C. Pour tout $\ell \geq 0$, on pose $C_\ell = C_{\ell \bmod d}$.

1) Pour toute fonction bornée sur E, montrer que

$$\forall \ell \geq 0,\ P(1_{C_\ell}\ f) = 1_{C_{\ell-1}}\ Pf.$$

2) Si f est une fonction P-harmonique bornée sur E, montrer que la fonction

$$H(x, \ell) = 1_{C_\ell}(x)\ f(x), \quad (x, \ell) \in E \times \mathbb{N}$$

vérifie

$$\forall (x, \ell) \in C \times \mathbb{N},\ H(x, \ell) = H(x, \ell + d)$$

et est harmonique pour la chaîne espace-temps associée à P.

3) Soit H une fonction harmonique pour la chaîne espace-temps associée à P telle que

$$\forall (x, \ell) \in E \times \mathbb{N},\ H(x, k) = H(x, k + d).$$

Pour toute racine d-ième de l'unité z, montrer que la fonction

$$h(x) = \frac{1}{d} \sum_{k=0}^{d-1} z^k\ H(x, k), \quad (x \in E)$$

vérifie $Ph = zh$ et $P^d h = h$.

4) Soit μ une mesure positive sur E.

a) Montrer que, pour tout $\ell \geq 0$,

$$(1_{C_{\ell-1}}\ \mu)P = 1_{C_\ell}\ \mu P.$$

b) Montrer que μ est P-invariante si et seulement si la mesure $\mu_0 = 1_{C_0}\ \mu$ est P^d-invariante. Dans ce cas montrer que l'on a $\mu = \sum_{0 \leq k \leq d-1}\ \mu_0\ P^k$.

c) Lorsque μ est une probabilité P-invariante, montrer que $\forall \ell \geq 0,\ \mu[C_\ell] = \frac{1}{d}$.

4.2 Propriété de récurrence

• **Récurrence et transience**

Proposition 4.2.1 *Soient u et v deux états tels que $h_v(u) = \mathbb{P}_u[\{N_v = +\infty\}] > 0$. Alors, pour tout $w \in E$, nous avons :*

$$\mathbb{P}_v[\{T_w < +\infty\}] = h_w(v) = \frac{\mathbb{P}_u[\{N_w = +\infty\} \cap \{N_v = +\infty\}]}{h_v(u)} = 0 \textit{ ou } 1.$$

En particulier, pour $w = v$, $h_v(v) = 1$.

<u>Preuve</u> : Nous savons (cf. ? ?(II.6)) que les processus $(f_w(X_n))_{n\geq 0}$ et $(h_w(X_n))_{n\geq 0}$ convergent $\mathbb{P}_z$-p.s. $\forall z \in E$ vers $1_{\{N_w=+\infty\}}$. Posons

$$A = \{N_v = +\infty\} \cap \{\lim_{n\to+\infty} f_w(X_n) = \lim_{n\to+\infty} h_w(X_n) = 1_{\{N_w=+\infty\}}\} \ ;$$

nous avons $\mathbb{P}_u[A] = \mathbb{P}_u[\{N_v = +\infty\}] > 0$.

Pour tout $\omega \in A$, les suites réelles $(f_w(X_n(\omega)))_{n\geq 0}$ et $(h_w(X_n(\omega)))_{n\geq 0}$ convergent vers $1_{\{N_w=+\infty\}}(\omega)$, mais aussi vers $f_w(v)$ et $h_w(v)$, car la suite $(X_n(\omega))_{n\geq 0}$ prend une infinité de fois la valeur v ; d'où :

$$\forall \omega \in A,\ f_w(v) = h_w(v) = 1_{\{N_w=+\infty\}}(\omega) = \frac{\mathbb{P}_u[\{N_w = +\infty\} \cap A]}{\mathbb{P}_u[A]}.$$

□

Définition 4.2.2 Un état $x \in E$ est dit **transitoire** si

$$h_x(x) = \mathbb{P}_x[\{N_x = +\infty\}] = 0.$$

Il est dit **récurrent** si

$$h_x(x) = \mathbb{P}_x[\{N_x = +\infty\}] = 1.$$

Une conséquence de la proposition 4.2.1 est la suivante.

Théorème 4.2.3 *Tout état $x \in E$ est soit transitoire, soit récurrent.*

Si x est récurrent, alors tout $y \in E$ tel que "$x \to y$" est aussi récurrent et nous avons :

$$\mathbb{P}_x[\{N_y = +\infty\}] = \mathbb{P}_y[\{N_x = +\infty\}] = 1.$$

Si E est fini, il existe au moins un état récurrent.

<u>Preuve</u> : Soit $x \in E$ tel que $h_x(x) > 0$. En appliquant la proposition 4.2.1, avec $u = v = w = x$, nous obtenons que $h_x(x) = 1$. D'où l'alternative annoncée.

Soit x un état récurrent. L'application de la proposition 4.2.1 nous donne :
pour $u = v = x$ et $w = y$, $f_y(x) = h_y(x) = h_y(x) = 1$;
puis pour $u = x$ et $v = w = y$, $h_y(y) = 1$;
et enfin pour $u = w = x$ et $v = y$, $f_x(y) = h_x(y) = h_x(x) = 1$.

Si E est fini, l'égalité

$$\Omega = \bigcup_{y\in E} \{N_y = +\infty\}$$

montre qu'il existe $y \in E$ tel que $h_y(x) > 0$; ce qui implique que y est récurrent. □

Revenons à la partition de E en classes d'équivalence pour la relation " $\leftrightarrow$ ".

1) Si une classe C possède un état récurrent, alors tous les états de C sont récurrents. Partant d'un état de C, la chaîne reste dans C et visite tous les états de C une infinité de fois :

$$\forall x, y \in C,\ \mathbb{P}_x[\{N_y = +\infty\}] = 1.$$

Nous dirons que la classe C est **récurrente**.

2) Si une classe C ne possède aucun état récurrent, alors nous dirons que la classe C est **transiente**.

• **Critères de récurrence**

Proposition 4.2.4 *Soit x un élément de E. Alors :*

$$\begin{array}{rcl} x \textit{ est récurrent} & \Leftrightarrow & \pi_x = \mathbb{P}_x[\{S_x < +\infty\}] = 1 \ ; \\ x \textit{ est transitoire} & \Leftrightarrow & \pi_x = \mathbb{P}_x[\{S_x < +\infty\}] < 1. \end{array}$$

Lorsque x est un état transitoire, la v.a.r. N_x suit une loi géométrique de paramètre π_x.

Preuve : De l'égalité

$$N_x = 1 + (N_x \circ \theta^{S_x})\ 1_{\{S_x < +\infty\}}, \mathbb{P}_x\text{-p.s.},$$

et de la propriété de Markov forte, il résulte que

$$\begin{array}{rcl} \mathbb{P}_x[\{N_x = k+1\}] & = \mathbb{P}_x[\{N_x \circ \theta^{S_x}\ 1_{\{S_x<+\infty\}} = k\} & \\ & = & \mathbb{E}_x[\mathbb{E}_x[1_{\{N_x = k\}} \circ \theta^{S_x}\ 1_{\{S_x<+\infty\}} | \mathcal{F}_{S_x}]] \\ & = & \mathbb{E}_x[1_{\{S_x<+\infty\}} \mathbb{P}_{X_{S_x}}[\{N_x = k\}]] \\ & = & \mathbb{P}_x[\{N_x = k\}\pi_x; \end{array}$$

d'où :

$$\begin{array}{rcl} \mathbb{P}_x[\{N_x = k+1\}] & = & (\pi_x)^k\ \mathbb{P}_x[\{N_x = 1\}] \\ & = & (\pi_x)^k\ \mathbb{P}_x[\{S_x = +\infty\}] \\ & = & (\pi_x)^k\ (1 - \pi_x). \end{array}$$

Si $\pi_x = 1$ alors $\forall k \geq 1,\ \mathbb{P}_x[\{N_x = k\}] = 0$ et par suite $\mathbb{P}_x[\{N_x = +\infty\}] = 1$. L'état x est donc récurrent.

Si $\pi_x < 1$, $\mathbb{P}_x[\{N_x = +\infty\}] = 1 - \sum_{k\geq 1} \mathbb{P}_x[\{N_x = k\}] = 0.$ $\square$

Le résultat précédent permet de donner une nouvelle caractérisation de la récurrence à l'aide du noyau potentiel G.

Proposition 4.2.5 *Soit x un élément de E. Alors :*

$$\begin{array}{rcl} x \textit{ est récurrent} & \Leftrightarrow & G(x,x) = +\infty \ ; \\ x \textit{ est transitoire} & \Leftrightarrow & G(x,x) < +\infty. \end{array}$$

Lorsque x est récurrent, alors $G(x,y) = +\infty$ ou 0 suivant que x mène ou non à y.

Lorsque y est transitoire, alors pour tout $x \in E$, $G(x,y) \leq G(y,y) < +\infty$.

Preuve : Toutes les assertions, exceptée la dernière, découlent de la proposition précédente et du théorème 4.2.3.

De l'égalité

$$N_y = (N_y \circ \theta^{T_y})\ 1_{\{T_y<+\infty\}}$$

et de la propriété de Markov forte, il résulte que

$$\begin{aligned} G(x,y) = \mathbb{E}_x[N_y] &= \mathbb{E}_x[1_{\{T_y<+\infty\}}\ \mathbb{E}_{X_{T_y}}[N_y]] \\ &= \mathbb{P}_x[\{T_y < +\infty\}]\ G(y,y); \end{aligned}$$

□

Remarque 4.2.6

Soit C une classe d'équivalence de la chaîne. Supposons que C soit récurrente. La chaîne, partant de C, ne quitte pas C et visite tous les états de C une infinité de fois. Cette chaîne induit alors sur C une chaîne récurrente irréductible (i.e. que la chaîne de Markov canonique sur C de probabilité de transition P est récurrente irréductible). Supposons en outre que C soit de période $d > 1$ et appelons $C_0, \ldots, C_{d-1}$ les sous-classes cycliques de C. Nous avons

$$\forall x \in E,\ G(x,x) = \sum_{n\geq 0} P^{nd}(x,x) = G^{(d)}(x,x)\ ;$$

ce qui montre que tous les états de C sont récurrents pour la chaîne de probabilité de transition P^d. Partant d'un état appartenant à une sous-classe C_ℓ, la chaîne de probabilité de transition P^d ne quitte pas C_ℓ et elle induit sur C_ℓ une chaîne récurrente irréductible apériodique.

Lorsque C est transiente, tous les états de C sont transitoires pour la chaîne de probabilité de transition P^d. La chaîne partant d'un état de C peut en général quitter C (c'est le cas par exemple si C est fini). Supposons que la chaîne ne puisse pas quitter C (i. e. $\forall x \in C,\ P(x,C) = 1$) et que C soit de période $d > 1$. Alors la chaîne de probabilité de transition P^d, partant d'un état d'une sous-classe cyclique C_ℓ de C, ne quitte pas C_ℓ et elle induit sur C_ℓ une chaîne transiente irréductible apériodique.

4.3 Comportement asymptotique d'une chaîne récurrente irréductible

Hypothèse : Dans cette section, nous supposons que **la chaîne est récurrente irréductible**.

• Fonctions P-surharmoniques positives

Lemme 4.3.1 *Les fonctions P-surharmoniques positives sont constantes.*

Preuve : Soient h une fonction P-surharmonique positive et x, y deux éléments de E. Puisque la chaîne est récurrente irréductible, nous savons que :

$$\mathbb{P}_x[\{N_x = +\infty\}] = \mathbb{P}_x[\{N_y = +\infty\}] = 1.$$

Pour $\mathbb{P}_x$-presque tout ω, la martingale $(h(X_n(\omega)))_{n\geq 0}$, converge et prend une infinité de fois les valeurs $h(x)$ et $h(y)$. □

Corollaire 4.3.2 *Les fonctions P-harmoniques bornées sont constantes.*

Preuve : Exercice.

• **Mesures positives P-invariantes**.

Théorème 4.3.3 *Il existe une mesure positive P-invariante μ sur E vérifiant*

$$\forall x \in E,\ 0 < \mu(x) < +\infty.$$

Toute mesure positive P-sous-invariante ν (i.e. telle que $\nu P \leq \nu$) s'écrit $c\ \mu$, pour $c \in [0, +\infty]$.

La mesure μ est bornée si, et seulement si,

$$\forall x \in E,\ \mathbb{E}_x[S_x] < +\infty.$$

Dans ce cas, il existe une unique probabilité P-invariante π donnée par

$$\pi(x) = \frac{1}{\mathbb{E}_x[S_x]}.$$

Le théorème 4.3.3 résulte des deux lemmes suivants.

Lemme 4.3.4 *Soit μ une mesure positive P-invariante sur E telle que, pour un état $x_0 \in E$, on ait $0 < \mu(x_0) < +\infty$ (μ est non triviale). Alors μ vérifie $\forall x \in E,\ 0 < \mu(x) < +\infty$ et toute mesure positive P-sous-invariante ν (i.e. telle que $\nu P \leq \nu$) s'écrit $c\ \mu$, pour $c \in [0, +\infty]$.*

Preuve : Soient x et y deux états de la chaîne. Comme la chaîne est irréductible, il existe $n \geq 1$ tel que $P^n(y, x) > 0$ et par suite nous avons :

$$\mu(x) = \mu P^n(x) = \sum_{z \in E} \mu(z)\, P^n(z, x) \geq \mu(y)\ P^n(y, x).$$

En prenant $x = x_0$, on obtient : $\mu(y) < +\infty$, $\forall y \in E$, et en prenant $y = x_0$: $\mu(x) > 0$, $\forall x \in E$. D'où la première assertion du lemme.

Considérons la probabilité de transition $\hat{P}$ définie par

$$\forall x, y \in E,\ \hat{P}(x, y) = \frac{\mu(y)}{\mu(x)}\ P(y, x).$$

On notera que, pour tout $p \in [1, +\infty]$, $\hat{P}$ est l'opérateur dual de P dans $L^p(E, \mu)$. Nous avons

$$\forall n \geq 0,\ \hat{P}^n(x, y) = \frac{\mu(y)}{\mu(x)}\ P^n(y, x) \text{ et } \hat{G}(x, y) = \frac{\mu(y)}{\mu(x)}\ G(y, x).$$

Ce qui montre (cf. proposition 4.2.5) que la chaîne de probabilité de transition $\hat{P}$ est récurrente irréductible.

Soit ν une mesure positive sur E, vérifiant $\nu P \leq \nu$. On vérifie facilement que la dérivée de Radon-Nikodym

$$f(x) = \frac{\nu(x)}{\mu(x)}, \quad (x \in E)$$

de ν par rapport à μ est $\hat{P}$-surharmonique positive. D'après le lemme 4.3.1, cette fonction est constante et il existe donc $c \in [0, +\infty]$ tel que $\forall x \in E,\ \nu(x) = c\ \mu(x)$. □

Lemme 4.3.5 *Pour tout $x \in E$, la mesure positive λ_x sur E définie par*

$$\forall y \in E,\ \lambda_x(y) = \mathbb{E}_x\Big[\sum_{0 \le k < S_x} 1_{\{X_k = y\}}\ \Big]$$

est P-invariante non triviale.

<u>Preuve</u> : Nous avons :

$$\begin{aligned}
\lambda_x P(y) &= \sum_{z \in E} \lambda_x(z)\ P(z,y) = \mathbb{E}_x[\sum_{0 \le k < S_x} P(X_k, y)] \\
&= \mathbb{E}_x[\sum_{0 \le k < S_x} \mathbb{P}_{X_k}[X_1 = y]] \\
&= \mathbb{E}_x[\sum_{0 \le k < S_x} \mathbb{P}_x[X_{k+1} = y | \mathcal{F}_k]] \quad \text{(propriété de Markov)} \\
&= \mathbb{E}_x[\sum_{k \ge 0} \mathbb{P}_x[\{X_{k+1} = y\} \cap \{S_x > k\} | \mathcal{F}_k]] \\
&= \sum_{k \ge 0} \mathbb{P}_x[\{X_{k+1} = y\} \cap \{S_x > k\}]] \\
&= \mathbb{E}_x[\sum_{1 \le k \le S_x} 1_{\{X_k = y\}}].
\end{aligned}$$

Or nous avons $1_{\{X_0 = y\}} = 1_{\{X_{S_x} = y\}} = \delta_x^y$, $\mathbb{P}_x$-p.s. D'où la P-invariance de λ_x.
Enfin, l'égalité $\lambda_x(x) = 1$ montre que λ_x est non triviale. □

• **Convergence en moyenne des itérés de P**

Proposition 4.3.6 *Nous avons l'alternative suivante :*
<u>ou bien</u> il existe une unique probabilité P-invariante π, et dans ce cas, pour tout $x \in E$ et toute fonction bornée f sur E,

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f(x) \ = \ \int_E f(u)\ \pi(du)\ ;$$

<u>ou bien</u> il n'existe pas de probabilité P-invariante, et dans ce cas, pour tout $x \in E$ et toute fonction f de $C_0(E)$,

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k f(x) \ = \ 0.$$

Avant de démontrer le résultat, faisons la remarque suivante qui nous sera utile par la suite :

Remarque 4.3.7

Toute suite bornée de fonctions sur l'ensemble dénombrable E est (procédé diagonal) relativement compacte pour la convergence simple (ou la convergence uniforme sur les parties finies de E). Pour montrer la convergence simple d'une telle suite, il suffit donc de prouver qu'elle possède une unique valeur d'adhérence.

Preuve de la proposition 4.3.6 : Pour toute fonction f_0 bornée sur E, la suite de fonctions

$$M_n f = \frac{1}{n}\sum_{k=0}^{n-1} \ P^k f_0)_{n\geq 0}$$

est bornée par $\|f_0\|_\infty$. Soit h_0 une valeur d'adhérence de cette suite et $(\varphi(n))_{n\geq 0}$ une suite strictement croissante d'entiers le long de laquelle la suite de fonctions $(M_n f_0)_{n\geq 0}$ converge simplement vers h_0.

L'espace $(C_0(E), \| \ \|_\infty)$ étant séparable et P étant une contraction de cet espace, on peut trouver, par le procédé diagonal, une suite strictement croissante d'entiers $(\psi(n))_{n\geq 0}$ telle que, pour toute $f \in C_0(E)$, la suite de fonctions $(M_{\varphi(\psi(n))} f)_{n\geq 0}$ converge simplement sur E. On en déduit que, pour tout $x \in E$, la suite de probabilités $(\frac{1}{\varphi(\psi(n))}\sum_{k=0}^{\varphi(\psi(n))} \ P^k(x,\cdot))$ converge vaguement vers une mesure positive ρ_x, nécessairement P-invariante et telle que, pour toute fonction f de $C_0(E)$, la fonction $x \to \rho_x(f)$ est P-harmonique bornée, donc constante d'après le corollaire 4.3.2. La mesure ρ_x ne dépend donc pas de $x \in E$; nous la notons ρ. Nous avons $\rho(E) \leq 1$.

S'il n'existe pas de probabilité P-invariante, alors ρ est nécessairement nulle, car, dans le cas contraire, $\frac{\rho}{\rho(E)}$ serait une probabilité P-invariante. On est alors dans le second cas de l'alternative.

Supposons qu'il existe au moins une probabilité P-invariante π. Pour toute f de $C_0(E)$, nous avons, d'après le théorème de convergence dominée,

$$\lim_{n\to+\infty} \frac{1}{\varphi(\psi(n))}\int_E \sum_{k=0}^{\varphi(\psi(n))} \ P^k f(x) \ \pi(dx) = \pi(f) = \rho(f).$$

Ceci montre l'unicité de π et le fait que l'on est alors dans le premier cas de l'alternative pour les fonctions f de $C_0(E)$. Le passage de $C_0(E)$ aux fonctions bornées est classique et laissé en exercice.

□

- **Fonctions harmoniques bornées pour la chaîne espace-temps**

Nous appelons $\mathcal{K}$ l'espace vectoriel des fonctions harmoniques bornées pour la chaîne espace-temps. Un élément $H \in \mathcal{K}$ est une fonction bornée sur $E \times \mathbb{N}$ telle que

$$\forall x \in E, \ \forall n \in \mathbb{N}, \ H(x,n) = \sum_{y\in E} \ P(x,y) \ H(y,n+1).$$

L'élément H s'identifie donc à une suite $(h_n)_{n\geq 0}$ de fonctions sur E telle que $\forall n \geq 0, \ Ph_{n+1} = h_n$.

Théorème 4.3.8 *Appelons d la période de la chaîne (cf. (4.1.8)). Alors, toute fonction H de $\mathcal{K}$ vérifie :*

$$\forall x \in E, \ \forall k \in \mathbb{N}, \ H(x,k) = H(x,k+d).$$

En particulier lorsque la chaîne est apériodique, $\mathcal{K}$ est réduit aux fonctions constantes.

Preuve : Pour tout $z \in E$ et tout entier naturel p, le processus $(H(X_n, n+p))_{n\geq 0}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P}_z)$ est une martingale bornée relativement à la filtration naturelle $(\mathcal{F}_n^X)_{n\geq 0}$ associée au processus $(X_n)_{n\geq 0}$. Ce processus converge donc $\mathbb{P}_z$-p.s. et nous avons (cf. ??(I.5)) :

$$\forall n \geq 0, \ H(z,p) = \mathbb{E}_z[H(X_n, n+p)] = \mathbb{E}_z[\lim_{k\to+\infty} \ H(X_k, k+p)]$$

$$\forall r \geq 0 \lim_{n\to+\infty} \sum_{y\in E} \left(H(y, r+n+p) - H(X_n, p+n)\right)^2 P^r(X_n, y) = 0, \mathbb{P}_z\text{-p.s..}$$

Appelons Ω_0 l'ensemble des $\omega \in \Omega$ pour lesquels, pour tous entiers $p, r \geq 0$, la limite $\lim_{n\to+\infty} H(X_n(\omega), n+p)$ existe et

$$\lim_{n\to+\infty} \sum_{y\in E} \left(H(y, r+n+p) - H(X_n(\omega), p+n)\right)^2 P^r(X_n(\omega), y) = 0.$$

Soit $x \in E$ et $A = \Omega_0 \cap \{N_x = +\infty\}$. Nous avons $\mathbb{P}_z[A] = 1$, $\forall z \in E$. Soit $\omega \in A$ et $(\varphi_\omega(n))_{n\geq 0}$ une suite d'entiers telle que $\forall n \geq 0,\ X_{\varphi_\omega(n)}(\omega) = x$.

Nous avons, pour tout $p \geq 0$,

$$\lim_{n\to+\infty} \sum_{y\in E} \left(H(y, r+\varphi_\omega(n)+p) - H(x, p+\varphi_\omega(n))\right)^2 P^r(x, y) = 0$$

Il existe (remarque 4.3.7) une suite strictement croissante d'entiers $(\psi(n))_{n\geq 0}$ telle que, pour tout $p \geq 0$, la suite de fonctions $(H(\cdot, \varphi_\omega(\psi(n)) + p))_{n\geq 0}$ converge simplement sur E vers une fonction L de $\mathcal{K}$ qui vérifie, pour tout $p \geq 0$

$$\sum_{y\in E} \left(L(y, r+p) - L(x, p)\right)^2 P^r(x, y) = 0$$

c'est-à-dire $L(y, r+p) = L(x, p)$, pour tout entier r tel que $P^r(x, y) > 0$.

Il en résulte que, pour tout $p \geq 0$, $L(x, d+p) = L(x, p)$.

Des égalités, vérifiées pour tout $\omega \in A$ et pour tout $p \geq 0$,

$$\lim_{n\to+\infty} H(X_n(\omega), n+p) = L(x, p) \text{ et } H(x, p) = \mathbb{E}_x[\lim_{n\to+\infty} H(X_n, n+p)]$$

il résulte alors que $\forall x \in E,\ \forall p \geq 0,\ H(x, d+p) = H(x, p)$.

Lorsque $d = 1$, la fonction H s'identifie à une fonction P-harmonique bornée, qui est constante d'après le corollaire 4.3.2. □

Corollaire 4.3.9 *Supposons que la chaîne soit apériodique. Alors nous avons l'alternative suivante :*
ou bien il existe une unique probabilité P-invariante π et dans ce cas, pour tout $x \in E$ et toute fonction bornée f sur E,

$$\lim_{n\to+\infty} P^n f(x) = \int_E f(u)\ \pi(du)\ ;$$

ou bien il n'existe pas de probabilité P-invariante et dans ce cas, pour tout $x \in E$ et toute fonction f de $C_0(E)$,

$$\lim_{n\to+\infty} P^n f(x) = 0.$$

Preuve : Soient $f_0 \in \mathcal{C}_0(E)$ et h_0 une valeur d'adhérence simple de la suite de fonction $(P^n f)_{n\geq 0}$ (cf. remarque (4.3.7)). Par le procédé diagonal, il existe une suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$ telle que pour tout $k \geq 0$, tout $x \in E$ et toute fonction f de $C_0(E)$, la suite $(P^{\varphi(n)-k} f(x))_{n\geq 0}$ converge vers une limite notée $H_f(x, k)$ vérifiant $H_{f_0}(\cdot, 0) = h_0$. Pour toute f de $C_0(E)$, il est clair que la fonction H_f appartient à $\mathcal{K}$ et est donc constante d'après le théorème

4.3.8. Nous avons $H_f = \rho(f)$, où ρ est une sous-probabilité limite vague de toutes les suites de probabilités

$$(P^{\varphi(n)-k}(x,\cdot))_{n\geq 0}, x \in E, k \geq 0.$$

S'il n'existe pas de probabilité P-invariante sur E, alors ρ est nécessairement la mesure nulle. La fonction h_0 est la fonction nulle ; d'où la seconde alternative.

S'il existe une unique mesure de probabilité P-invariante π, alors nous avons nécessairement $\rho = \pi$. La fonction h_0 est constante et égale à $\pi(f_0)$; d'où la première alternative pour les éléments de $\mathcal{C}_0(E)$.

Le passage de $\mathcal{C}_0(E)$ aux fonctions bornées est laissé en exercice. □

Corollaire 4.3.10 *Supposons que la chaîne admette une période $d > 1$ et appelons $C_0, \ldots, C_{d-1}$ ses sous-classes cycliques. Nous avons alors l'alternative suivante :*
<u>Ou bien</u> il existe une unique probabilité P-invariante π. Dans ce cas, pour chaque $0 \leq k \leq d-1$, $\pi(C_k) = \frac{1}{d}$ et la restriction de la chaîne de probabilité de transition P^d à C_k possède l'unique probabilité P^d-invariante $\pi_k = \frac{1}{d}\, 1_{C_k}\pi$. Pour tout $x \in C_k$, tout $k \geq 0$, tout $j \geq 0$ et toute fonction bornée f sur E, nous avons alors :

$$\lim_{n\to+\infty} P^{nd+j}f(x) \;=\; \int_E f(u)\;\pi_{k+j}(du)\;;$$

<u>Ou bien</u> il n'existe pas de probabilité P-invariante et dans ce cas, pour tout $x \in E$ et toute fonction f de $\mathcal{C}_0(E)$, nous avons :

$$\lim_{n\to+\infty} P^n f(x) \;=\; 0.$$

<u>Preuve</u> : Pour toute classe cyclique C_k, $0 \leq k < d$, la chaîne de probabilité de transition P^d est récurrente apériodique (cf. (4.2.6)). De plus, pour tous $n, k, j \geq 0$ et tout $x \in C_k$, la probabilité $P^{nd+j}(x,.)$ est portée par C_{k+j} (cf. lemme 4.1.9).

S'il n'existe pas de probabilité P-invariante, alors, d'après l'exercice 4.1.11, il n'existe pas de probabilité P^d-invariante. Dans ce cas (corollaire 4.3.9), pour tout $f \in \mathcal{C}_0(E)$ et tout $x \in C_k$, nous avons :

$$P^{nd+j}f(x) = P^j[P^{nd}f_{|C_k}](x) \;\; n \vec{\to} +\infty\; 0.$$

S'il existe une probabilité P-invariante π, alors, d'après l'exercice 4.1.11, les $\pi_k = \dfrac{1}{d}1_{C_k}\pi$ sont des probabilités P^d-invariantes et nous avons : $\pi_k P^j = \pi_{k+j}, \forall j \geq 0$.

D'après le corollaire 4.3.9 et l'exercice 4.1.11, pour toute fonction f bornée sur E et tout $x \in C_k$, nous avons :

$$P^{nd+j}f(x) = P^j[P^{nd}f_{|C_k}](x) \;\;\overset{n\to+\infty}{\to}\; (\pi_k P^j)(f) = \pi_{k+j}(f).$$

□

Définition 4.3.11 Une chaîne récurrente irréductible est dite **récurrente positive**, si elle possède une unique probabilité P-invariante. Dans le cas contraire, elle est dite **récurrente nulle**.

4.4 Exercices et problèmes

Nous donnons ici à titre d'exercice quelques exemples de modèles markoviens.

Exercice 4.4.1 *Fluctuation de deux liquides incompressibles entre deux récipients (Daniel Bernoulli, 1769).*

Deux urnes contiennent $2N$ boules, N blanches et N noires, qui représentent les molécules des deux liquides. L'incompressibilité se traduit par le fait que chacune des urnes contient N boules à tout instant. La fluctuation des molécules entre les deux urnes est modélisée de la façon suivante : à chaque instant $n = 0, 1, \ldots$, on choisit une boule dans chacune des urnes et on les échange. On note X_n le nombre de boules blanches dans la première urne à l'instant n.

1) Pour tout entier $n \geq 0$, calculer $\mathbb{P}[X_{n+1} = x_{n+1} | X_0 = x_0, \ldots X_n = x_n]$ pour des entiers $x_0, \ldots, x_n$ de $\{0, \ldots, N\}$ tels que $\mathbb{P}[\{X_0 = x_0, \ldots X_n = x_n\}] > 0$.

En déduire que le processus $(X_n)_{n\geq 0}$ est une chaîne de Markov de probabilité de transition P définie par :

$$\forall 1 \leq i \leq N,\ P(i-1, i) = (1 - \frac{i}{N})^2 \ \text{ et } \ P(i, i-1) = (\frac{i}{N})^2 \ ;$$

$$\forall 0 \leq i \leq N,\ P(i, i) = 2(\frac{i}{N})(1 - \frac{i}{N}).$$

2) Montrer que la chaîne est récurrente positive irréductible et trouver l'unique probabilité P-invariante.

Exercice 4.4.2 *Modèle d'échange de chaleur entre deux corps (Tatiana et Paul Ehrenfest, 1907).*

Deux urnes contiennent un total de $2N$ boules numérotées de 1 à $2N$. Les températures sont représentées par le nombre de boules. L'échange de chaleur est modélisé comme suit : à chaque instant $n = 0, 1, \ldots$, on choisit au hasard un entier de $\{1, \ldots, 2N\}$ et la boule dont le numéro est tiré est changée d'urne. On appelle $(X_n)_{n\geq 0}$ le nombre de boules se trouvant dans la première urne à l'instant n.

1) Montrer que $(X_n)_{n\geq 0}$ est une chaîne de Markov. Préciser l'espace des états et la probabilité de transition.

2) Montrer que cette chaîne est récurrente positive irréductible et déterminer son unique probabilité invariante.

3) Pour tout état initial k et tout entier $n \geq 0$, calculer $\mathbb{E}_k[X_n]$.

Etudier le comportement asymptotique de la suite $(\mathbb{E}_k[X_n])_{n\geq 0}$ et interpréter le résultat.

Exercice 4.4.3

I Sur l'ensemble $E = \{0, 1, \ldots, N\}$, $N \geq 2$, on considère la probabilité de transition P définie par :

$$P(0,0) = P(N,N) = 1 \ \text{ et } \ \forall 1 \leq i \leq N-1, \forall j \in E, P(i,j) = C_N^j(\frac{i}{N})(1 - \frac{i}{N})^{N-j}.$$

On désigne par $(\Omega, \mathcal{F}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ la chaîne de Markov canonique de probabilité de transition P.

1) a) Montrer que la fonction identité I de E est P-harmonique.

b) En déduire que le processus $(X_n)_{n\geq 0}$ converge, $\forall x \in E$, $\mathbb{P}_x$-p.s., vers une v.a.r. Z à valeurs dans E.

2) a) Pour tout $p \geq 0$ et tout $k \in \{1, \ldots, N-1\}$, montrer que

$$\forall x \in E,\ \mathbb{P}_x[\bigcap_{n\geq p} \{X_n = k\}] = 0.$$

b) En déduire que, $\forall x \in E$, $\mathbb{P}_x$-p.s., la v.a.r. Z est à valeurs dans l'ensemble $\{0, N\}$.

3) a) Etablir les relations : $\forall x \in E,\ \forall n \geq 0,\ x = \mathbb{E}_x[X_n] = \mathbb{E}_x[Z]$.

b) En déduire les valeurs des probabilités $\mathbb{P}_x[\{Z = 0\}]$ et $\mathbb{P}[\{Z = N\}]$, pour tout $x \in E$.

4) Déterminer toutes les fonctions P-harmoniques sur E.

II. Une urne contient X_0 boules blanches et $(N - X_0)$ boules noires. A chaque instant $n = 0, 1, \ldots$, on fait l'opération suivante : on effectue N tirages indépendants et avec remise. Si X_n est le nombre de boules blanches tirées, on vide l'urne et on y place X_n boules blanches et $(N - X_n)$ boules noires.

1) Montrer que le processus $(X_n)_{n\geq 0}$ est une chaîne de Markov homogène. Donner l'espace des états et la probabilité de transition.

2) Décrire le comportement de la chaîne.

Exercice 4.4.4

Soit μ une probabilité sur l'ensemble $\mathbb{N}$ des entiers naturels. On considère la probabilité de transition sur $\mathbb{N}$ définie par : pour tout $j \in \mathbb{N},\ P(0, j) = \mu(\{j\})$ et pour tout $k \in \mathbb{N}^*$, $P(k, \cdot)$ est la loi uniforme sur $\{0, 1, \ldots, k-1\}$ On désigne par $(\Omega, \mathcal{F}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ la chaîne de Markov canonique de probabilité de transition P.

1) On suppose que le support de μ est fini. Classifier les états. Donner les états récurrents. Décrire le comportement de la chaîne.

2) On suppose que le support de μ est infini. Montrer que la chaîne est récurrente irréductible.

3) a) Montrer que la mesure positive ν est P-invariante, si et seulement si,

$$\forall k \geq 1,\ \nu(k) = \nu(0)[1 - \mu(0) + \mu(k)] - \sum_{j=1}^{k} \frac{\nu(j)}{j}$$

$$\text{et } \nu(0)(1 - \mu(0)) = \sum_{j\geq 1} \frac{\nu(j)}{j}.$$

b) Montrer que l'on a alors :

$$\forall k \geq 1,\ \nu(k) = \frac{\nu(0)}{k+1}(k\mu(k) + \sum_{j\geq k} \mu(j)).$$

c) A quelle condition a-t-on récurrence positive ?

d) Donner un exemple de récurrence positive et un exemple de récurrence nulle.

Exercice 4.4.5

Soient $(a_n)_{n\geq 0}$ et $(b_n)_{n\geq 0}$ deux suites de nombres réels appartenant à]0,1[, vérifiant $a_0 + b_0 = 1$. Sur l'ensemble des entiers relatifs $\mathbb{Z}$, on considère la probabilité de transition P définie par :

$$\begin{aligned} \forall n > 0, \quad & P(n, n+1) = a_n \text{ et } P(n,0) = 1 - a_n \ ; \\ \forall n > 0, \quad & P(-n, -n-1) = b_n \text{ et } P(-n,0) = 1 - b_n \ ; \\ & P(0,1) = a_0 \text{ et } P(0,-1) = b_0 = 1 - a_0. \end{aligned}$$

On désigne par $(\Omega, \mathcal{F}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ la chaîne de Markov canonique de probabilité de transition P sur $\mathbb{Z}$.

1) Montrer que la chaîne est irréductible.

2) a) Soit $S_0 = \inf\{n \geq 1 : X_n = 0\}$ le temps de retour en zéro. Déterminer la loi de S_0 sous $\mathbb{P}_0$.

b) Donner des conditions nécessaires et suffisantes de récurrence de la chaîne.

3) On se place dans la situation où la chaîne est récurrente.

a) Déterminer l'unique (à constante multiplicative près) mesure positive P-invariante.

b) Donner des conditions nécessaires et suffisantes de récurrence positive de la chaîne.

c) Donner la valeur de $\mathbb{E}_0[S_0]$.

4) On se place dans la situation où la chaîne est transiente. Soit x un entier relatif.

a) Montrer que

$$\Omega = \{X_n \to +\infty\} \cup \{X_n \to -\infty\},\ \mathbb{P}_x - p.s..$$

b) Montrer que l'on a, $\mathbb{P}_x$-p.s. :

$$\text{si } x \leq 0, \{X_n \to +\infty\} = \bigcup_{r\geq 0} \{X_{r+n} = n, \forall n \geq 0\} \ ;$$

$$\text{si } x > 0, \{X_n \to +\infty\} = \{X_n = x + n, \forall n \geq 0\} \bigcup \Big(\bigcup_{r\geq 0} \{X_{r+n} = n, \forall n \geq 0\} \Big).$$

c) Etablir des égalités analogues pour l'évènement $\{X_n \to -\infty\}$.

d) Déterminer les probabilités des évènements $\{X_n \to -\infty\}$ et $\{X_n \to +\infty\}$ sous $\mathbb{P}_x$ ainsi que la valeur du potentiel $G(x, 0)$.

Exercice 4.4.6

I. Soient N et M deux entiers non nuls. On appelle E le pavé de $\mathbb{N}^2$ défini par :

$$E = \{(x,y) \in \mathbb{N}^2 : 0 \leq x \leq N, 0 \leq y \leq M\}.$$

On pose :

$$\overset{\circ}{E} = \{(x,y) \in E : x < N \text{ et } y < M\} \quad \text{et} \quad \partial E = \{(x,y) \in E : x = N \text{ ou } y = M\}.$$

Soient p et q deux réels positifs vérifiant $p + q = 1$. Nous considérons sur E la probabilité de transition P définie par :

$$\forall\ (x,y) \in E,\ P((x,y),(x+1,y)) = p \text{ et } P((x,y),(x,y+1)) = q$$

$$\forall\ (x,y) \in \partial E,\ \ P((x,y),(x,y)) = 1.$$

Nous désignons par

$$(\Omega = E^{\mathbb{N}}, \mathcal{F} = \bigotimes_{\mathbb{N}} \mathcal{P}(E), (Z_n = (X_n, Y_n))_{n\geq 0}, (\mathbb{P}_z)_{z\in E})$$

la chaîne de Markov canonique sur $E^{\mathbb{N}}$ de probabilité de transition P.

1) Pour tout $t \in \mathbb{R}$, on appelle h_t la fonction sur E définie par

$$h_t(x,y) = t^x (\frac{1-pt}{q})^y,\ \ (x,y) \in E.$$

Vérifier que h_t est une fonction P-harmonique bornée sur E. Pour tout $t \in \mathbb{R}$ et tout $z \in E$, que peut-on dire du processus $(h_t(Z_n))_{n\geq 0}$ défini sur $(\Omega, \mathcal{F}, \mathbb{P}_z)$?

2) Montrer que le processus $(Z_n)_{n\geq 0}$ à valeurs dans E converge, $\mathbb{P}_z$-p.s. $\forall z \in E$, vers une v.a.r. $Z_\infty = (X_\infty, Y_\infty)$ à valeurs dans $\partial E \setminus \{(N,M)\}$ vérifiant

$$\forall\ t \in \mathbb{R},\ \forall\ z \in E,\ \ \mathbb{E}_z[h_t(Z_\infty)] = h_t(z).$$

3) En déduire que l'on a, pour tout réel t,

$$\begin{aligned} t^N \quad & \sum_{k=0}^{M-1} (\frac{1-pt}{q})^k\ \mathbb{P}_0[\{Z_\infty = (N,k)\}] \\ & +(\frac{1-pt}{q})^M \sum_{k=0}^{N-1} t^k\ \mathbb{P}_0[\{Z_\infty = (k,M)\}] \ = \ 1. \end{aligned}$$

4) a) Montrer que le développement en éléments simples de la fraction rationnelle $\dfrac{1}{x^N(1-x)^M}$ est donné par :

$$\frac{1}{x^N(1-x)^M} = \sum_{k=0}^{N-1} C_{M-1+k}^{M-1} \frac{1}{x^{N-k}} + \sum_{k=0}^{M-1} C_{N-1+k}^{N-1} \frac{1}{(1-x)^{M-k}}.$$

b) En déduire la loi de la v.a. Z_∞.

II. Une épreuve entre deux joueurs A et B conduit au résultat suivant : A est gagnant avec une probabilité p et B est gagnant avec une probabilité $q = 1-p$.

On se donne deux entiers naturels N et M. Le jeu consiste en une répétition d'épreuves indépendantes. Le jeu s'arrête lorsque : ou bien le joueur A a obtenu N victoires (A est alors le gagnant du jeu) ; ou bien le joueur B a obtenu M victoires (B est alors le gagnant du jeu).

a) Donner les probabilités de gagner pour chacun des joueurs.

b) Si $p_A(N,M)$ désigne la probabilité de gagner du joueur A, vérifier que :

$$\lim_{M\to+\infty} p_A(N,M) = 1 \quad \text{et} \quad \lim_{N\to+\infty} p_A(N,M) = 0.$$

c) Déterminer le nombre moyen d'épreuves nécessaires pour que le jeu se termine.

Exercice 4.4.7

Sur l'ensemble des entiers $E = \{0, 1, \dots, N\}$, on considère la chaîne de Markov canonique $(\Omega = E^{\mathbb{N}}, \mathcal{F}, (X_n)_{n\geq 0}, (\mathbb{P}_x)_{x\in E})$ de probabilité de transition définie par :

$$P(0,0) = P(N,N) = 1$$

$$\forall 1 \leq k \leq N, P(k, k+1) = p_k \in]0,1[\quad \text{et} \quad P(k, k-1) = q_k = 1 - p_k.$$

On définit sur E les fonctions r et R suivantes :

$$r(0) = 1; R(0) = 0; \forall n \geq 1, r(n) = \prod_{k=1}^{n} \frac{q_k}{p_k} \text{ et } R(n) = \sum_{k=1}^{n} r(k-1).$$

1) Montrer qu'une fonction f sur E est P-harmonique si et seulement si

$$\forall 1 \leq k \leq N-1, \ f(k+1) - f(k) = \frac{q_k}{p_k} \left(f(k) - f(k-1)\right).$$

Déterminer alors toutes les fonctions P-harmoniques. Montrer qu'elles constituent un espace vectoriel de dimension 2, dont on donnera une base. [On se servira des fonctions r et R introduites précédemment].

2) Montrer que, pour toute fonction P-harmonique h, la suite de v.a.r. $(h(X_n))_{n\geq 0}$ converge $\mathbb{P}_x$-p.s., pour tout $x \in E$. En déduire que la suite de v.a.r. $(X_n)_{n\geq 0}$ elle-même converge $\mathbb{P}_x$-p.s., $\forall x \in E$, vers une v.a.r. Z à valeurs dans $\{0, N\}$.

3) Montrer que, pour toute fonction P-harmonique h, on a :

$$\forall x \in E, \ \forall n \geq 0, \ h(x) = \mathbb{E}_x[h(Z)].$$

En déduire, pour $x \in E$, les probabilités $\mathbb{P}_x[\{Z = 0\}]$ et $\mathbb{P}_x[\{Z = N\}]$.

4) On pose $T = \inf\{n \geq 0 : X_n \in \{0, N\}$ (autrement dit T est le temps d'entrée dans le sous-ensemble $\{0, N\}$ de E). On désigne par f la fonction sur E définie par : $f(x) = \mathbb{E}_x[T]$.

a) En utilisant la propriété de Markov, montrer que l'on a :

$$\forall x \in \{1, \dots, N-1\}, \ f(x) = Pf(x) + 1.$$

b) Déterminer alors la fonction f.

Chapitre 5

Fonctions μ-harmoniques sur un groupe abélien

5.1 Notations et définitions

Soient $(S, +)$ un semi-groupe abélien localement compact à base dénombrable et σ une mesure de Radon positive sur S. On suppose que S possède un élément neutre, noté 0.

Nous munissons l'espace vectoriel $C(S)$ des fonctions continues sur S, à valeurs réelles, de la topologie de la convergence uniforme sur les compacts. Si $(K_n)_{n\geq 0}$ est une suite croissante de compacts telle que $\bigcup_{n\geq 0} K_n = S$, alors cette topologie est celle définie par la distance :

$$d(f,g) = \sum_{n\geq 0} \frac{1}{2^n} \sup_{s\in K_n} \left(|f(s) - g(s)| \wedge 1\right).$$

Pour tout entier $n \geq 0$, l'application $\varphi_n : f \in C(S) \to \int_{K_n} f(s)\ \sigma(ds)$ est une forme linéaire continue sur $C(S)$.

Définition 5.1.1 Nous appelons **exponentielle** sur $(S, +)$, tout homomorphisme de semi-groupe continu χ de $(S, +)$ dans $(\mathbb{R}_+^*, \times)$ (i.e. tel que $\forall s, t \in S,\ \chi(s + t) = \chi(s)\ \chi(t)$). Muni du produit des fonctions, l'ensemble $\mathcal{E}$ des exponentielles est (pour la topologie de la convergence uniforme sur les compacts) un groupe topologique.

Nous appelons H_+ le cône des fonctions σ-**harmoniques** continues positives ; c'est-à-dire des fonctions h de $C(E)$ positives vérifiant l'équation fonctionnelle

$$\forall x \in S,\ h(x) = \int_S h(x + y)\ \sigma(dy). \tag{5.1}$$

Le cône H_+ contient les **exponentielles harmoniques** (lorsqu'il y en a !) ; c'est-à-dire les exponentielles χ vérifiant $\int_S \chi(s)\ \sigma(ds) = 1$. Si h est une fonction de H_+, alors, pour tout $t \in S$, la translatée $h^t : s \to h(t + s)$ est encore une fonction de H_+.

Nous notons $\mathcal{E}_\sigma$ l'ensemble des exponentielles harmoniques. Nous avons

$$\mathcal{E}_\sigma = \{\chi \in \mathcal{E} : \lim_{n\to+\infty} \uparrow \varphi_n(\chi) = 1\}.$$

Ceci implique que $\mathcal{E}_\sigma$ est un borélien de $\mathcal{E}$. Pour toute mesure positive μ sur les boréliens de $\mathcal{E}$ portée par $\mathcal{E}_\sigma$, la fonction h définie par

$$h(s) = \int_{\mathcal{E}_\sigma} \chi(s)\ \mu(d\chi)$$

est à valeurs dans $[0, +\infty]$, vérifie l'équation fonctionnelle (5.1 et

$$\begin{aligned} h(0) &= \int_S h(s)\ \sigma(ds) = \int_S \int_{\mathcal{E}_\sigma} \chi(s)\ \mu(d\chi)\ \sigma(ds) \\ &= \int_{\mathcal{E}_\sigma} \int_S \chi(s)\ \sigma(ds)\ \mu(d\chi) = \mu(\mathcal{E}_\sigma). \end{aligned}$$

Il s'ensuit que $h(0) < +\infty$ (et par suite que h est finie σ-p.p.) si, et seulement si, $\mu(\mathcal{E}_\sigma) < +\infty$.

5.2 Le théorème de Choquet-Deny

Nous nous proposons de montrer le résultat suivant (cf ? ?) :

Théorème 5.2.1 *Soient $(G, +)$ un* **groupe** *abélien localement compact à base dénombrable et σ une mesure de Radon positive sur G.*

Si l'ensemble $\mathcal{E}_\sigma$ des exponentielles harmoniques sur G est vide, alors le cône H_+ est réduit à la fonction nulle.

Si $\mathcal{E}_\sigma$ n'est pas vide, pour toute fonction non nulle h du cône H_+, il existe une unique mesure positive bornée μ sur les boréliens de $\mathcal{E}_\sigma$ telle que

$$\forall s \in G,\ \ h(s) = \int_{\mathcal{E}_\sigma} \chi(s)\ \mu(d\chi).$$

Lorsque le sous-semi-groupe fermé S_σ engendré par le support de σ est G tout entier, $\mathcal{E}_\sigma$ est compact.

Pour montrer ce résultat, nous commençons par établir un théorème (moins précis) portant sur les semi-groupes abéliens. La démonstration qui suit permet d'étendre le résultat précédent à certains semi-groupes (cf ? ?).

Soit h une fonction du cône H_+ telle que $h(0) > 0$. Nous désignons par Ω l'espace produit $S^{\mathbb{N}^*}$, par $\mathcal{F}$ la tribu des boréliens de Ω et par $(Y_n)_{n\geq 1}$ les applications coordonnées de Ω. Pour tout entier $n \geq 1$, nous notons $\mathcal{F}_n$ la tribu engendrée par les variables aléatoires $\{Y_k : 1 \leq k \leq n\}$ et par $\mathcal{F}_0$ la tribu triviale $\{\emptyset, \Omega\}$. Nous posons : $X_0 = e$ et $\forall n \geq 1,\ X_n = Y_1 + \ldots + Y_n$.

Considérons (théorème de Kolmogorov) l'unique probabilité ${}^h\mathbb{P}$ sur $(\Omega, \mathcal{F})$ vérifiant, pour tout entier $n \geq 1$ et toute fonction borélienne bornée (ou positive) f sur S^n,

$$\int_\Omega f(Y_1, \ldots, Y_n)\ {}^h\mathbb{P}(d\omega) = \int_{S^n} f(x_1, \ldots, x_n)\ \frac{h(x_1 + \ldots + x_n)}{h(0)}\ \sigma(dx_1) \ldots \sigma(dx_n).$$

On voit facilement que :

i) $\forall n \geq 0,\ {}^h\mathbb{P}[\{h(X_n) = 0\}] = 0$; autrement dit,

$${}^h\mathbb{P}\text{-p.s.}, \forall n \geq 0,\ \ h(X_n) > 0.$$

ii) Le processus $(X_n)_{n\geq 0}$, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, {}^h\mathbb{P})$, est, relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$, une chaîne de Markov partant de 0, de probabilité de transition ${}^hP$ définie par

$$ {}^hPf(s) = \frac{1}{h(s)} \int_S f(s+x)\ h(s+x)\ \sigma(dx) $$

pour $s \in S$ vérifiant $h(s) > 0$ et arbitraire pour $s \in S$ vérifiant $h(s) = 0$. Nous choisissons ${}^hPf(s) = f(s)$, pour $s \in S$ tel que $h(s) = 0$.

Une fonction f est ${}^hP$-harmonique si, et seulement si, la fonction $h\ f$ est σ-harmonique. En particulier la fonction $x \to \frac{h(t+x)}{h(x)}$ est ${}^hP$-harmonique.

Pour tout $t \in S$, le processus $(\dfrac{h(t+X_n)}{h(X_n)})_{n\geq 0}$ défini sur $(\Omega, \mathcal{F}, {}^h\mathbb{P})$ est donc une martingale positive relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. D'après le théorème de convergence presque-sûre des martingales, ce processus converge, ${}^h\mathbb{P}$-p.s., vers une v.a.r. positive ${}^h\xi(t,\cdot)$ vérifiant (cf. 2.3.9), pour tout $t \in S$,

$$ h(t+X_n) \geq h(X_n)\ {}^h\mathbb{E}[{}^h\xi(t,\cdot)|\mathcal{F}_n],\ \ {}^h\mathbb{P}\text{-p.s.}\ \ \text{et}\ \ h(t) \geq h(e)\ {}^h\mathbb{E}[{}^h\xi(t,\cdot)]. $$

En fait, nous allons montrer le résultat suivant :

Théorème 5.2.2 *Appelons S_σ le sous-semi-groupe fermé de S engendré par le support de σ. Soit h une fonction de H_+ vérifiant $h(0) > 0$. Avec les notations ci-dessus nous avons, pour tous éléments s et t de S_σ,*
i) ${}^h\xi(s+t,\cdot) = {}^h\xi(s,\cdot)\ {}^h\xi(t,\cdot)$, ${}^h\mathbb{P}$-p.s..
ii) $h(t+X_n) = h(X_n)\ {}^h\mathbb{E}[{}^h\xi(t,\cdot)|\mathcal{F}_n]$, ${}^h\mathbb{P}$-p.s. et $h(t) = h(e)\ {}^h\mathbb{E}[{}^h\xi(t,\cdot)]$,
et, pour tout $t \in S_\sigma$, la martingale $\Big(\dfrac{h(t+X_n)}{h(X_n)}\Big)_{n\geq 0}$ converge aussi en norme $L^p(\Omega, \mathcal{F}, {}^h\mathbb{P})$, $p \geq 1$, vers la v.a.r. ${}^h\xi(t,\cdot)$ et, pour tout $t \in S_\sigma$,

$$ h(t)\ {}^{h^t}\mathbb{P}\ =\ {}^h\xi(t,\cdot)\ h(0)\ {}^h\mathbb{P}, $$

en convenant que la mesure figurant à gauche est la mesure nulle si $h(t) = 0$.

De plus nous avons :

$$ \forall s \in S_\sigma,\ \text{vérifiant } h(s) > 0,\ {}^{h^s}\xi(t,.) = {}^h\xi(t,\cdot),\ {}^{h^s}\mathbb{P}\text{-p.s.} $$

$$ \text{et} \quad \int_S {}^h\xi(t,\cdot)\ \sigma(dt) = 1,\ {}^h\mathbb{P}\text{-p.s.} $$

5.3 Preuve des théorèmes

Pour tout entier $p \geq 1$, nous désignons : par $\mathcal{B}(S^p)$ la tribu des boréliens de l'espace produit S^p ; par σ^p la mesure produit $\otimes_p \sigma$ sur les boréliens de S^p et par σ^{*p} la p-ième convolée de σ (i.e. l'image de σ^p par l'application qui au p-uple $(u_1, \ldots, u_p)$ de S^p associe l'élément $u_1 + \ldots + u_p$ de S).

Pour tout $p \geq 1$ et tout $A \in \mathcal{B}(S^p)$, le processus

$$ (\ H_A(X_n) = \frac{1}{h(X_n)} \int_A h(s_1 + \ldots + s_p + X_n)\ \ \sigma(ds_1)\ldots\sigma(ds_p)\ )_{n\geq 0}, $$

défini sur $(\Omega, \mathcal{F}, {}^h\mathbb{P})$, est une martingale bornée par 1, relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. D'après les résultats du chapitre 2, ce processus converge donc, ${}^h\mathbb{P}$-p.s. et en norme L^p, $p \geq 1$, vers une v.a.r. $\lambda(A,\cdot)$ vérifiant

$$\forall n \geq 0,\ \ H_A(X_n) = {}^h\mathbb{E}[\lambda(A,\cdot)|\mathcal{F}_n] \tag{5.2}$$

$$\text{et} \quad H_A(0) = {}^h\mathbb{E}[\lambda(A,\cdot)]. \tag{5.3}$$

Lemme 5.3.1 *Soient p et q deux entiers supérieurs à 1. Pour tous boréliens $A \in \mathcal{B}(S^p)$ et $B \in \mathcal{B}(S^q)$, nous avons :*

$$\lambda(A \times B, \cdot) = \lambda(A,\cdot)\ \lambda(B,\cdot),\ \ {}^h\mathbb{P}\text{-}p.s..$$

Preuve : Pour tout entier $n \geq 0$, posons

$$U_n = \frac{1}{h(X_n)} \ \int_S [H_A(u+X_n) - H_A(X_n)]^2\ h(u+X_n)\ \sigma^{*q}(du).$$

Nous avons

$$U_n = {}^hP^q([H_A(\cdot) - H_A(X_n)]^2)(X_n) = {}^hP^q H_A^2\ (X_n)\ \ -\ \ H_A^2(X_n)\ ;$$

et par suite

$${}^h\mathbb{E}[U_n] = {}^hP^{q+n}H_A^2(0) - {}^hP^n H_A^2(0).$$

Comme la suite $({}^hP^nH_A^2(0))_{n\geq 0}$ est croissante et majorée par 1, la série $\sum_{n\geq 0} {}^h\mathbb{E}[U_n]$ converge. On en déduit en particulier que la suite de v.a.r. $(U_n)_{n\geq 0}$ converge ${}^h\mathbb{P}$-p.s. vers zéro.

Cela dit, nous avons, en posant $\Delta_n = H_{A\times B}(X_n)\ \ -\ \ H_A(X_n)\ H_B(X_n)$,

$$\begin{aligned}
|\Delta_n| \ &= \ |\int_B \frac{h(u_1+\ldots+u_q+X_n)}{h(X_n)}\ [H_A(u+X_n) - H_A(X_n)]\sigma(du_1)\ldots\sigma(du_q)| \\
&\leq \ \int_B \frac{h(u_1+\ldots+u_q+X_n)}{h(X_n)}\ [H_A(u_1+\ldots+u_q+X_n) - H_A(X_n)]^2\sigma(du_1)\ldots\sigma(du_q) \\
&\leq \ \int_S \frac{h(u_1+\ldots+u_q+X_n)}{h(X_n)}\ [H_A(u+X_n) - H_A(X_n)]^2\sigma^{*q}(du) = U_n.
\end{aligned}$$

□

Du lemme 5.3.1 et des relations (5.3), on déduit immédiatement le corollaire :

Corollaire 5.3.2 *Soient p et q deux entiers supérieurs à 1. Pour tous boréliens $A \in \mathcal{B}(S^p)$ et $B \in \mathcal{B}(S^q)$, vérifiant $\sigma^p(A)\ \sigma^q(B) > 0$, nous avons :*

$${}^h\mathbb{E}\Big[\big(\frac{\lambda(A,\cdot)}{\sigma^p(A)}\ \ -\ \ \frac{\lambda(B,\cdot)}{\sigma^q(B)}\big)^2\Big]\ \ =\ \ \frac{H_{A\times A}(0)}{[\sigma^p(A)]^2} - \frac{2\ H_{A\times B}(0)}{\sigma^p(A)\ \sigma^q(B)}\ \ +\ \ \frac{H_{B\times B}(0)}{[\sigma^p(B)]^2}.$$

Preuve du théorème 5.2.2 : Soient p un entier ≥ 1 et $s_1, \ldots, s_p$ des éléments de *Supp* σ. Pour tout $i \in \{1, \ldots, p\}$, choisissons une suite de voisinages compacts de s_i dans S, $(V_n^i)_{n\geq 0}$, telle que $\bigcap_{n\geq 0} V_n^i = \{s_i\}$. Pour tout entier $n \geq 0$, nous appelons V_n le voisinage compact $V_n^1 \times \ldots \times V_n^p$ de $(s_1, \ldots, s_p)$ dans S^p.

Le corollaire 5.3.2 montre que :

1) La suite de v.a.r. $\Big(\dfrac{\lambda(V_n,\cdot)}{\sigma^p(V_n)}\Big)_{n\geq 0}$ est une suite de Cauchy dans L^2. Ce processus converge donc en norme L^2 vers une v.a.r. $\psi(s_1, \ldots, s_p\ ;\cdot)$.

2) La v.a.r. $\psi(s_1, \ldots, s_p\,; \cdot)$ ne dépend pas du choix des suites de voisinages $(V_n^i)_{n\geq 0}$, $i \in \{1, \ldots, p\}$. Elle ne dépend des éléments $s_1, \ldots, s_p$ que par leur somme $s = s_1 + \ldots + s_p$. Nous notons $\psi(s, \cdot)$ cette v.a.r.

En posant $M = \bigcup_{p\geq 1} Supp\ (\sigma^{*p})$, nous avons alors, pour tous $s, t \in M$,

$$h(0)\ {}^h\mathbb{E}[(\psi(s,\cdot) - \psi(t,\cdot))^2] = h(2s) - 2\ h(s+t) + h(2t), \tag{5.4}$$

conséquence du corollaire 5.3.2,

$$h(t + X_n) = h(X_n)\ {}^h\mathbb{E}[\psi(t,\cdot)|\mathcal{F}_n], {}^h\mathbb{P}\text{-p.s.} \quad \text{et} \quad h(t) = h(e)\ {}^h\mathbb{E}[\psi(t,\cdot)], \tag{5.5}$$

conséquence des relations (5.2 et (5.3) et

$$\psi(s+t,\cdot) = \psi(s,\cdot)\ \psi(t,\cdot), {}^h\mathbb{P}\text{-p.s.}, \tag{5.6}$$

conséquence du lemme 5.3.1.

Soit $t \in S_\sigma$. Il existe une suite $(t_n)_{n\geq 0}$ d'éléments de M convergeant vers t. Les relations (3.4.1) montrent que la suite de v.a.r. $(\psi(t_n,\cdot))_{n\geq 0}$ est de Cauchy, donc convergente dans L^2, et cette limite ne dépend pas de la suite $(t_n)_{n\geq 0}$ choisie. Nous notons $\psi(t,.)$ la v.a.r. limite. Les égalités (5.4), (5.5) et (5.6) sont encore vraies pour des éléments s et t de S_σ.

D'après (5.6), pour tout entier $p \geq 1$ et tout $t \in S_\sigma$, nous avons :

$$(\psi(t,\cdot))^p = \psi(pt,\cdot), {}^h\mathbb{P}\text{-p.s.},$$

ce qui montre que la v.a.r. $\psi(t,\cdot)$ appartient à tous les espaces $L^r(\Omega, \mathcal{F}, {}^h\mathbb{P})$, $r \geq 1$.

Les relations (5.5) montrent, via la théorie des martingales, que, pour tout $t \in S_\sigma$, le processus $(\frac{h(t+X_n)}{h(X_n)})_{n\geq 0}$ converge ${}^h\mathbb{P}$-p.s. et en norme L^r, $r \geq 1$, vers $\psi(t,\cdot)$. En particulier, on en déduit que

$$\psi(t,\cdot) = {}^h\xi(s,\cdot), {}^h\mathbb{P}\text{-p.s.}$$

Deux mesures positives sur les boréliens de $S^{\mathbb{N}^*}$ sont égales si, et seulement si, elle coïncident sur les fonctions boréliennes bornées (ou positives) ne dépendant que d'un nombre fini de coordonnées. Soit f une fonction borélienne sur S^n, $n \geq 1$. Nous avons, si $h(t) > 0$,

$$\begin{aligned}
\int_\Omega f(Y_1, \ldots, Y_n) \qquad {}^h\xi(t,\cdot)\ d^h\mathbb{P} &= \int_\Omega f(Y_1, \ldots, Y_n)\ {}^h\mathbb{E}[{}^h\xi(t,\cdot)|\mathcal{F}_n]\ d^h\mathbb{P} \\
&= \int_\Omega f(Y_1, \ldots, Y_n)\ \frac{h(t+X_n)}{h(X_n)}\ d^h\mathbb{P} \\
&= \int_{S^n} f(x_1, \ldots, x_n)\ \frac{h(t + x_1 + \ldots + x_n)}{h(0)}\ \sigma(dx_1)\ldots\sigma(dx_n) \\
&= \frac{h(t)}{h(0)} \int_\Omega f(Y_1, \ldots, Y_n)\ d^{h^t}\mathbb{P}.
\end{aligned}$$

Lorsque $h(t) = 0$, alors ${}^h\xi(t\cdot) = 0$, ${}^h\mathbb{P}$-p.s. ; si bien que la mesure ${}^h\xi(t\cdot)\ h(0)\ {}^h\mathbb{P}$ est la mesure nulle.

Pour tous $s, t \in S_\sigma$, nous avons d'une part :

$$h(s+t)\ {}^{h^{s+t}}\mathbb{P} = {}^h\xi(s+t,\cdot)\ h(0)\ {}^h\mathbb{P} = {}^h\xi(s,\cdot)\ {}^h\xi(t,\cdot)\ h(0)\ {}^h\mathbb{P} = {}^h\xi(t,\cdot)\ h(s)\ {}^{h^s}\mathbb{P}$$

et d'autre part :

$$h(s+t)\ {}^{h^{s+t}}\mathbb{P} = {}^{h^s}\xi(t,\cdot)\ h(s)\ {}^{h^s}\mathbb{P};$$

d'où l'on déduit ${}^h\xi(s,\cdot) = {}^{h^s}\xi(t,\cdot)$, ${}^{h^s}\mathbb{P}$-p.s., si $h(s) > 0$.

Enfin, nous avons,

$$\begin{aligned} 1 &= \int_S \frac{h(t+X_n)}{h(X_n)}\ \sigma(dt) \\ &= \int_S {}^h\mathbb{E}[{}^h\xi(t,\cdot)|\mathcal{F}_n]\ \sigma(dt) \\ &= {}^h\mathbb{E}[\int_S {}^h\xi(t,\cdot)\ \sigma(dt)|\mathcal{F}_n] \quad \text{(Fubini)} \\ &\overset{n\to+\infty}{\longrightarrow} \int_S {}^h\xi(t,\cdot)\ \sigma(dt). \end{aligned}$$

□

<u>Preuve</u> du théorème 5.2.1 : On suppose à présent que S est un groupe abélien G. Soit h une fonction non nulle de H_+. Quitte à remplacer h par une translatée par un élément de G, nous pouvons supposer que $h(0) > 0$; ce qui implique (d'après l'équation fonctionnelle (5.1)) que $\forall t \in S_\sigma$, $h(-t) > 0$.

Première étape : Nous montrons que le théorème 5.2.2 reste vrai pour les éléments de G.

Soit t un élément de S_σ. D'après le théorème 5.2.2 appliqué à h^{-t}, nous avons :

$$\begin{aligned} h(0)\ {}^h\mathbb{P} &= {}^{h^{-t}}\xi(t,\cdot)\ h(-t)\ {}^{h^{-t}}\mathbb{P} \\ &\text{et } \forall u \in S_\sigma,\ {}^{h^{-t}}\xi(u,\cdot) = {}^h\xi(u,\cdot),\ {}^h\mathbb{P}\text{-p.s..} \end{aligned}$$

On en déduit que ${}^h\mathbb{P}[\{{}^h\xi(t,\cdot) > 0\}] = 1$ et par suite $h(t) > 0$. Nous venons donc de montrer que h est strictement positive sur S_σ ; ce qui implique que h est strictement positive sur le sous-groupe $S_\sigma - S_\sigma$ de G. D'autre part, nous avons :

$$h(-t)\ {}^{h^{-t}}\mathbb{P} = \frac{1}{{}^h\xi(t,\cdot)}\ h(0)\ {}^h\mathbb{P}.$$

Cette égalité appliquée à h^s, pour $s \in S_\sigma$, nous donne :

$$h(s-t)\ {}^{h^{s-t}}\mathbb{P} = \frac{1}{{}^{h^s}\xi(t,\cdot)}\ h(s)\ {}^{h^s}\mathbb{P} = \frac{1}{{}^h\xi(t,\cdot)}\ h(s)\ {}^{h^s}\mathbb{P} = \frac{{}^h\xi(s,\cdot)}{{}^h\xi(t,\cdot)}\ h(0)\ {}^h\mathbb{P}.$$

Pour tout entier $n \geq 0$, l'égalité de ces mesures en restriction à la tribu $\mathcal{F}_n$, nous donne

$${}^h\mathbb{E}[\frac{{}^h\xi(s,\cdot)}{{}^h\xi(t,\cdot)}|\mathcal{F}_n] = \frac{h(s-t+X_n)}{h(X_n)},\ {}^h\mathbb{P}\text{-p.s..}$$

D'où la convergence ${}^h\mathbb{P}$-p.s. et en norme L^1 de la martingale $(\frac{h(s-t+X_n)}{h(X_n)})_{n\geq 0}$ vers la v.a.r. $\frac{{}^h\xi(s,\cdot)}{{}^h\xi(t,\cdot)}$, nécessairement égale à ${}^h\xi(s-t,\cdot)$.

Nous venons de prouver que le théorème 5.2.2 reste vrai pour les éléments du sous-groupe $S_\sigma - S_\sigma$ de G. Le passage du sous-groupe $S_\sigma - S_\sigma$ à son adhérence G_σ se fait avec les même arguments qui ont permis dans la preuve du théorème 5.2.2 de passer de M à son adhérence S_σ.

Deuxième étape : Nous montrons que h est une moyenne d'exponentielles harmoniques.

Pour cela, il faut montrer l'existence d'une version ${}^h\overline{\xi}$ de ${}^h\xi$ telle que, pour ${}^h\mathbb{P}$-presque tout $\omega \in \Omega$, ${}^h\overline{\xi}(\cdot,\omega)$ soit une fonction de $\mathcal{E}_\sigma$. Pour simplifier l'écriture nous écrivons ξ au lieu de ${}^h\xi$.

Les relations

$$\forall g \in G_\sigma = G,\ \xi(g,\cdot)\ \xi(-g,\cdot)\ =\ \xi(0,\cdot) = 1,\ {}^h\mathbb{P}\text{-p.s.}$$

montrent que, pour tout $g \in G$, ${}^h\mathbb{P}[\{\xi(g,\cdot) > 0\}] = 1$ et h est strictement positive sur G. Nous commençons par prendre, pour tous $(g,\omega) \in G \times \Omega$,

$${}^h\xi(g,\omega) = \limsup_{n\to+\infty} \frac{h(g + X_n(\omega))}{h(X_n(\omega))} \in [0,+\infty].$$

Désignons par m une mesure de Haar sur le groupe abélien G. Il existe (Fubini) un borélien Ω_0 de Ω de ${}^h\mathbb{P}$-mesure pleine tel que pour tout $\omega \in \Omega_0$, $\xi(\cdot,\omega) > 0$, m-presque partout. Soient β une fonction continue positive non nulle, à support compact sur G. Il s'ensuit que, pour tout $\omega \in \Omega_0$, $\int_G \xi(u,\omega)\ \beta(u)\ m(du) > 0$. Comme

$${}^h\mathbb{E}[\int_G \xi(u,\omega)\ \beta(u)\ m(du)] < +\infty,$$

quitte à remplacer Ω_0 par un sous-borélien de ${}^h\mathbb{P}$-mesure pleine, nous pouvons supposer que, pour tout $\omega \in \Omega_0$, $0 < \int_G \xi(u,\omega)\ \beta(u)\ m(du) < +\infty$.

Pour tout $g \in G$ et tout $\omega \in \Omega_0$, posons

$$\overline{\xi}(g,\omega) = \frac{\int_G \xi(g+u,\omega)\ \beta(u)\ m(du)}{\int_G \xi(u,\omega)\ \beta(u)\ m(du)}.$$

Du théorème de Fubini et des égalités

$$\begin{aligned}&\forall g,x \in G,\ \xi(g+x,\cdot) = \xi(g,\cdot)\ \xi(x,\cdot),\ {}^h\mathbb{P}\text{-p.s.}\\ &\text{et } \int_S \xi(t,\cdot)\ \sigma dt\ =\ 1,\ {}^h\mathbb{P}\text{-p.s.};\end{aligned}$$

il en résulte que

$$\begin{gathered}\forall g \in G,\ \overline{\xi}(g,\cdot) = \xi(g,\cdot),\ {}^h\mathbb{P}\text{-p.s.}\\ \forall g,x \in G,\ \overline{\xi}(g+x,\cdot) = \overline{\xi}(g,\cdot)\ \overline{\xi}(x,\cdot),\ {}^h\mathbb{P}\text{-p.s.}\\ \text{et } \int_S \overline{\xi}(t,\cdot)\ \sigma dt\ =\ 1,\ {}^h\mathbb{P}\text{-p.s..}\end{gathered}$$

Soit K un compact de G. Pour tous $g,x \in K$ et pour toute fonction ψ sur G à support compact, strictement positive sur un voisinage de $K + Supp\ \beta$, nous avons

$$\begin{aligned}|\overline{\xi}(x,\omega) - \overline{\xi}(g,\omega)| &\leq \int_G \xi(u,\omega)\ |\beta(u-x) - \beta(u-g)|\ m(du)\\ &\leq \sup_{u\in G}\{\frac{|\beta(u-x) - \beta(u-g)|}{\psi(u)}\}\ \int_G \xi(u,\omega)\ \psi(u)\ m(du).\end{aligned}$$

Ce qui montre que pour $\omega \in \Omega_0$ tel que $\int_G \xi(u,\omega)\ \psi(u)\ m(du) < +\infty$, $\overline{\xi}(\cdot,\omega)$ est continue sur K. En prenant une suite de compacts $(K_n)_{n\geq 0}$ telle que $\bigcup_{n\geq 0} K_n = G$, on obtient alors qu'il existe un sous-borélien Ω_1 de Ω_0 de ${}^h\mathbb{P}$-mesure pleine tel que, pour tout $\omega \in \Omega_1$, la fonction $\overline{\xi}(\cdot, \omega)$ est continue sur G.

Soient D un sous-ensemble dénombrable dense dans G, il existe un sous-borélien Ω_2 de Ω_1 de ${}^h\mathbb{P}$-mesure pleine tel que, pour tout $\omega \in \Omega_2$,

$$\begin{aligned} \forall g, x \in D, \qquad \overline{\xi}(g+x, \omega) &= \overline{\xi}(g, \omega)\ \overline{\xi}(x, \omega) \\ \text{et} \quad \int_S \overline{\xi}(t, \omega)\ \sigma dt &= 1. \end{aligned}$$

De la continuité des fonctions $\overline{\xi}(\cdot, \omega)$, $\omega \in \Omega_2$, il résulte alors que, pour tout $\omega \in \Omega_2$, la fonction $\overline{\xi}(\cdot, \omega)$ est une exponentionnelle harmonique.

Considérons l'application de Ω_2 dans $\mathcal{E}_\sigma$ qui à ω associe l'exponentielle harmonique $\overline{\xi}(\cdot, \omega)$. Appelons μ l'image de la mesure ${}^h\mathbb{P}$ par cette application, nous avons :

$$\forall g \in G, \ h(g) = \int_{\mathcal{E}_\sigma} \chi(g)\ \mu(d\chi).$$

Troisième étape : Nous montrons que si $S_\sigma = G$, alors $\mathcal{E}_\sigma$ est compact.

Posons $\lambda = \sum_{k \geq 1} \frac{1}{2^k}\ \sigma^k$. La fonction h est λ-harmonique continue, strictement positive ; par suite la mesure λ est une mesure de Radon. Quitte à remplacer σ par λ, nous pouvons supposer que le support de σ est G tout entier. Nous avons besoin du lemme élémentaire suivant :

Lemme 5.3.3 *Pour toute fonction α continue positive, à support compact sur G, il existe une fonction continue positive δ, nulle en 0, telle que*

$$\forall g, x \in G, \ |h^\alpha(g+x) - h^\alpha(x)| \leq \delta(g)\ h^\alpha(x)$$

où h^α désigne la fonction qui à $g \in G$ associe $\int_G h(x+g)\ \alpha(x)\ m(dx)$.

Preuve : Puisque le support de σ est G tout entier, la fonction

$$\Phi(g) = \int_G \alpha(g - x)\ \sigma(dx)$$

est continue et strictement positive sur G. Nous avons alors :

$$\begin{aligned} |h^\alpha(g+x) - h^\alpha(x)| &\leq \int_G h(u) |\alpha(u-g-x) - \alpha(u-x)|\ m(du) \\ &\leq \delta(g) \int_G h(u)\ \Phi(u)\ m(du) \\ &\leq \delta(g) \int_G h^\alpha(x+u)\ \sigma(du) = \delta(g)\ h^\alpha(x)\ ; \end{aligned}$$

en posant, $\delta(g) = \sup_{u \in G} \frac{|\alpha(u-g) - \alpha(u)|}{\Phi(u)}$. □

Cela dit, soit α une fonction continue positive, à support compact sur G. La martingale

$$\Big(\frac{h^\alpha(x + X_n)}{h(X_n)} = {}^h\mathbb{E}[\int_G \overline{\xi}(u+x, \cdot)\ \alpha(u)\ m(du) | \mathcal{F}_n]\Big)_{n \geq 0}$$

converge ${}^h\mathbb{P}$-p.s. et en norme L^1 vers

$$\int_G \overline{\xi}(x+u, \cdot)\ \alpha(u)\ m(du) = \overline{\xi}(x, \cdot) \int_G \overline{\xi}(u, \cdot)\ \alpha(u)\ m(du).$$

On en déduit que le processus, $(\dfrac{h^\alpha(x+X_n)}{h^\alpha(X_n)})_{n\geq 0}$ converge ${}^h\mathbb{P}$-p.s. vers $\overline{\xi}(x,\cdot)$. D'après le lemme, via la séparabilité de G et la continuité de $\overline{\xi}(\cdot,\omega)$, pour $\omega \in \Omega_2$, il existe alors un sous-borélien Ω_3 de Ω_2 de ${}^h\mathbb{P}$-mesure 1, tel que :

$$\forall\omega \in \Omega_3,\ \forall g \in G,\ |\overline{\xi}(x,\cdot) - 1| \leq \delta(x).$$

Ceci montre, d'après le théorème d'Ascoli, la relative compacité de la famille $\{\overline{\xi}(.,\omega) : \omega \in \Omega_3\}$. Comme δ ne dépend pas de h, il s'ensuit que $\mathcal{E}_\sigma$ est compact.

Quatrième étape : Nous montrons l'unicité de la mesure μ.

• **Cas particulier** : Nous commençons par traiter le cas où $\mathcal{E}_\sigma$ est compact.

Appelons A l'espace vectoriel sur $\mathbb{R}$ des fonctions sur $\mathcal{E}_\sigma$ de la forme

$$\chi \longrightarrow \sum_{k=1}^{n} \alpha_k\ \chi(g_k)\ ;\ \text{avec } n \in \mathbb{N}^* \text{ et } \forall 1 \leq k \leq n,\ (\alpha_k, g_k) \in \mathbb{R}\times G.$$

Muni du produit des fonctions, A est une algèbre qui sépare les points de $\mathcal{E}_\sigma$ et contient les constantes.

D'après le théorème de Stone-Weierstrass, A est dense dans l'espace $C(\mathcal{E}_\sigma)$ des fonctions continues sur $\mathcal{E}_\sigma$, pour la topologie de la convergence uniforme.

Deux mesures positives sur les boréliens de $\mathcal{E}_\sigma$, qui représentent la même fonction non nulle h de H_+, coïncide sur A et donc sur $C(\mathcal{E}_\sigma)$. D'où l'égalité de ces deux mesures.

• **Cas général** :

Il existe une suite croissante $(V_p)_{p\geq 0}$ d'ouverts symétriques (i.e. $-V_p = V_p$), relativement compacts de G telle que $\bigcup\limits_{p\geq 0} V_p = G$. A chaque entier $p \geq 0$, nous associons le sous-groupe ouvert, et donc fermé, de G défini par $G_p = \mathbb{Z}\ V_p$. Nous avons $G = \bigcup\limits_{p\geq 0} G_p$. Pour tout sous-groupe fermé H de G, l'ensemble des exponentielles $\mathcal{E}(H)$ sur H s'identifie, via la fonction logarithme, à l'espace vectoriel $E(H)$ des homomorphismes de groupe de H dans $(\mathbb{R}, +)$.

Soit p un entier naturel. Pour tout $\chi \in E(G_p)$, posons $\|\chi\| = \sup\limits_{g\in V_p} |\chi(g)|$. On définit ainsi une norme sur $E(G_p)$. Pour cette norme, la restriction à tout compact de G_p des fonctions de la boule unité constitue une famille de fonctions bornée et équicontinue (noter que $\forall r \in \mathbb{N}^*,\ \sup\limits_{g\in rV_p} |\chi(g)| = r\ \|\chi\|$).

Il s'ensuit (théorème d'Ascoli) que cette boule unité est compacte pour la convergence uniforme sur les compacts.

D'après le théorème de Riesz, l'espace vectoriel normé $E(G_p)$ est de dimension finie. On en déduit alors que $E(G)$ est une limite projective d'une suite croissante d'espaces vectoriels de dimension finie. Autrement dit $E(G)$ s'identifie : soit à $\mathbb{R}^q$, $q \in \mathbb{N}^*$; soit à $\mathbb{R}^{\mathbb{N}^*}$.

Nous nous plaçons dans le cas le plus défavorable et, pour tout $p \geq 0$, nous notons π_p la projection de $E(G)$ sur $E(G_p)$ (i.e. pour $\chi \in E(G)$, $\pi_p(\chi)$ est la restriction de χ à G_p). D'après le théorème de Kolmogorov, pour montrer que deux mesures de probabilités sur les boréliens de $E(G)$ sont égales, il suffit de montrer que leurs projections sur $E(G_p)$, $p \geq 0$, sont égales.

Pour tout $g \in G$, il existe un entier $p \geq 0$ tel que g appartienne à G_p. La fonction continue $L_g\ : \chi \in E(G) \longrightarrow \chi(g)$ est donc la composée par π_p d'une fonction continue $\overline{L_g}$ sur $E(G_p)$. (Autrement dit, pour tout $g \in G$, la fonction L_g sur $E(G)$ identifié à $\mathbb{R}^{\mathbb{N}^*}$ ne dépend que d'un nombre fini de coordonnées).

Cela dit soit μ une probabilité sur les boréliens de $E(G)$, portée par les exponentielles harmoniques, telle que

$$\forall g \in G,\ h(g) = h(0) \int_{E(G)} e^{\chi(g)}\ \mu(d\chi).$$

Soit p un entier ≥ 0. nous appelons $C_0(E(G_p))$ l'espace des fonctions continues sur $E(G_p)$ tendant vers zéro à l'infini. Pour toute fonction positive F de $C_0(E(G_p))$, la fonction

$$f(g) = h(0) \int_{E(G)} F(\pi_p(\chi))\ e^{\chi(g)}\ \mu(d\chi), (g \in G),$$

est σ-harmonique continue positive. La martingale $\left(\dfrac{f(X_n)}{h(X_n)}\right)_{n\geq 0}$ est bornée ; elle converge donc ${}^h\mathbb{P}$-p.s. et en norme L^q, $q \geq 1$, vers une v.a.r. $\theta(\cdot, F)$. L'espace $C_0(E(G_p))$ étant séparable, il existe un sous-borélien Ω_4 de Ω_3, de ${}^h\mathbb{P}$-mesure pleine tel que, pour tout $\omega \in \Omega_4$ et tout $F \in C_0(E(G_p))$, la suite

$$\frac{\int_{E(G)} F(\pi_p(\chi))\ e^{\chi(X_n(\omega))}\ \mu(d\chi)}{\int_{E(G)} e^{\chi(X_n(\omega))}\ \mu(d\chi)}$$

converge vers $\theta(\omega, F)$; autrement dit, la suite de probabilités sur $E(G_p)$

$$\rho_n(\omega, d\chi) = \pi_p\Big[\frac{e^{\chi(X_n(\omega))}}{h(X_n(\omega))}\ \mu(d\chi)\Big]$$

converge vaguement vers la mesure $\theta(\omega, d\chi)$. Pour toute fonction $F \in C_0(E(G_p))$, nous avons :

$$\int_\Omega \rho_n(\cdot, F)\ d^h\mathbb{P} = \int_{E(G_p)} F\ d(\pi_p(\mu)) = \int_\Omega \theta(\cdot, F)\ d\ {}^h\mathbb{P}.$$

Ce qui montre (en prenant une suite de fonctions de $C_0(E(G_p))$ croissant vers la fonction identique à 1) que la convergence vague de la suite de probabilités $\rho_n(\omega, d\chi)$ vers $\theta(\omega, d\chi)$ est étroite, pour tout ω appartenant à un sous-borélien Ω_5 de Ω_4 de ${}^h\mathbb{P}$-mesure pleine.

Pour tout $g \in G_p$, la fonction $\overline{L_g}$ est une fonction positive de $C(E(G_p))$. Nous avons donc, pour tout $\omega \in \Omega_5$,

$$\liminf_{n\to+\infty} \rho_n(\omega, \overline{L_g}) \geq \theta(\omega, \overline{L_g}).$$

Or nous avons :

$$\rho_n(\omega, \overline{L_g}) = \frac{h(g + X_n)}{h(X_n)} \overset{{}^h\mathbb{P}\text{-p.s.}}{\longrightarrow} \overline{\xi}(g, \omega);$$

d'où $\theta(\omega, \overline{L_g}) \leq \overline{\xi}(g, \omega)$. Les deux membres de cette égalité ayant même espérance par rapport à ${}^h\mathbb{P}$, il s'ensuit que l'on a égalité sur un sous-borélien Ω_6 de Ω_5 de ${}^h\mathbb{P}$-mesure pleine. Cette égalité implique alors que, pour tout entier $r \geq 1$ et tout $g \in G_p$,

$$\overline{\xi}(g, \omega) = [\overline{\xi}(r\ g, \omega)]^{\frac{1}{r}} = \theta\Big(\omega, (\overline{L_g})^r]\Big)^{\frac{1}{r}}.$$

En faisant tendre r vers $+\infty$, on voit que $\theta(\omega, d\chi)$ est la mesure de Dirac en l'exponentielle $\pi_p(\overline{\xi}(\cdot, \omega))$. D'où l'on déduit que

$$\pi_p(\mu)(d\chi) = \int_\Omega \theta(\cdot, d\chi)\ d^h\mathbb{P} = \int_\Omega \varepsilon_{\pi_p(\overline{\xi}(\cdot,\omega))}\ d^h\mathbb{P}.$$

Ce qui détermine de façon unique la probabilité $\pi_p(\mu)$. $\square$

5.4 Exemples

Exemple 1 : Considérons le groupe $G = (\mathbb{R}^d, +)$. Les exponentielles sur G sont les fonctions χ_u, $u \in \mathbb{R}^d$, définies par :

$$\forall x \in \mathbb{R}^d, \ \chi_u(x) = e^{\langle u,x \rangle} \ ;$$

où $\langle \ , \ \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^d$. $E(G)$ s'identifie donc à $\mathbb{R}^d$. Nous avons :

$$\mathcal{E}_\sigma = \{\chi_u : \int_{\mathbb{R}^d} e^{\langle u,x \rangle} \ dx = 1\}.$$

La fonction $F(u) = \int_{\mathbb{R}^d} e^{\langle u,x \rangle} \ \sigma(dx)$ est Log-strictement convexe.

Pour $d = 1$, il existe $0, 1$ ou 2 exponentielles harmoniques suivant que le minimum de F est $>, =$ ou $<$ à 1. Dans le cas général, $\mathcal{E}_\sigma$ est un fermé de $\mathbb{R}^d$ et est donc localement compact à base dénombrable.

Exemple 2 : Considérons le groupe multiplicatif G des rationnels strictement positifs $(\mathbb{Q}_+^*, \times)$, muni de la topologie discrète. Désignons par $\mathcal{P}$ l'ensemble des entiers naturels premiers. Tout $r \in \mathbb{Q}_+^*$ s'écrit de façon unique sous la forme $r = p_1^{\alpha_1} \cdots p_N^{\alpha_n}$ avec $N \in \mathbb{N}^*$, $p_1, \ldots p_N \in \mathcal{P}$ et $\alpha_1, \ldots \alpha_N \in \mathbb{Z}$. Toute exponentielle χ sur $\mathbb{Q}_+^*$ est déterminée par ses valeurs sur les entiers premiers. Réciproquement si l'on se donne une famille $\{a_p : p \in \mathcal{P}\}$ de nombre réels strictement positifs, la fonction définie par $\chi(r) = a_1^{\alpha_1} \cdots a_N^{\alpha_n}$, si $r = p_1^{\alpha_1} \cdots p_N^{\alpha_n}$, est une exponentielle sur $\mathbb{Q}_+^*$. L'espace vectoriel $E(G)$ s'identifie donc à $]0, +\infty[^{\mathcal{P}}$ muni de la topologie produit ; $E(G)$ n'est donc pas localement compact.

Chapitre 6

Opérateur de transition (cas équicontinu)

Sous des hypothèses d'équicontinuité, nous faisons l'étude spectrale d'un opérateur de transition, dans un cadre plus général que celui de la quasi-compacité. (Voir [Rau92])

6.1 Ensembles P-invariants et fonctions harmoniques

Dans toute la suite, nous désignons par (E, d) un espace métrique compact et par P une probabilité de transition sur E qui **respecte les fonctions continues** ; c'est-à-dire telle que pour toute fonction continue f, la fonction Pf est continue.

Définitions - Notations 6.1.1 Nous notons $P^n, n \geq 1$, les itérées de P, en convenant que P^0 est l'opérateur identité ou encore que, pour tout $x \in E, P^0(x, \cdot)$ est la mesure de Dirac en x.

Si K est un sous ensemble de E et x un élément de E, nous notons $d(x, K)$ la distance de x à K : $d(x, K) = \inf_{y \in K} d(x, y)$.

Un borélien B de E est dit P-**invariant** (ou simplement **invariant**) si, $\forall x \in B$, $P(x, B) = 1$.

Notons que cette condition équivaut à $P1_B \geq 1_B$. Il n'y a pas nécessairement égalité ; par contre, il y a égalité μ-p.p. pour toute mesure de probabilité μ P-invariante. Certains auteurs utilisent le terme "P-absorbant".

Si σ est une mesure positive sur les boréliens de E, nous appelons **support** de σ et nous notons $\mathrm{supp}(\sigma)$ l'ensemble des éléments x de E tels que $\sigma(V) > 0$, pour tout voisinage V de x. On voit facilement que $\mathrm{supp}(\sigma)$ est aussi l'ensemble des éléments x de E tels que $\int_E f(y)\,\sigma(dy) > 0$ pour toute fonction continue positive sur E vérifiant $f(x) > 0$.

Pour tout $x \in E$, nous notons $G(x, \cdot)$ la mesure **potentiel** définie par : $G(x, \cdot) = \sum_{r \geq 0} P^r(x, \cdot)$.

Lemme 6.1.2 *Pour tout élément u de E et tout entier $r \geq 1$,*
i) $x \in \mathrm{supp}\, P(u, \cdot)$ et $y \in \mathrm{supp}\, P^r(x, \cdot)$, implique $y \in \mathrm{supp}\, P^{r+1}(u, \cdot)$
ii) $x \in \mathrm{supp}\, P(u, \cdot)$ et $y \in \mathrm{supp}\, G(x, \cdot)$, implique $y \in \mathrm{supp}\, G(u, \cdot)$.

Preuve : Soit f une fonction continue positive telle que $f(y) > 0$. Ecrivons :

$$P^{r+1}f(u) = \int_E P^r f(z)\, P(u, dz).$$

Comme P respecte $C(E)$, la fonction $P^r f$ est continue positive. De plus, puisque $y \in \operatorname{supp} P^r(x, \cdot)$, nous avons $P^r f(x) > 0$. Il s'ensuit qu'il existe un voisinage V de x tel que $\forall\, z \in V,\ P^r f(z) > \frac{1}{2} P^r f(x)$. Il vient alors

$$P^{r+1}f(u) \geq \int_V P^r f(z)\, P(u, dz) \geq \frac{1}{2} P^r f(x) P(u, V) > 0.$$

D'où la première assertion. La seconde affirmation résulte de la première, après avoir remarqué que, pour tout élément x de E, $\operatorname{supp} G(x, \cdot)$ est la fermeture dans E de la réunion $\bigcup_{r \geq 0} \operatorname{supp} P^r(x, \cdot)$. $\square$

• **Etude des fonctions P-harmoniques continues**

Rappelons qu'une fonction borélienne bornée h, à valeurs réelles ou complexes, est dite P-harmonique si elle vérifie la relation $Ph = h$. Nous désignons par $\mathcal{H}$ l'espace vectoriel des fonctions P-harmoniques boréliennes bornées et par $\mathcal{H}_c$ celui des fonctions P-harmoniques continues.

Si $h \in \mathcal{H}$, la fonction h^2 vérifie $Ph^2 \geq (Ph)^2 = h^2$. Il s'ensuit que la suite de fonctions $(P^n(h^2))_{n \geq 0}$ est croissante et converge vers une fonction de $\mathcal{H}$ que nous noterons $h * h$. Pour deux fonctions h_1 et h_2 de $\mathcal{H}$, nous définissons alors

$$\begin{aligned} h_1 * h_2 &= \frac{1}{4}\left[(h_1 + h_2) * (h_1 + h_2) - (h_1 - h_2) * (h_1 - h_2)\right] \\ &= \lim_{n \to +\infty} P^n(h_1 h_2). \end{aligned}$$

Muni de ce produit et de la norme uniforme $\| \ \|_\infty$, $\mathcal{H}$ est une algèbre de Banach. Nous notons $\mathrm{Alg}(\mathcal{H}_c)$ la sous-algèbre *fermée* de $\mathcal{H}$ engendrée par $\mathcal{H}_c$.

6.2 Description de l'algèbre $\mathcal{H}_c$

Soit h une fonction de $\mathcal{H}_c$. Alors l'ensemble des points de E, où h atteint son maximum (resp. son minimum) est un compact invariant de E. Si h n'est pas constante, on obtient ainsi deux compacts invariants disjoints. Dans ce qui suit nous précisons le lien existant entre les fonctions harmoniques continues et les sous-ensembles compacts invariants de E.

Nous mettrons en évidence une famille $\mathcal{G}$ de sous-ensembles compacts invariants de E, deux à deux disjoints, telle que, pour presque toute trajectoire $\omega \in \Omega$, il existe un compact invariant $\eta(\omega)$ de $\mathcal{G}$ vérifiant

$$\lim_{n \to +\infty} d(X_n(\omega), \eta(\omega)) = 0.$$

Les éléments de $\mathcal{G}$ jouent en quelque sorte le rôle des "classes ergodiques" dans la théorie classique des chaînes de Markov vérifiant la condition de Doeblin ([Doo53].

Nous munirons $\mathcal{G}$ d'une distance qui en fait un espace métrique compact et nous construirons une probabilité de transition ρ de E dans $\mathcal{G}$ telle que l'application

$$\hat{h} \to h(x) = \int_{\mathcal{G}} \hat{h}(\gamma)\, \rho(x, d\gamma)$$

soit une isométrie d'algèbre de Banach de $C(\mathcal{G})$ sur $\mathrm{Alg}(\mathcal{H}_c)$.

Une suite $(x_n)_{n\geq 0}$ d'éléments de E converge vers $\gamma \in \mathcal{G}$, si la suite $(x_n)_{n\geq 0}$ a toutes ses valeurs d'adhérence dans γ, c'est-à-dire $d(x_n, \gamma) \to 0$. L'espace $\mathcal{G}$ apparaît alors comme une frontière de E et la formule $h(x) = \int_{\mathcal{G}} \hat{h}(\gamma)\, \rho(x, d\gamma)$, comme une formule de Poisson. A tout élément h de $\mathcal{H}_c$, nous avons associé une fonction continue $\hat{h}$ sur la frontière $\mathcal{G}$ de E, qui intégrée par le "noyau de Poisson" ρ permet de retrouver h.

La description des fonctions harmoniques continues a été ramenée à la recherche de sous-ensembles compacts invariants disjoints de E. Dans de nombreux cas, cette recherche revêt un caractère géométrique.

Fonctions harmoniques et variables aléatoires invariantes

En reprenant les notions introduites au chapitre 3 et les notations de 3.4, nous désignons par $\mathcal{J} = \{B \in \mathcal{F} : \theta^{-1}(B) = B\}$ la tribu formée des sous-ensembles boréliens de Ω invariants par le décalage θ. Une variable aléatoire Z est $\mathcal{J}$-mesurable si et seulement si $Z \circ \theta = Z$; une telle variable sera dite invariante par θ. Pour tout entier $n \geq 0$, nous notons $\mathcal{F}_n$ la sous-tribu de $\mathcal{F}$ engendrée par les variables aléatoires $\{X_k : 0 \leq k \leq n\}$.

Deux variables aléatoires, Z et Z', invariantes par θ sont dites équivalentes si $\forall\, x \in E$, $\mathbb{P}_x[Z \neq Z'] = 0$. Nous appelons $\mathcal{V}$ l'algèbre des classes d'équivalence des variables aléatoires invariantes par θ.

D'après la proposition 3.5.1 du chapitre 3, nous avons :

Proposition 6.2.1 *L'application qui à la classe Z de $\mathcal{V}$ associe la fonction P-harmonique $h(x) = \mathbb{E}_x[Z]$ défini un isomorphisme d'algèbre de Banach.*

Lemme 6.2.2 *Soit h une fonction de $\mathcal{H}$. Pour tout entier $r \geq 1$ et tout $x \in E$, nous avons :*

$$\sum_{n\geq 0} \mathbb{E}_x\Big[\int_E [h(y) - h(X_n)]^2\, P^r(X_n, dy)\big] \;<\; +\infty.$$

<u>Preuve</u> : Il suffit d'observer que :

$$\mathbb{E}_x\Big[\int_E [h(y) - h(X_n)]^2\, P^r(X_n, dy)\big] = \mathbb{E}_x[h^2(X_{n+r})] - \mathbb{E}_x[h^2(X_n)].$$

□

Corollaire 6.2.3 *Pour tout $x \in E$, tout $h \in \mathcal{H}$ et tout entier $r \geq 1$, la suite de variables aléatoires $\left(\int_E [h(y) - h(X_n)]^2\, P^r(X_n, dy)\right)_{n\geq 0}$ converge vers zéro, $\mathbb{P}_x$-presque sûrement.*

• Classes ergodiques

Notation : Soit $h \in \mathcal{H}_c$. Nous appelons Ω_h le sous-ensemble $\mathcal{F}$-mesurable et θ-invariant de Ω constitué des éléments ω tels que :

i) la suite $\Big(h(X_n(\omega))\Big)_{n\geq 0}$ converge ;
ii) pour tout entier $r \geq 1$, la suite

$$\Big(\int_E [h(x) - h(X_n(\omega))]^2 \, P^r(X_n(\omega), dx)\Big)_{n\geq 0}$$

converge vers zéro.

D'après ce qui précède, nous savons que $\mathbb{P}_x[\Omega_h] = 1, \forall x \in E$.

L'espace $C(E)$ est séparable (i.e possède une suite d'élements dense), E étant un espace métrique compact. Comme $\mathcal{H}_c$ est un sous-espace vectoriel fermé de $C(E)$, il est également séparable. Nous désignons par $(h_p)_{p\geq 0}$ une famille dénombrable de fonctions non nulles de $\mathcal{H}_c$ dense dans $\mathcal{H}_c$.

En posant $\Omega_0 = \cap_{p\geq 0}\Omega_{h_p}$, nous obtenons alors un sous-ensemble de Ω, $\mathcal{F}$-mesurable, invariant par θ de $\mathbb{P}_x$-mesure 1 pour tout $x \in E$ et vérifiant $\Omega_0 = \cap_{h\in\mathcal{H}_c}\Omega_h$.

Soient ω un élément de Ω_0 et u une valeur d'adhérence dans E de la suite $(X_n(\omega))_{n\geq 0}$. Puisque P especte $C(E)$, il résulte de la propriété *ii)* de (2.7) que, pour tout $h \in \mathcal{H}_c$ et tout $r \geq 1$, $\int_E [h(x) - h(u)]^2 \, P^r(u, dx) = 0$. Autrement dit, *toute fonction P-harmonique continue h est constante et égale à $h(u)$ sur* $\operatorname{supp} G(u, \cdot)$.

Ceci nous amène à considérer le sous-ensemble compact, non vide, de E défini par :

$$F = \{u \in E : \forall\ h \in \mathcal{H}_c, \forall\ x \in \operatorname{supp} G(u, \cdot), h(x) = h(u)\}.$$

Du lemme 6.1.2, il résulte que F est invariant.

Deux éléments u et v de F sont dits **équivalents** si $\forall\ h \in \mathcal{H}_c, h(u) = h(v)$. Il est clair que nous définissons ainsi une relation d'équivalence sur F. Chaque classe d'équivalence fournit un sous-ensemble compact invariant de F. Nous désignons par $\mathcal{G}$ l'ensemble de ces classes d'équivalences. Pour tout $h \in \mathcal{H}_c$ et tout $\gamma \in \mathcal{G}$, notons $\hat{h}(\gamma)$ la valeur de h sur γ. Nous munissons $\mathcal{G}$ de la distance δ définie à l'aide d'une suite (h_p) dense dans $\mathcal{H}_c$ par :

$$\delta(\gamma_1, \gamma_2) = \sum_{p\geq 0} \frac{1}{2^p} \frac{|\widehat{h_p}(\gamma_1) - \widehat{h_p}(\gamma_2)|}{\|h_p\|_\infty}.$$

Définition 6.2.4 Nous donnerons le nom de **classe ergodique** à chaque élément de $\mathcal{G}$.

Lemme 6.2.5 *$(\mathcal{G}, \delta)$ est un espace métrique compact. Il existe une application mesurable η de Ω dans $\mathcal{G}$ telle que :*
i) $\forall\omega \in \Omega_0, \lim_{n\to+\infty} d(X_n(\omega), \eta(\omega)) = 0.$
ii) $\eta \circ \theta = \eta$.

Preuve : Soit $(\gamma_n)_{n\geq 0}$ une suite de classes ergodiques. Pour tout entier naturel n, nous choisissons un élément u_n de γ_n. Le procédé diagonal nous fournit une suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$ telle que pour tout entier $p \geq 0$, la suite $\Big(h_p(u_{\varphi(n)})\Big)_{n\geq 0}$ converge. Quitte à prendre une sous-suite de $(\varphi(n))_{n\geq 0}$, on peut supposer que la suite $(u_{\varphi(n)})_{n\geq 0}$ converge vers un élément u de F. Si γ est la classe ergodique qui contient u, il est alors clair que : $\lim_{n\to+\infty} \delta(\gamma_{\varphi(n)}, \gamma) = 0$. Ceci établit la compacité de $\mathcal{G}$ muni de la distance δ.

Soit $\omega \in \Omega_0$. Nous avons vu que toutes les valeurs d'adhérence de la suite $(X_n(\omega))_{n\geq 0}$ appartiennent à F. Or nous savons aussi que, pour tout $h \in \mathcal{H}_c$, la suite $\Big(h(X_n(\omega))\Big)_{n\geq 0}$ converge. Il s'ensuit que toutes les valeurs d'adhérence appartiennent nécessairement à la même classe ergodique $\eta(\omega)$.

Il est clair que l'application η de Ω_0 dans $\mathcal{G}$ ainsi obtenue est invariante par θ. Nous montrons qu'elle est $\mathcal{F}$-mesurable. Soit C une boule fermée de $\mathcal{G}$. Le sous-ensemble $B = \bigcup_{\gamma\in C}\gamma$ de F est compact. [En effet si x est un élément de F adhérent à B, il existe une suite $(x_n)_{n\geq 0}$ d'éléments de B convergeant vers x telle que : $\forall\ n \geq 0, x_n \in \gamma_n \in C$. D'après la compacité de $\mathcal{G}$, il existe une suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$ telle que $\gamma_{\varphi(n)} \xrightarrow{\delta} \gamma \in C$, ce qui implique $x \in \gamma \subset B$.]

Pour tout $\omega \in \Omega_0$, nous avons alors :

$$\eta^{-1}(C) = \Big\{\omega \in \Omega_0 : \lim_{n\to+\infty} d(X_n(\omega), B) = 0\Big\};$$

ce qui montre que η est $\mathcal{F}$-mesurable.

On donne à η une valeur arbitraire sur le complémentaire de Ω_0, qui est aussi un sous-ensemble de Ω invariant par θ. $\square$

Exemples :

Considérons une probabilité de transition N de la forme

$$Nf(x) = \sum_{s\in S} p(s,x)f(s.x), \tag{6.1}$$

où S est une famille finie ou dénombrable de transformations continues sur un espace métrique compact E et les poids $p(s,.)$, $s \in S$, sont des fonctions continues positives telles que

$$\sum_{s\in S} p(s,x) = 1, \forall x \in E.$$

Supposons que les deux conditions suivantes soient satisfaites :

i) $d(s\cdot x, s\cdot y) < d(x,y), \forall s \in S, x,y \in E$,

ii) il existe $\delta > 0$ tel que, pour tous $x, y \in E$ vérifiant $d(x,y) < \delta$, on ait $p(s,x)p(s,y) > 0$, pour au moins un élément s de S.

Proposition 6.2.6 *Sous les conditions i) et ii) précédentes, si F et G sont deux compacts invariants disjoints de E, alors $d(F,G) \geq \delta$.*

<u>Preuve</u> : Supposons qu'on ait $d(F,G) < \delta$. Soient $x \in F$ et $y \in G$ tels que $d(x,y) = d(F,G)$.

D'après les hypothèses, il existe $s \in S$ tel que $s\cdot x \in F$ et $s\cdot y \in G$. On en déduit que $d(F,G) = 0$. $\square$

Sous les conditions de la proposition, il n'y a donc qu'un nombre fini de classes ergodiques pour le noyau N. C'est en particulier le cas pour les opérateurs de transfert des exemples 2.1 et 2.2 du chapitre 3.

• Mesures invariantes et classes ergodiques

Il est clair que, pour tout compact invariant K et tout élément $x \in K$, les valeurs d'ahérence de la suite de mesures $(\frac{1}{n}\sum_0^{n-1} P^k(x,\cdot)$ sont des mesures de probabilité P-invariantes portées par K.

Proposition 6.2.7 *Toute mesure de probabilité P-invariante ν est portée par le compact invariant F défini en (2.7). Si ν est extrémale, elle est portée par une classe ergodique.*

Preuve : Soit $h \in \mathcal{H}_c$. Pour tout entier $r \geq 1$, nous avons :

$$P^r(h^2) \geq (P^r h)^2 = h^2 \text{ et } \int_E [P^r(h^2) - h^2](u)\, \nu(du) = 0.$$

Il s'ensuit que, pour tout $u \in \operatorname{supp}(\nu)$ et tout $x \in \operatorname{supp} G(u, \cdot), h(x) = h(u)$; d'où $\operatorname{supp}(\nu) \subset F$.

Nous avons alors :

$$\begin{aligned}
\int_\gamma Pf(u)\, \nu(du) &= \int_\gamma P(1_\gamma f)(u)\, \nu(du) \quad (\gamma \text{ est un compact invariant}) \\
&= \int_F P(1_\gamma f)(u)\, \nu(du) \quad (\text{F est réunion de compacts invariants}) \\
&= \int_E P(1_\gamma f)(u)\, \nu(du) \quad (\nu \text{ est portée par} F) \\
&= \int_\gamma f(u)\, \nu(du).
\end{aligned}$$

ce qui montre que la restriction de ν à chaque classe ergodique est encore une mesure invariante par P.

• Représentation intégrale des fonctions harmoniques

Nous appelons ρ le noyau de E dans $\mathcal{B}$ image par η du noyau N de E dans Ω défini par $N(x, A) = \mathbb{P}_x[A]$. Par définition, $\rho(x, B)$ est la probabilité, partant de x, que la chaîne ait toutes ses valeurs d'adhérence dans B.

Théorème 6.2.8 *Le noyau ρ vérifie les propriétés suivantes :*
i) $\forall\, x \in E,\ \int_E \rho(u, \cdot)\, P(x, du) = \rho(x, \cdot)$;
ii) $\forall\, \gamma \in \mathcal{G}, \forall\, x \in \gamma,\ \rho(x, \cdot)$ *est la mesure de Dirac en* γ.

Toute fonction h de $\mathcal{H}_c$ s'écrit

$$h(x) = \int_{\mathcal{G}} \hat{h}(\gamma)\, \rho(x, d\gamma).$$

L'application $h \to \hat{h}$ est une isométrie d'algèbres de Banach de $\mathrm{Alg}(\mathcal{H}_c)$ sur l'espace $C(\mathcal{G})$ des fonctions continues sur $\mathcal{G}$.

Preuve : Pour toute fonction continue $\hat{f}$ sur $\mathcal{G}$, nous avons :

$$\begin{aligned}
\int_{\mathcal{G}} \hat{f}(\gamma)\, \rho(x, d\gamma) &= \mathbb{E}_x[\hat{f} \circ \eta] = \mathbb{E}_x[\hat{f} \circ \eta \circ \theta] \quad (\theta\text{-invariance de } \eta) \\
&= \mathbb{E}_x[\mathbb{E}_x[\hat{f} \circ \eta \circ \theta \mid \mathcal{F}_1]] = \mathbb{E}_x[\mathbb{E}_{X_1}[\hat{f} \circ \eta]] \quad (\text{propriété de Markov}) \\
&= \int_E \mathbb{E}_u[\hat{f} \circ \eta]\, P(x, du) = \int_E \int_{\mathcal{G}} \hat{f}(\gamma)\, \rho(u, d\gamma) P(x, du);
\end{aligned}$$

d'où la première propriété du noyau ρ.

Soient $\gamma \in \mathcal{G}$ et $x \in \gamma$. Comme γ est un compact invariant, pour tout entier $n \geq 0$, nous avons $\mathbb{P}_x[X_n \in \gamma] = 1$. Par suite $\mathbb{P}_x - p.s.,\ \eta(\cdot) = \gamma$ et $\rho(x, \cdot)$ est la mesure de Dirac en γ.

La relation

$$h(x) = \mathbb{E}_x[\lim_{n \to +\infty} h(X_n)] = \mathbb{E}_x[\hat{h} \circ \eta] = \int_{\mathcal{G}} \hat{h}(\gamma)\, \rho(x, d\gamma).$$

Cette formule montre que $\|h\|_\infty \leq \|\hat{h}\|_\infty$. Comme l'inégalité opposée est évidente, l'application $h \to \hat{h}$ est bien une isométrie.

Pour toutes fonctions h_1 et h_2 de $\mathcal{H}_c$, nous avons $\widehat{h_1 * h_2} = \hat{h_1}\hat{h_2}$. D'après la construction même de l'espace $\mathcal{G}$, l'ensemble des fonctions continues $\{\hat{h} : h \in \mathcal{H}_c\}$ sépare les points de $\mathcal{G}$ et contient les fonctions constantes. D'après le théorème de Stone-Weierstrass, l'algèbre engendrée par ces fonctions est dense dans $C(\mathcal{G})$. On en déduit que l'application $h \to \hat{h}$ de $\mathrm{Alg}(\mathcal{H}_c)$ dans $C(\mathcal{G})$ est surjective.

□

Remarque Nous venons de montrer que l'isométrie surjective définie en (6.2.8) envoie la sous-algèbre $\mathrm{Alg}(\mathcal{H}_c)$ de $\mathcal{H}$ sur la sous-algèbre formée par les classes des variables aléatoires invariantes de la forme $f \circ \eta$, pour $f \in \mathcal{C}(\mathcal{B})$.

L'espace $\mathcal{B}$ apparaît ainsi comme une frontière de E et la formule

$$h(x) = \int_{\mathcal{G}} \hat{h}(\gamma)\, \rho(x, d\gamma)$$

comme une formule de Poisson.

Corollaire 6.2.9 *Pour tout $\gamma \in \mathcal{G}$, la fonction*

$$\mu_x(\{\gamma\}) = \mathbb{P}_x[\lim_{n \to +\infty} d(X_n, \gamma) = 0],$$

appartient à $\mathrm{Alg}(\mathcal{H}_c)$ et est semi-continue supérieurement. Cette fonction est continue si, et seulement si, γ est un point isolé de $\mathcal{B}$.

En particulier, lorsque $\mathcal{B}$ est fini (i.e. l'algèbre $\mathrm{Alg}(\mathcal{H}_c)$ est de dimension finie) toutes les fonctions $\{x \to \mu_x(\{\gamma\}) : \gamma \in \mathcal{G}\}$ sont continues, $\mathrm{Alg}(\mathcal{H}_c) = \mathcal{H}_c$ et ces fonctions constituent alors une base de $\mathcal{H}_c$.

Preuve : Soit $\gamma \in \mathcal{G}$. Considérons une suite de fonctions continues sur $\mathcal{G}$ qui décroît vers l'indicatrice du singleton $\{\gamma\}$. D'après la propriété d'isométrie du théorème 6.2.7, la fonction P-harmonique $x \to \mu_x(\{\gamma\})$ est une limite décroissante de fonctions P-harmoniques continues et est donc semi-continue supérieurement. De plus cette fonction est continue si, et seulement si, la fonction indicatrice du singleton $\{\gamma\}$ est continue ; ce qui équivaut à dire que γ est un point isolé de $\mathcal{G}$. Le résultat est alors clair. □

En général, l'égalité $\mathrm{Alg}(\mathcal{H}_c) = \mathcal{H}_c$ n'est pas assurée, mais est obtenue sous une hypothèse plus forte que nous allons introduire.

Définition 6.2.10 On dit que P vérifie la condition d'**équicontinuité en moyenne**, si, pour toute fonction continue f sur E, la famille de fonctions $\{\frac{1}{n}\sum_{k=0}^{n-1} P^k f : n \geq 1\}$ est équicontinue et donc relativement compacte dans $C(E)$.

Proposition 6.2.11 *Si P vérifie la condition d'équicontinuité en moyenne, alors l'algèbre $\mathrm{Alg}(\mathcal{H}_c)$ est égale à $\mathcal{H}_c$.*

Preuve : Soient h_1 et h_2 deux fonctions de $\mathcal{H}_c$. De l'hypothèse faite et de la relation

$$h_1 * h_2 = \lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k(h_1 h_2),$$

il résulte que la fonction $h_1 * h_2$ est continue. $\square$

6.3 Mesures invariantes et convergence en moyenne de Césaro

Nous désignons par $\mathcal{M}(P)$ l'ensemble des mesures de probabilité sur les boréliens de E qui sont invariantes par P. Nous nous proposons dans cette section de donner, sous l'hypothèse d'équicontinuité en moyenne, une description de $\mathcal{M}(P)$.

Pour toute fonction continue f sur E et tout entier $n \geq 1$, nous désignons par $M_n(f)$ la fonction continue sur E définie par $M_n(f) = \frac{1}{n}\sum_{k=0}^{n-1} P^k f$.

Proposition 6.3.1 *Lorsque P vérifie la condition d'équicontinuité en moyenne,*
i) chaque classe ergodique γ porte une unique mesure ν_γ de $\mathcal{M}(P)$;
ii) tout élément λ de $\mathcal{M}(P)$ s'écrit $\int_{\mathcal{G}} \nu_\gamma \, \mu_\lambda(d\gamma)$ avec $\mu_\lambda(\cdot) = \int_E \rho(x,\cdot)\lambda(dx)$;
iii) pour toute fonction f de $C(E)$, la suite de fonctions $(M_n(f))_{n\geq 1}$ converge uniformément sur E vers $\int_{\mathcal{G}} \nu_\gamma(f)\, \rho(\cdot, d\gamma)$.

Preuve : Soit $(\varphi(n))_{n\geq 1}$ une suite strictement croissante d'entiers telle que, pour tout $f \in C(E)$, la suite de fonctions $(M_{\varphi(n)}(f))_{n\geq 1}$ converge uniformément sur E vers une fonction notée h_f. [L'existence de telles suites est assurée par le procédé diagonal, compte-tenu de la séparabilité de $C(E)$.] Chaque fonction h_f appartient à $\mathcal{H}_c$ et s'écrit donc (théorème 6.2.7 : $h_f(x) = \int_{\mathcal{G}} \widehat{h_f}(\gamma)\, \rho(x, d\gamma)$).

Il est clair que, pour tout $\gamma \in \mathcal{G}$, l'application $f \mapsto \widehat{h_f}(\gamma)$ définit une mesure de probabilité ν_γ sur E, invariante par P. (Rappelons que $\widehat{h_f}(\gamma)$ est la valeur prise par h_f sur γ). Si λ est une mesure quelconque de $\mathcal{M}(P)$, nous avons, pour tout entier $n \geq 1$,

$$\begin{aligned}\int_E f(u)\,\lambda(du) &= \int_E M_{\varphi(n)}(f)(u)\,\lambda(du)\\ &= \int_E \int_{\mathcal{G}} \nu_\gamma(f)\rho(x,d\gamma)\,\lambda(dx).\end{aligned}$$

En restriction à une classe ergodique γ, la suite de fonctions $(M_{\varphi(n)}(f))_{n\geq 1}$ converge uniformément vers $\widehat{h_f}(\gamma) = \nu_\gamma(f)$. [Noter que, pour tout $x \in \gamma$, $\rho(x,\cdot)$ est la mesure de Dirac en γ]. Il s'ensuit que ν_γ est l'unique mesure de $\mathcal{M}(P)$ portée par γ. On en déduit alors que les fonctions limites $h_f, f \in C(E)$ ne dépendent pas de la suite d'entiers $(\varphi(n))_{n\geq 1}$ choisie ; d'où la convergence des moyennes. $\square$

Proposition 6.3.2 *Supposons que P vérifie la condition d'équicontinuité en moyenne. Alors, il existe un sous-ensemble mesurable θ invariant $\widetilde{\Omega}_0$ de Ω_0 de $\mathbb{P}_x$-mesure 1, pour tout $x \in E$, tel que, pour tout $\omega \in \widetilde{\Omega}_0$ et toute fonction f de $C(E)$, la suite de variables aléatoires $(\frac{1}{n}\sum_{k=0}^{n-1} f(X_k(\omega)))_{n\geq 1}$ converge vers $\nu_{\eta(\omega)}(f)$.*

Preuve : Soient $f \in C(E)$ et $x \in E$. On voit aisément que le processus $\left(\sum_{k=1}^n \frac{f(X_k)-Pf(X_{k-1})}{k}\right)_{n\geq 1}$ est une martingale bornée dans $L^2(\Omega, \mathcal{F}, \mathbb{P}_x)$, relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Ce processus converge donc $\mathbb{P}_x$-p.s. et il s'ensuit (lemme de Kronecker) que le processus

$$\Big(S_n(f) = \frac{1}{n}\sum_{k=1}^{n}(f(X_k) - Pf(X_{k-1}))\Big)_{n\geq 1}$$

converge $\mathbb{P}_x$-p.s. vers zéro.

Considérons le sous-ensemble Ω_0 définie en (2.8) et appelons $\widetilde{\Omega_0}$ l'ensemble des éléments ω de Ω_0 pour lesquels la suite $(S_n(f)(\omega))_{n\geq 1}$ converge vers zéro, pour toutes les fonctions f de $C(E)$. L'espace $C(E)$ étant séparable, $\widetilde{\Omega_0}$ est un borélien de $\mathcal{F}$ de $\mathbb{P}_x$-mesure 1, pour tout $x \in E$.

Soient $\omega \in \widetilde{\Omega_0}$ et λ_ω une valeur d'adhérence étroite de la suite de mesures de probabilité $\left(\frac{1}{n}\sum_{k=1}^n \varepsilon_{X_k(\omega)}\right)_{n\geq 1}$.

D'après ce qui précède, λ_ω est invariante par P. D'autre part, d'après le lemme 6.2.5, $\lim_{n\to+\infty} d(X_n(\omega), \eta(\omega)) = 0$; la mesure λ_ω est nécessairement portée par la classe ergodique $\eta(\omega)$.

On en déduit que $\lambda_\omega = \nu_{\eta(\omega)}$. □

Définition 6.3.3 Un compact M de $\mathbb{R}^d$ invariant est dit **minimal**, s'il ne contient pas de sous-ensemble fermé invariant non vide autre que lui-même.

Dans le cas général, l'existence d'un compact invariant minimal dans tout compact invariant s'établit à l'aide du lemme de Zorn. Sous l'hypothèse d'équicontinuité en moyenne, il résulte de la proposition 6.3.2 que chaque classe ergodique γ contient un unique compact invariant minimal, noté γ_{min}, qui est le support de ν_γ.

D'après la proposition 6.3.2 appliquée pour ω fixé à la fonction $d(\cdot, (\eta(\omega))_{min})$, pour tout $\omega \in \widetilde{\Omega}_0$, on a la convergence vers 0, en moyenne de Césaro, de la suite $d(X_n(\omega), (\eta(\omega))_{min}$, soit encore la convergence vers 0 de cette suite sur un ensemble de densité 1.

Une question naturelle est de savoir si l'on a $\lim_n d(X_n(\omega), (\eta(\omega))_{min} = 0$, ce qui permettrait dans l'étude précédente de remplacer chaque classe ergodique γ par le compact invariant minimal qu'elle contient.

Exemples : Cas régulier.

On considère la transformation τ définie sur le cercle E identifié à $[0,1[$ par $\tau : x \to px$ mod 1. On note $S = \{s_i, i = 0, \ldots, p-1\}$, $s_i x = \dfrac{x+i}{p}$, la famille des inverses de τ et on considère la probabilité de transition P_u définie par :

$$P_u f(x) = \sum_{i=0}^{p-1} u(s_i x) f(s_i x),$$

où u est une fonction continue positive sur E telle que

$$\sum_{s\in S} u(sx) = 1, \forall x \in E.$$

Etant donnée une fonction ϕ continue sur $[0,1]$, à valeurs positives, croissante et nulle en zéro, nous notons L_ϕ le sous-espace de E formé des fonctions f telles que :

$$m_\phi(f) = \sup\{\frac{|f(x)-f(y)|}{\phi(|x-y|)}, x, y \in [0,1], x \neq y\} < \infty.$$

Nous munissons L_ϕ de la norme $\|\|_\phi$ définie par :

$$\|f\|_\phi = \|f\|_\infty + m_\phi(f).$$

L'espace L_ϕ est un espace de Banach et, d'après le théorème d'Ascoli, tout sous-ensemble borné de L_ϕ est relativement compact dans E.

Définition 6.3.4 Nous dirons que u vérifie l'hypothèse de régularité (R_ϕ) si $u \in L_\phi$, pour une fonction ϕ vérifiant

$$\sum_{k\geq 0} \phi(p^{-k}) < +\infty. \tag{6.2}$$

On notera que l'hypothèse (R_ϕ) est satisfaite, avec $\phi(x) = x^\alpha$, pour un $\alpha > 0$, si u est **höldérienne**, mais qu'elle est plus générale. On peut par exemple prendre ϕ de la forme

$$\phi(x) = (1 + |log(x)|)^{-(1+\varepsilon)}, \text{ pour un } \varepsilon > 0.$$

• **Continuité en variation de la famille** $(\mathbb{P}_x, x \in [0,1])$

Nous allons montrer que lorsque u vérifie l'hypothèse (R_ϕ), l'opérateur P_u la propriété d'équicontinuité suivante est vérifiée :
$\forall f \in \mathcal{C}(E)$, la suite $(P_n f)$ est équicontinue.

Pour $x, y \in E$ et tout entier $n \geq 1$, posons

$$\delta_n(x,y) = \sum_{(s_1,\dots s_n)\in S^n} |u(s_n \dots s_1 x)\dots u(s_1 x) - u(s_n \dots s_1 x)\dots u(s_1 y)|.$$

Nous avons :

$$\delta_n(x,y) \leq p m_\phi(u) \sum_{k=1}^{n} \phi(|x-y|p^{-k}).$$

Il en résulte que, pour toute fonction $f \in \mathcal{C}(E)$, on a :

$$|P^n f(x) - P^n f(y)| \leq \delta_n(x,y) + V(f, p^{-n}|x-y|);$$

où, pour tout $\alpha > 0$, on pose $V(f,\alpha) = \sup_{d(x,y)\leq\alpha} |f(x) - f(y)|$.

L'hypothèse (R_ϕ) permet de montrer une propriété de continuité en variation pour la famille de probabilités $(\mathbb{P}_x, x \in [0,1])$ introduite dans la section précédente.

Considérons la chaîne des sauts définie dans le chapitre 3 associée à P_u. Nous avons :

Proposition 6.3.5 *Si la fonction u vérifie la condition de régularité (R_ϕ), il existe une constante D, $0 \leq D < \infty$, telle que, pour toute variable aléatoire bornée U sur Ω,*

$$|\mathbb{E}_x(U) - \mathbb{E}_y(U)| \leq D \sum_{k\geq 1} \phi(|x-y)|p^{-k})\ \|U\|_\infty, \text{ pour } x, y \in [0,1].$$

Preuve : Si U ne dépend que d'un nombre fini n de coordonnées, on a :

$$\begin{aligned}
|\mathbb{E}_x[U(Y_1,...,Y_n)] - \mathbb{E}_y[U(Y_1,...,Y_n)]| &= |\sum_{s\in S^n} U(s_1,\dots,s_n)u(s_n\dots s_1 x)\dots u(s_1 x) \\
&\qquad - \sum_{s\in S^n} U(s_1,\dots,s_n)u(s_n\dots s_1 y)\dots u(s_1 y)| \\
&\leq \delta_n(x,y)\|U\|_\infty.
\end{aligned}$$

Dans le cas général, posons $\mathbb{Q} = \frac{1}{2}(\mathbb{P}_x + \mathbb{P}_y)$, et soit U_n l'espérance conditionnelle $\mathbb{E}_{\mathbb{Q}}[U|\mathcal{F}_n]$. Les variables aléatoires $U_n, n \geq 1$, sont bornées par $\|U\|_\infty$ et la suite converge $\mathbb{Q}$-p.s., et donc $\mathbb{P}_x$ et $\mathbb{P}_y$-p.s., vers U. Il résulte de la majoration précédente :

$$|\mathbb{E}_x(U_n) - \mathbb{E}_y(U_n)| \leq \|U\|_\infty \delta_n(x,y) \leq D \sum_{k\geq 1} \phi(|x-y|p^{-k})\|U\|_\infty,$$

et l'inégalité cherchée résulte du théorème de convergence dominée de Lebesgue. □

Ce résultat montre que, pour toute variable aléatoire bornée U définie sur Ω θ-invariante, la fonction $x \to \mathbb{E}_x(U)$ est une fonction harmonique continue.

Si C est un compact invariant, nous avons

$$\mathbb{P}_x[\lim_n d(Z_n(x,\cdot),C)=0] = \mathbb{P}_x[\lim_n d(Z_n(0,\cdot),C)=0],$$

ce qui implique que la fonction $x \to \mathbb{P}_x[\lim_n d(Z_n(x,\cdot),C)=0]$ est dans $\mathcal{H}_c$.

Dans la théorie précédente, on peut donc remplacer les classes ergodiques par les compacts invariants minimaux.

Exemple 6.3.6

Nous reprenons le même exemple avec une fonction u moins régulière. Soient $p=2$ et u définie par $u(x) = \log(2x)/\log x$ pour $x \in]0, \frac{1}{2}]$, $u(0)=1$ et $u(x+\frac{1}{2}) = 1-u(x)$, pour $x \in]0, \frac{1}{2}]$.

Nous allons montrer que, dans cet exemple, la mesure de Dirac en 0 est l'unique mesure de probabilité invariante ; ce qui implique l'équicontinuité en moyenne de P. L'unique classe ergodique est l'espace E tout entier. Le singleton $\{0\}$ est l'unique compact invariant minimal. Cependant la convergence de la chaîne vers 0 n'a lieu qu'en moyenne de Césaro.

Soit μ une mesure invariante extrémale pour la chaîne canonique associée à P.

Pour $\varepsilon \in]0, \frac{1}{2}[$, posons $S_\varepsilon = \inf\{n \geq 1 : X_n \in [\varepsilon, 1[\}$. Nous avons :

$$\mathbb{P}_x[S_\varepsilon \geq n] = u(\frac{x}{2}) \ldots u(\frac{x}{2^{n-1}}) = \frac{\log x}{\log \frac{x}{2^{n-1}}}.$$

On a donc, pour tout $x \in E$,

$$\mathbb{E}_x[S_\varepsilon] = \sum_{n\geq 1} \mathbb{P}_x[S_\varepsilon \geq n] = \sum_{n \geq 0} \frac{1}{1+n\frac{\log 2}{\log 1/x}} = +\infty.$$

Le système $(\Omega, \mathcal{B}(\Omega), \mathbb{P}_\mu, \theta)$ est ergodique d'après le lemme ? ?, chapitre I. Du lemme ? ?, chapitre I, appliqué à ce système et au borélien $A = [\varepsilon, 1[$, il résulte que nécessairement $\mu(]\varepsilon, 1]) = 0$.

La mesure μ est la mesure de Dirac en 0 et l'espace $\mathcal{H}_c$ est donc réduit aux constantes ce qui implique que E est l'unique classe ergodique. Cependant, on a, pour $x \in]0,1[$, $\mathbb{P}_x[S_\varepsilon < \infty] = 1$ et par suite (propriété de Markov forte), partant de $x \in]0,1[$, la chaîne visite une infinité de fois $]\varepsilon, 1[$.

Cet exemple montre que l'on ne peut pas, sans hypothèse supplémentaire, remplacer les classes ergodiques "grossières" définies par $\mathcal{H}_c$ par les ensembles minimaux.

Dans le cas d'équicontinuité en moyenne, il existe une application $\widetilde{\eta} : \Omega \to \widetilde{\mathcal{G}}$, où $\widetilde{\mathcal{G}}$ est l'ensemble des compacts invariants minimaux telle que, pour tout $\omega \in \widetilde{\Omega}_0$ la suite $d(X_n(\omega), \widetilde{\eta}(\omega))$ converge vers 0 en moyenne de Césaro.

6.4 Fonctions harmoniques pour la chaîne espace-temps

• Définitions et notations

Nous appelons Q la probabilité de transition sur l'espace produit $E \times \mathbb{Z}$ définie par :

$$QF(x,n) = \int_E F(y,n+1)\, P(x,dy), \quad (x,n) \in E \times \mathbb{Z}.$$

Nous désignons par $\mathcal{K}$ l'espace vectoriel des fonctions boréliennes bornées Q-harmoniques sur $E\times\mathbb{Z}$. Tout élément H de $\mathcal{K}$ est une suite bornée $(h_n = H(\cdot,n))_{n\in\mathbb{Z}}$ de fonctions boréliennes sur E telle que $\forall\, n \in \mathbb{Z}$, $Ph_n = h_{n-1}$.

Nous notons $\mathcal{K}_c$ le sous-espace vectoriel de $\mathcal{K}$ constitué des fonctions H telles que la famille de fonctions $\{H(\cdot,n) : n \in \mathbb{Z}\}$ soit équicontinue (et donc relativement compacte dans C(E)).

Comme en 3.10, on définit le produit de deux fonctions de $\mathcal{K}$ par :

$$H_1 * H_2 = \lim_{n\to+\infty} Q^n(H_1 H_2).$$

Muni de ce produit et de la norme uniforme $\| \ \|_\infty$, définie par

$$\|H\|_\infty = \sup_{n\in\mathbb{Z}} \|H(\cdot,n)\|_\infty,$$

$\mathcal{K}$ est une algèbre de Banach. A tout élément h de $\mathcal{H}$ est associé un élément $H = (h_k)_{k\in\mathbb{Z}}$ de $\mathcal{K}$ défini par $h_k = h$, pour tout entier relatif k. Cette correspondance permet d'identifier l'algèbre $\mathcal{H}$ à une sous-algèbre de $\mathcal{K}$.

• Fonctions Q-harmoniques et variables aléatoires asymptotiques

Pour tout entier $n \geq 0$, nous notons $\mathcal{A}_n$ la sous-tribu de $\mathcal{F}$ engendrée par les variables aléatoires $\{X_k : k \geq n\}$. Nous appelons $\mathcal{A}$ la **tribu asymptotique** ; c'est-à-dire la tribu intersection des tribus $\mathcal{A}_n$. La tribu $\mathcal{I}$ ses boréliens invariants par θ est contenue dans la tribu asymptotique. Une variable aléatoire U est mesurable par rapport à la tribu $\mathcal{A}$ si, et seulement si, pour tout entier naturel $n \geq 1$, il existe une variable aléatoire U_n telle que $U = U_n \circ \theta^n$; une telle variable sera dite **asymptotique**. Soit n un entier supérieur où égal à 1. Pour tous éléments $x_1, \ldots, x_n$ de E et tout $\omega \in \Omega$, il vient :

$$\forall\, k \geq 1, U(x_1, \ldots, x_n, \omega) = U_n(\omega) = U_{n+k} \circ \theta^k(\omega).$$

Il s'ensuit que les variables aléatoires $\{U_n : n \geq 1\}$ sont déterminées de façon unique et sont asymptotiques. Nous écrirons donc $\forall\, n \geq 1$, $U_n = U \circ \theta^{-n}$.

Par une preuve analogue à celle de la proposition 6.2.1, on montre :

Proposition 6.4.1 *([Rev75], ch. 6, prop. 2.3) Pour tout $H \in \mathcal{K}$ et tout $(x,k) \in E\times\mathbb{Z}$, le processus $(H(X_n, n{+}k))_{n\geq 0}$, défini sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P}_x)$, est une martingale bornée relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$. Ce processus converge $\mathbb{P}_x$-p.s. vers une variable aléatoire asymptotique $U\circ\theta^{-k}$, vérifiant, pour tout entier $n \geq 0$,*

$$H(x,k) = \mathbb{E}_x[H(X_n, n+k)] = \mathbb{E}_x[U \circ \theta^{-k}]$$

et

$$H(X_n, n+k) = \mathbb{E}_x[U \circ \theta^{-k} \mid \mathcal{F}_n] \quad \mathbb{P}_x\text{-p.s.}$$

Réciproquement, si U est une variable aléatoire asymptotique bornée, alors la fonction H, définie par

$$H(x,k) = \mathbb{E}_x[U \circ \theta^{-k}] \ \ ((x,k) \in E \times \mathbb{Z}),$$

appartient à $\mathcal{K}$ et la martingale $(H(X_n, k+n))_{n\geq 0}$ converge, $\mathbb{P}_x$-p.s. $\forall \in E$, vers $U \circ \theta^{-k}$.

Deux variables aléatoires asymptotiques, U et U', sont dites équivalentes si $\forall\,(x,k) \in E\times\mathbb{Z}$, $\mathbb{P}_x[U\circ\theta^{-k} \neq U'\circ\theta^{-k}] = 0$. Nous appelons $\mathcal{W}$ l'algèbre des classes d'équivalence des variables aléatoires asymptotiques.

Corollaire 6.4.2 *L'application qui à la classe U de $\mathcal{W}$ associe la fonction Q-harmonique $H(x,k) = \mathbb{E}_x[U\circ\theta^{-k}]$ est un isomorphisme d'algèbre de Banach.*

Lemme 6.4.3 *Soit H une fonction de $\mathcal{K}$. Pour tout $r \geq 1$, tout $k \in \mathbb{Z}$ et tout $x \in E$, nous avons*

$$\sum_{n\geq 0} \mathbb{E}_x\Big[\int_E [H(y,r+n+k) - H(X_n,n+k)]^2\, P^r(X_n,dy)\big] \;<\; +\infty.$$

Preuve : La démonstration est identique à celle du lemme (2.7). □

Corollaire 6.4.4 *Pour tout fonction H de $\mathcal{K}$, tout entier $k \in \mathbb{Z}$ et tout entier $r \geq 1$, la suite de variables aléatoires $\left(\int_E [H(y,r+n+k) - H(X_n,n+k)]^2\, P^r(X_n,dy)\right)_{n\geq 0}$ converge vers zéro, $\mathbb{P}_x$-p.s. $\forall\, x \in E$.*

• Lemmes fondamentaux

Dorénavant on suppose que P vérifie l'hypothèse d'équicontinuité suivante notée (H) :

pour toute fonction continue f, la famille de fonctions $\{P^n f \;:\; n \geq 0\}$ est équicontinue et donc relativement compacte dans $C(E)$.

Nous allons établir une représentation intégrale des fonctions de $\mathcal{K}_c$ analogue à celle obtenue pour les éléments de $\mathcal{H}_c$.

Lemme 6.4.5 *De toute suite d'entiers naturels, $(\varphi(n))_{n\geq 0}$, on peut extraire une sous-suite $(\varphi\circ\psi(n))_{n\geq 0}$, telle que : pour tout $f \in C(E)$ et tout $k \in \mathbb{Z}$, la suite de fonctions $(P^{\varphi\circ\psi(n)-k} f)_{n\geq 0}$ converge uniformément sur E. La fonction H définie par,*

$$H(x,k) = \lim_{n\to+\infty} P^{\varphi\circ\psi(n)-k} f(x),\;\; (x,k) \in E\times\mathbb{Z}.$$

appartient alors à $\mathcal{K}_c$.

Preuve : Compte-tenu de la séparabilité de $C(E)$ et de la propriété de contraction de P, la première assertion découle du procédé diagonal. La seconde assertion est évidente. □

Lemme 6.4.6 *Soit $(\varphi(n))_{n\geq 0}$ une suite strictement croissante d'entiers naturels telle que, pour tout $f \in C(E)$ et tout $k \in \mathbb{Z}$, la suite de fonctions $(P^{\varphi(n)-k} f)_{n\geq 0}$ converge uniformément sur E.*

Soit $H = (h_k)_{k\in\mathbb{Z}}$ une fonction de $\mathcal{K}_c$. Soit (procédé diagonal) $(\psi(n))_{n\geq 0}$ une suite d'entiers telle que, pour tout $k \in \mathbb{Z}$, la suite de fonctions $(h_{\varphi\circ\psi(n)+k})_{n\geq 0}$, converge uniformément vers une fonction notée l_k. Alors $L = (l_k)_{k\in\mathbb{Z}}$ est un élément de $\mathcal{K}_c$ tel que, pour tout entier k de $\mathbb{Z}$, la suite de fonctions $(l_{k-\varphi(n)})_{n\geq 0}$, converge uniformément vers h_k.

L'espace $\mathcal{K}_c$ est une sous-algèbre séparable de $\mathcal{K}$.

Preuve : Il est clair que L appartient à $\mathcal{K}_c$.

Nous avons :

$$\begin{aligned} h_k &= P^{\varphi(n)} h_{\varphi(n)+k} = P^{\varphi\circ\psi(n)} h_{\varphi\circ\psi(n)+k} \\ &= \lim_{n\to+\infty} P^{\varphi\circ\psi(n)} l_k = \lim_{n\to+\infty} P^{\varphi(n)} l_k \\ &= \lim_{n\to+\infty} l_{k-\varphi(n)}, \end{aligned}$$

les convergences étant uniformes sur E. D'où, la première assertion du lemme.

Ce qui précède montre que pour tout élément H de $\mathcal{K}_c$ et tout entier k de $\mathbb{Z}$, la fonction $h_k = H(\cdot, k)$ s'obtient comme limite uniforme d'une suite de fonctions du type $(P^{\varphi(n)-k} f)_{n\geq 0}$, avec $f \in C(E)$ (Il suffit de prendre $f = l_0$). La séparabilité de $\mathcal{K}_c$ est alors claire.

Soient $H_1 = (h_{1,k})_{k\in\mathbb{Z}}$, $H_2 = (h_{2,k})_{k\in\mathbb{Z}}$ deux éléments de $\mathcal{K}_c$ et k un entier relatif. Nous avons

$$\begin{aligned} H_1 * H_2(\cdot, k) &= \lim_{n\to+\infty} Q^{n-k}(H_1 H_2)(\cdot, k) \\ &= \lim_{n\to+\infty} P^{n-k}(h_{1,n} h_{2,n}). \end{aligned}$$

Soit $(\psi(n))_{n\geq 0}$ une suite d'entiers telle que, pour $i \in \{1, 2\}$, la suite de fonctions $(h_{i,\varphi\circ\psi(n)})_{n\geq 0}$, converge uniformément sur E vers une fonction notée f_i. Il vient alors

$$H_1 * H_2(\cdot, k) = \lim_{n\to+\infty} P^{\varphi(n)-k}(f_1 f_2);$$

ce qui montre, puisque P vérifie l'hypothèse (H), que $H_1 * H_2$ est encore un élément de $\mathcal{K}_c$. $\square$

- **Classes cycliques**

Notation : Soit $H \in \mathcal{K}_c$. Nous appelons Ω_H le sous-ensemble de Ω, $\mathcal{F}$-mesurable et invariant par θ, constitué des éléments ω tels que :
i) pour tout entier $k \in \mathbb{Z}$, la suite $\Big(H(X_n(\omega), k+n)\Big)_{n\geq 0}$ converge
ii) pour tout entiers $r \geq 1$ et $k \in \mathbb{Z}$, la suite

$$\Big(\int_E [H(y, r+n+k) - H(X_n, n+k)]^2 \, P^r(X_n, dy)\Big)_{n\geq 0}$$

converge vers zéro.

D'après ce qui précède, nous savons que $\mathbb{P}_x[\Omega_H] = 1, \forall x \in E$. Comme $\mathcal{K}_c$ est séparable, en posant $\Omega_1 = \cap_{H\in\mathcal{K}_c}\Omega_H$, nous obtenons un sous-ensemble de Ω, $\mathcal{F}$-mesurable, invariant par θ, de $\mathbb{P}_x$-mesure 1 pour tout $x \in E$.

Lemme 6.4.7 *Soient ω un élément de Ω_1 et u une valeur d'adhérence dans E de la suite $(X_n(\omega))_{n\geq 0}$. Pour tout élément H de $\mathcal{K}_c$ et tout entier k de $\mathbb{Z}$, nous avons :*

$$\forall\, r \geq 1,\ \forall\, x \in \operatorname{supp}(P^r(u, \cdot)),\quad H(u, k) = H(x, k+r).$$

Preuve : Soit $(\varphi(n))_{n\geq 0}$ une suite d'entiers naturels telle que $X_{\varphi(n)}(\omega) \to u$. Soit ψ une application strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ elle que, pour tout $f \in C(E)$ et tout $k \in \mathbb{Z}$, la suite de fonctions $(P^{\varphi\circ\psi(n)-k} f)_{n\geq 0}$ converge uniformément sur E (lemme 6.4.5).

Soit $H = (h_k)_{k\in\mathbb{Z}}$ une fonction de $\mathcal{K}_c$. D'après le lemme 6.4.6, il existe une application strictement croissante ξ de $\mathbb{N}$ dans $\mathbb{N}$ et une fonction $L = (l_k)_{k\in\mathbb{Z}}$ de $\mathcal{K}_c$ telles que, pour tout $k \in \mathbb{Z}$:

a) $l_k = \lim_{n\to+\infty} h_{\varphi\circ\psi\circ\xi(n)+k}$;
b) $h_k = \lim_{n\to+\infty} l_{k-\varphi\circ\psi(n)}$.

Des propriétés *ii*) de la définition de Ω_H et *a*) ci-dessus, il résulte que, pour tous entiers $r \geq 1$ et $k \in \mathbb{Z}$,

$$\lim_{n\to+\infty} \int_E \left[L(x, r+k) - L(X_{\varphi\circ\psi\circ\xi(n)}, k)\right]^2 P^r(X_{\varphi\circ\psi\circ\xi(n)}, dx) \quad = \quad 0;$$

et par suite, puisque P respecte $C(E)$,

$$\forall\, x \in \operatorname{supp} P^r(u,\cdot),\ L(x, r+k) = L(u,k).$$

De la propriété *b*) ci-dessus, il vient alors :

$$\forall\, x \in \operatorname{supp} P^r(u,\cdot),\ H(x, r+k) = H(u,k).$$

□

Ceci nous amène à considérer le sous-ensemble compact, non vide, de E défini par :

$$D = \{u \in E : \forall H \in \mathcal{K}_c, \forall k \in \mathbb{Z}, \forall r \geq 1, \forall x \in \operatorname{supp} P^r(u,\cdot),\ H(x,r+k) = H(u,k)\}.$$

Comme $\mathcal{H}_c$ s'identifie à une sous-algèbre de $\mathcal{K}_c$, le compact D est contenu dans le compact invariant F de (6.2.4). Du lemme 6.1.2, il résulte que D aussi est invariant.

D'après le lemme 6.4.7, D contient toutes les valeurs d'adhérence des suites $\{(X_n(\omega))_{n\geq 0} : \omega \in \Omega_1\}$. Tout ce qui a été fait dans la section 2, peut alors se faire en remplaçant F par D. Dans la suite nous supposerons que cette substitution a été faite et donc $D = F$.

Définition 6.4.8 Nous disons que deux éléments u et v de F sont **équivalents** si, pour tout $H \in \mathcal{K}_c$ et tout $k \in \mathbb{Z}$, $H(u,k) = H(v,k)$.

Nous définissons ainsi une relation d'équivalence sur F. Chaque classe d'équivalence fournit un sous-ensemble compact de F. Nous notons $\mathcal{C}$ l'ensemble de ces compacts. Comme $\mathcal{H}_c$ s'identifie une sous algèbre de $\mathcal{K}_c$, cette partition de F est obtenue par découpage de chaque classe ergodique γ de $\mathcal{B}$. Pour toute classe ergodique γ, nous notons $\mathcal{C}_\gamma$ l'ensemble des éléments de $\mathcal{C}$ qui constituent une partition de γ.

Pour tout $H \in \mathcal{K}_c$, et tout entier $k \in \mathbb{Z}$, nous notons $\hat{H}(\varsigma, k)$ la valeur de $H(\cdot, k)$ sur la classe ς de $\mathcal{C}$. Soit $(H_p)_{p\in\mathbb{Z}}$ une suite d'éléments, non nuls, de $\mathcal{K}_c$ dense dans $\mathcal{K}_c$. Nous munissons $\mathcal{C}$ de la distance δ définie par :

$$\delta(\varsigma_1, \varsigma_2) = \sum_{p\geq 0} \frac{1}{2^p} \frac{1}{\|H_p\|_\infty} \sup_{k\in\mathbb{Z}} |\widehat{H_p}(\varsigma_1, k) - \widehat{H_p}(\varsigma_2, k)|.$$

Nous avons alors :

Lemme 6.4.9 $(\mathcal{C}, \delta)$ *est un espace métrique compact. Pour toute classe* ς *de* $\mathcal{C}$*, les supports des mesures de probabilités* $\{P(u,\cdot) : u \in \varsigma\}$ *sont tous contenus dans une unique classe de* $\mathcal{C}$*, notée* $\tau \cdot \varsigma$*. L'opérateur* P *induit alors une isométrie, notée* τ*, sur l'espace* $(\mathcal{C}, \delta)$*. Pour tout* $H \in \mathcal{K}_c$*, tout* $\varsigma \in \mathcal{C}$ *et tout* $k \in \mathbb{Z}$*,* $\hat{H}(\varsigma, k) = \hat{H}(\tau^{-k} \cdot \varsigma, 0)$*.*

Preuve : Soient $(\varsigma_n)_{n\geq 0}$ une suite de classes de $\mathcal{C}$ et, pour tout entier naturel n, un élément u_n de ς_n. Par le procédé diagonal nous construisons une suite strictement croissante d'entiers $(\varphi(n))_{n\geq 0}$

telle que pour tous entiers $p \geq 0$ et $k \in \mathbb{Z}$, la suite $\Big(H_p(u_{\varphi(n)}, k)\Big)_{n\geq 0}$ converge. Quitte à prendre une sous-suite de $(\varphi(n))_{n\geq 0}$, on peut supposer que la suite $(u_{\varphi(n)})_{n\geq 0}$ converge vers un élément u de F. Si ς est la classe de $\mathcal{C}$ qui contient u, il est alors clair que : $\lim_{n\to+\infty} \delta(\varsigma_{\varphi(n)}, \varsigma) \;=\; 0$. D'où la compacité.

La deuxième et la dernière assertions résultent de la définition même du compact absorbant $D = F$.

Les relations
$$\forall\, k \in \mathbb{Z}, \forall\, \varsigma \in \mathcal{C}, \hat{H}(\tau\cdot\varsigma, k+1) = \hat{H}(\varsigma, k),$$
impliquent que, pour tous ς_1 et ς_2 de $\mathcal{C}$, $\delta(\tau\cdot\varsigma_1, \tau\cdot\varsigma_2) = \delta(\varsigma_1, \varsigma_2)$. Comme toute isométrie dans un espace métrique compact est nécessairement surjective et donc bijective, la démonstration du lemme est terminée.

□

L'ensemble des isométries de l'espace métrique compact $(\mathcal{C}, \delta)$ muni de la loi de composition des applications est un groupe topologique compact. Nous désignons par G le sous-groupe fermé du groupe des isométries, engendré par l'isométrie τ ; G est un groupe compact abélien. En outre comme tout semi-groupe compact d'un groupe topologique est un groupe, G coïncide avec le semi-groupe fermé engendré par l'isométrie τ. En particulier, il existe des suites d'entiers $(\varphi(n))_{n\geq 0}$ telles que la suite des isométries $(\tau^{\varphi(n)})_{n\geq 0}$ de G converge vers l'identité (ou élément neutre de G).

Lemme 6.4.10 *Il existe une application $\mathcal{F}$-mesurable ζ de Ω dans $(\mathcal{C}, \delta)$ telle que :*
i) Pour tout $\omega \in \Omega_1$ et toute suite d'entiers $(\varphi(n))_{n\geq 0}$ pour laquelle, la suite des isométries $(\tau^{\varphi(n)})_{n\geq 0}$ converge vers l'identité,
$$\lim_{n\to+\infty} \delta(X_{\varphi(n)}(\omega), \zeta(\omega)) \;=\; 0;$$

ii) Pour tout $\omega \in \Omega_1$, $\zeta \circ \theta(\omega) = \tau \circ \zeta(\omega)$.

Nous appelons **classe cyclique** chaque élément de $\mathcal{C}$.

<u>Preuve</u> du lemme Soient $(\varphi(n))_{n\geq 0}$ et $(\psi(n))_{n\geq 0}$ deux suites d'entiers naturels telle que les suites des isométries $(\tau^{\varphi(n)})_{n\geq 0}$ et $(\tau^{\psi(n)})_{n\geq 0}$ convergent vers l'identité. Soit $\omega \in \Omega_1$. Soient u_1 et u_2 des valeurs d'adhérence respectivement des suites $(X_{\varphi(n)}(\omega))_{n\geq 0}$ et $(X_{\psi(n)}(\omega))_{n\geq 0}$.

Nous appelons ς_1 et ς_2 les classes cycliques contenant respectivement u_1 et u_2.

Soit H une fonction de $\mathcal{K}_c$. Il existe des applications strictement croissantes ξ_1 et ξ_2 de $\mathbb{N}$ dans $\mathbb{N}$ telles que, tout entier $k \in \mathbb{Z}$,
$$\lim_{n\to+\infty} H(X_n(\omega), n+k) = \lim_{n\to+\infty} H(u_1, \varphi\circ\xi_1(n)+k) = \lim_{n\to+\infty} H(u_2, \psi\circ\xi_2(n)+k);$$
et par suite,
$$\lim_{n\to+\infty} H(X_n(\omega), n+k) = \hat{H}(\tau^{-k}\cdot\varsigma_1, 0) = \hat{H}(\tau^{-k}\cdot\varsigma_2, 0).$$
On en déduit que :
$$\forall\, g \in G,\ \forall\, k \in \mathbb{Z},\ \hat{H}(g\cdot\varsigma_1, k) = \hat{H}(g\cdot\varsigma_2, k);$$
et donc $\varsigma_1 = \varsigma_2$.

Nous venons de montrer qu'il existe une classe cyclique $\zeta(\omega)$ qui contient toutes les valeurs d'adhérence de la suite $(X_{\varphi(n)}(\omega))_{n\geq 0}$, pour toute suite d'entiers naturels $(\varphi(n))_{n\geq 0}$ pour laquelle la suite $(\tau^{\varphi(n)})_{n\geq 0}$ converge vers l'identité. Nous obtenons ainsi une application mesurable ζ de Ω_1 dans $\mathcal{C}$ vérifiant l'assertion i). [La mesurabilité de ζ se démontrant de la même façon que celle de η dans le lemme 6.2.6.] Nous étendons cette application à Ω en lui donnant une valeur arbitraire sur le complémentaire de Ω_1.

D'autre part, pour tout $H \in \mathcal{K}_c$, tout $k \in \mathbb{Z}$ et tout $\omega \in \Omega_1$, la relation

$$\lim_{n\to+\infty} H(X_n \circ \theta(\omega), n+k) = \lim_{n\to+\infty} H(X_n(\omega), n-1+k)$$

implique, d'après ce qui précède,

$$\hat{H}(\tau^{-k} \cdot \zeta \circ \theta(\omega), 0) = \hat{H}(\tau^{-k+1} \cdot \zeta(\omega), 0);$$

et par suite $\forall\, g \in G,\ \forall\, k \in \mathbb{Z}$,

$$\hat{H}(g \cdot \zeta \circ \theta(\omega), k) = \hat{H}(g\tau \cdot \zeta(\omega), k).$$

D'où l'on déduit l'assertion ii). □

• Représentation intégrale

Pour tout $x \in E$, nous appelons ρ_x l'image de la probabilité $\mathbb{P}_x$ par l'application ζ du lemme précédent.

Nous notons plus simplement $\hat{H}$ la fonction $\hat{H}(\cdot, 0)$; nous avons (lemme 6.4.9) $\forall\, x \in \varsigma, \forall\, k \in \mathbb{Z},\ H(x,k) = \hat{H}(\tau^{-k} \cdot \varsigma)$.

Théorème 6.4.11 *La famille de mesures de probabilité $\{\rho_x : x \in E\}$ sur $\mathcal{C}$ vérifie les propriétés suivantes :*
i) $\forall\, x \in E,\ \int_E \rho_u\, P(x,du) = \tau(\rho_x)$;
ii) $\forall\, \varsigma \in \mathcal{C}, \forall\, x \in \varsigma,\ \rho_x$ *est la mesure de Dirac en* ς.

Toute fonction H de $\mathcal{K}_c$ s'écrit :

$$H(x,k) = \int_{\mathcal{C}} \hat{H}(\tau^{-k} \cdot \varsigma)\, \rho_x(d\varsigma) \quad (x,k) \in E \times \mathbb{Z}.$$

L'application $H \to \hat{H}$ est une isométrie d'algèbres de Banach de $\mathcal{K}_c$ sur $C(\mathcal{C})$.

Preuve : Pour toute fonction continue $\hat{f}$ sur $\mathcal{C}$, nous avons (cf. preuve du théorème (6.2.7)) :

$$\begin{aligned}\int_{\mathcal{C}} \hat{f}(\tau \cdot \varsigma)\, \rho_x(d\varsigma) &= \mathbb{E}_x[\hat{f} \circ \tau \circ \zeta] = \mathbb{E}_x[\hat{f} \circ \zeta \circ \theta] \\ &= \int_E \int_{\mathcal{C}} \hat{f}(\varsigma)\, \rho_u(d\varsigma) P(x,du).\end{aligned}$$

D'où, la première propriété de la famille $\{\rho_x : x \in E\}$.

Soient $\varsigma \in \mathcal{C}$ et $x \in \varsigma$. Pour tout entier $n \geq 0$, nous avons $\mathbb{P}_x[X_n \in \tau^n \cdot \varsigma] = P^n(x, \tau^n \cdot \varsigma) = 1$. Si bien que, $\mathbb{P}_x$-p.s., $\zeta = \lim_{n\to+\infty} \tau^{\varphi(n)} \cdot \varsigma = \varsigma$ et par suite ρ_x est la mesure de Dirac en ς.

Soit $(\varphi(n))_{n\geq 0}$ une suite d'entiers naturels telle que la suite des isométries $(\tau^{\varphi(n)})_{n\geq 0}$ converge vers l'identité.

Pour tout $(x,k) \in E \times \mathbb{Z}$, nous avons :

$$H(x,k) \quad = \quad \mathbb{E}_x[\lim_{n\to+\infty} H(X_n, n+k)]$$

$$\begin{aligned} &= \mathbb{E}_x[\lim_{n\to+\infty} H(X_{\varphi(n)}, \varphi(n)+k)] \\ &= \mathbb{E}_x[\hat{H}(\tau^{-k}\cdot\zeta] \\ &= \int_{\mathcal{C}} \hat{H}(\tau^{-k}\cdot\varsigma)\,\rho_x(d\varsigma). \end{aligned}$$

On prouve enfin que l'application $H \to \hat{H}$ est une isométrie de $\mathcal{K}_c$ sur $C(\mathcal{C})$ en reprenant les arguments de la démonstration du théorème 6.2.7. □

6.5 Décomposition spectrale

Définitions et notations

Nous notons $\mathbb{T}$ le groupe multiplicatif des nombres complexes de module 1. Pour tout $\alpha \in \mathbb{T}$, nous désignons par $\mathcal{H}_\alpha$ [resp. $\mathcal{H}_{\alpha,c}$] l'espace vectoriel des fonctions boréliennes bornées [resp. des fonctions continues] f sur E telles que $Pf = \alpha f$. Appelons $\mathcal{P}$ [resp. $\mathcal{P}_c$] le fermé de $C(E)$ constitué par la réunion des $\mathcal{H}_\alpha, \alpha \in \mathbb{T}$ [resp. des $\mathcal{H}_{\alpha,c}, \alpha \in \mathbb{T}$].

Soient $\alpha \in \mathbb{T}$ et $h \in \mathcal{H}_\alpha$. La fonction H définie par $H(x,k) = \alpha^{-k}h(x),\ (x,k) \in E \times \mathbb{Z}$ appartient à $\mathcal{K}$. Soient $h_1 \in \mathcal{H}_\alpha$ et $h_2 \in \mathcal{H}_\beta$, avec $\alpha, \beta \in \mathbb{T}$. Le produit $H_1 * H_2$ des éléments de $\mathcal{K}$ correspondants respectivement à h_1 et h_2 s'écrit $H_1 * H_2(x,k) = (\alpha\beta)^{-k}h(x)$; où h est un élément de $\mathcal{H}_{\alpha\beta}$.

Le fermé $\mathcal{P}$ [resp. $\mathcal{P}_c$] s'identifie alors à un fermé de $\mathcal{K}$, stable pour le produit. Le sous-espace vectoriel de $\mathcal{K}$ resp. $\mathcal{K}_c$] engendré par ce fermé constitue donc une sous-algèbre de l'algèbre $\mathcal{K}$ [resp. $\mathcal{K}_c$].

Nous notons $Sp(P)$ l'ensemble des nombres complexes α de $\mathbb{T}$ tels que l'espace vectoriel $\mathcal{H}_{\alpha,c}$ ne soit pas réduit à la fonction nulle.

Appelons **caractère** sur G tout homomorphisme continu de groupe de G dans $\mathbb{T}$ et désignons par $\hat{G}$ le groupe des caractères sur G.

Etude de $\mathcal{P}_c$

Théorème 6.5.1 *Soit h un élément de $\mathcal{P}_c$. Pour toute classe cyclique ς nous notons $\hat{h}(\varsigma)$ la valeur de la fonction h sur ς.*

La formule

$$h(x) \ = \int_{\mathcal{C}} \hat{h}(\varsigma)\,\rho_x(d\varsigma)$$

établit une isométrie de $\mathcal{P}_c$ sur l'ensemble des fonctions $\hat{h}$ de $C(\mathcal{C})$ pour lesquelles il existe un caractère χ sur G tel que :

$$\forall \varsigma \in \mathcal{C},\ \forall g \in G,\ \ \hat{h}(g\cdot\varsigma) = \chi(g)\hat{h}(\varsigma).$$

Preuve : Soient $\alpha \in Sp(P)$ et $h \in \mathcal{H}_{\alpha,c}$. D'après le théorème 6.4.11, nous avons, pour tout $(x,k) \in E \times \mathbb{Z}$,

$$\alpha^{-k}h(x) \ = \int_{\mathcal{C}} \hat{h}(\tau^{-k}\cdot\varsigma)\ \rho_x(d\varsigma).$$

Si bien que $\forall \varsigma \in \mathcal{C}$, $\forall k \in \mathbb{Z}$,

$$\hat{h}(\tau^{-k}\cdot\varsigma) = \alpha^{-k}\,\hat{h}(\varsigma).$$

Pour achever la démonstration du théorème, il suffit alors de montrer qu'il existe un caractère χ sur G tel que $\chi(\tau) = \alpha$.

De la dernière égalité, il résulte que, pour toute suite d'entiers $(\varphi(n))_{n\geq 0}$ telle que, la suite des isométries $(\tau^{\varphi(n)})_{n\geq 0}$ converge vers l'identité, la suite $(\alpha^{\varphi(n)})_{n\geq 0}$ converge vers 1. Soit g une isométrie de G. Considérons une suite d'entiers naturels $(\psi(n))_{n\geq 0}$ telle que $\tau^{\psi(n)} \to g$. A l'aide de la remarque précédente, on voit aisément que la suite $(\alpha^{\psi(n)})_{n\geq 0}$ converge et que cette limite ne dépend pas du choix de la suite $(\psi(n))_{n\geq 0}$. En posant $\chi(g) = \lim_{n\to+\infty} \alpha^{\psi(n)}$, on définit un caractère sur G qui vérifie la propriété voulue.

□

Corollaire 6.5.2 *La sous-algèbre (ou sous-espace) de $\mathcal{K}_c$ engendré par $\mathcal{P}_c$ est dense dans $\mathcal{K}_c$.*

Preuve : Il suffit de montrer que la famille de fonctions $\{\hat{h} : h \in \mathcal{P}_c\}$ sépare les points de $\mathcal{C}$.

Désignons par κ la mesure de Haar normalisée sur le groupe compact abélien G. Pour tout caractère χ sur G et toute fonction H de $\mathcal{K}_c$, soit $\hat{h}_\chi$ la fonction de $C(\mathcal{C})$ définie par

$$\hat{h}_\chi(\varsigma) = \int_G \chi(g)\hat{H}(g\cdot\varsigma)\,\kappa(dg) \quad (\varsigma \in \mathcal{C}).$$

Pour tout $g \in G$, nous avons $\hat{h}_\chi(g\cdot\varsigma) = \chi^{-1}(g)\hat{h}_\chi(\varsigma)$. Les fonctions $\{\hat{h}_\chi : \chi \in \hat{G}\}$ correspondent donc à des fonctions de $\mathcal{P}_c$ par l'isométrie du théorème 6.5.1.

Soient ς_1 et ς_2 deux classes cycliques telles que $h(\varsigma_1) = h(\varsigma_2)$ pour toute fonction h e $\mathcal{P}_c$. De ce qui précède, il résulte que

$$\forall\, H \in \mathcal{K}_c, \forall\, \chi \in \hat{G},\ \hat{h}_\chi(\varsigma_1) = \hat{h}_\chi(\varsigma_2).$$

Autrement dit, pour tout $H \in \mathcal{K}_c$, les fonctions continues f_1 et f_2 sur G définies par $f_i(g) = \hat{H}(g\cdot\varsigma_i) \quad (g \in G) \quad (i \in \{1,2\})$, ont les mêmes coefficients de Fourier. Elles sont donc égales ([]Rudin) et l'on a $\forall\, H \in \mathcal{K}_c, \forall\, g \in G,\ \hat{H}(g\cdot\varsigma_1) = \hat{H}(g\cdot\varsigma_2)$; ce qui montre que $\varsigma_1 = \varsigma_2$. □

• Décomposition spectrale

Théorème 6.5.3 *Supposons que, pour toute fonction continue f, la famille de fonctions $\{P^n f : n \geq 0\}$ soit équicontinue. Alors l'espace des fonctions continues $C(E)$ sur E se décompose en la somme directe de deux sous-espaces fermés $\mathcal{E}_1$ et $\mathcal{E}_2$ tels que :*
i) $\mathcal{E}_1$ est le sous-espace fermé de $C(E)$ engendré par les fonctions propres de P associées à des valeurs propres de module 1.
ii) pour toute fonction f de $\mathcal{E}_2$, la suite de fonctions $(P^n f)_{n\geq 0}$ converge uniformément vers la fonction nulle.

Preuve : Soit $(\varphi(n))_{n\geq 0}$ une suite d'entiers naturels telle que :
i) la suite $(\tau^{\varphi(n)})_{n\geq 0}$ converge vers l'identité ;
ii) pour toute fonction f de $C(E)$ et tout entier k de $\mathbb{Z}$, la suite de fonctions $(P^{\varphi(n)-k} f)_{n\geq 0}$ converge uniformément sur E.

En posant, pour tout $(x,k) \in E \times \mathbb{Z}$, $H_f(x,k) = \lim_{n\to+\infty} P^{\varphi(n)-k} f(x)$, nous obtenons une fonction H_f de $\mathcal{K}_c$ qui s'écrit donc (théorème 6.4.11)

$$H_f(x,k) = \int_{\mathcal{C}} \widehat{H_f}(\tau^{-k}\cdot\varsigma)\,\rho_x(d\varsigma).$$

Pour $f \in C(E)$, posons :

$$\Pi f = \lim_{n\to+\infty} P^{\varphi(n)} f = H_f(\cdot, 0) = \int_{\mathcal{C}} \widehat{H_f}(\varsigma)\, \rho_{\cdot}(d\varsigma).$$

Des égalités

$$\forall\, x \in E,\, \forall\, n \geq 0,\;\; P^{\varphi(n)}\Pi f(x) = \int_{\mathcal{C}} \widehat{H_f}(\tau^{\varphi(n)} \cdot \varsigma)\, \rho_x(d\varsigma),$$

il résulte que l'on définit ainsi un projecteur sur $C(E)$ tel que la suite de fonctions $(P^{\varphi(n)}(f - \Pi f))_{n\geq 0}$ converge uniformément sur E vers la fonction nulle. Comme P est une contraction de $C(E)$, on en déduit que la suite de fonctions $(P^n(f - \Pi f))_{n\geq 0}$ converge uniformément sur E vers la fonction nulle. Enfin d'après le corollaire 6.5.2, l'image du projecteur Π n'est autre que le sous-espace fermé de $C(E)$ engendré par les fonctions propres de P associées à des valeurs propres de module 1. $\square$

Chapitre 7

Exemple de chaîne de Markov (problème de Dirichlet)

Nous étudions ici une classe d'exemples de chaînes de Markov reliées au problème de Dirichlet dans des domaines bornés de $\mathbb{R}^d$.

7.1 Fonctions harmoniques

Soit D un ouvert du plan complexe. Une fonction $f : z = (x, y) \to f(x, y)$ définie sur D, différentiable comme fonction de x et de y, est holomorphe si et seulement si elle vérifie l'équation de Cauchy-Riemann,

$$\frac{\partial}{\partial x} f = -i \frac{\partial}{\partial y} f.$$

Si $f = u + iv$, u et v étant la partie réelle et la partie imaginaire de f, l'équation de Cauchy-Riemann s'écrit : $\frac{\partial}{\partial x} u = \frac{\partial}{\partial y} v$, $\frac{\partial}{\partial y} u = -\frac{\partial}{\partial x} v$.

Introduisons les opérateurs de dérivation partielle :

$$\partial = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}), \overline{\partial} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}).$$

L'équation de Cauchy-Riemann s'écrit alors : $\overline{\partial} f(z) = 0$.

On note Δ le Laplacien, opérateur aux dérivées partielles du second ordre défini par :

$$\Delta f = \frac{\partial^2}{\partial x^2} f + \frac{\partial^2}{\partial y^2} f.$$

Si une fonction f a des dérivées partielles du second ordre continues, elles vérifient $\frac{\partial^2}{\partial y \partial x} f = \frac{\partial^2}{\partial x \partial y} f$.
On a donc : $\Delta f = 4\partial\overline{\partial} f$.

Définition 7.1.1 Une fonction f est dite **harmonique** sur D, si elle vérifie $\partial\overline{\partial} f = 0$.

Une fonction à valeurs complexes est harmonique si, et seulement si, ses parties réelles et imaginaires sont harmoniques. Les fonctions holomorphes, leur partie réelle, sont donc des fonctions harmoniques.

Prenons maintenant comme domaine D le disque unité ouvert $U = \{z \in \mathbb{C} : |z| < 1\}$. Le cercle $\mathbb{T}$ peut être vu comme la frontière de U. La restriction au bord d'une fonction définie sur le disque définit une fonction périodique. Inversement, nous allons étudier une méthode de prolongement au disque d'une fonction périodique, vue comme fonction définie sur le bord du disque.

Prolongement à l'intérieur du disque d'une fonction définie sur le cercle unité

Dans la suite, r est un paramètre réel, $0 \le r < 1$. Le noyau de Poisson que nous noterons P_r est la fonction définie, pour $t \in \mathbb{R}$, par :

$$P_r(t) = \sum_{-\infty}^{+\infty} r^{|n|} e^{2\pi int} = \frac{1-r^2}{|1-e^{2\pi it}r|^2} = \frac{1-r^2}{1+r^2-2r\cos 2\pi t}.$$

Proposition 7.1.2 *La famille $(P_r, 0 < r < 1)$ est une famille de fonctions positives, d'intégrale égale à 1 sur $[0,1]$, formant dans la classe des fonctions périodiques de période 1 une identité approchée quand r tend vers 1 par valeurs inférieures.*

Preuve : La première forme montre qu'on a $\int_0^1 P_r(t)\, dt = 1$. La deuxième forme montre que $P_r(t)$ est un noyau positif.

En étudiant le comportement de $P_r(t)$, quand $r \to 1^-$, pour $\varepsilon \le t \le 1-\varepsilon$ où ε est un nombre fixé vérifiant $0 < \varepsilon < 1$, on obtient le résultat. $\square$

Si f une fonction continue, périodique, de période 1, on en déduit que $\lim_{r\to 1^-} P_r * f = f$.

Si u est une fonction continue et réelle sur le disque unité fermé $\overline{D}$, harmonique à l'intérieur, elle est égale dans le disque ouvert à l'intégrale de Poisson de sa restriction au bord et elle est la partie réelle de la fonction holomorphe sur D définie par :

$$f(z) = \int_0^1 \frac{e^{2\pi it}+z}{e^{2\pi it}-z}\, u(e^{2\pi it})\, dt.$$

Corollaire 7.1.3 *Toute fonction harmonique réelle est localement la partie réelle d'une fonction holomorphe, qui est unique à une constante imaginaire pure près.*

Preuve : On applique le résultat précédent à un petit disque. $\square$

7.2 Propriété de la moyenne

Définition 7.2.1 Une fonction g définie et continue sur un ouvert D a la **propriété de la valeur moyenne**, si, pour tout $z \in D$, il existe une suite (r_n) de nombres > 0, tendant vers 0, telle que $u(z)$ soit égal à la valeur moyenne de u sur les cercles de rayons r_n de centre z :

$$u(z) = \int_0^1 u(z + r_n e^{2\pi it})\, dt.$$

D'après la formule de Poisson, si u est harmonique, la relation précédente est vérifiée, $\forall r > 0$ tel que le disque fermé de rayon r et de centre z soit contenu dans D. Inversement, on le théorème suivant :

Théorème 7.2.2 *Si une fonction continue u a la propriété de la moyenne dans un ouvert D, elle est harmonique dans D.*

Preuve : On peut supposer u réelle. Soit $\overline{B}(a,r)$ un disque de centre a inclus dans D. Par l'intégrale de Poisson, on peut construire une fonction h harmonique à l'intérieur de ce disque et égale à u sur sa frontière. Nous allons montrer que, pour tout $a \in D$, u coïncide avec h qui est harmonique sur $B(a,r)$, ce qui prouvera que u est harmonique sur D.

Posons $v = u - h$. La fonction v vérifie encore la propriété de la moyenne. Soit $m = \sup\{v(z), z \in \overline{D}(a,r)\}$. Supposons qu'on ait $m > 0$. Soit M l'ensemble des z dans $\overline{D}(a,r)$ tels que $v(z) = m$. Comme v est nulle sur la frontière, M est un sous-ensemble compact inclus dans la disque ouvert $D(a,r)$. Soit $z_0 \in M$, tel que $|z_0 - a| \geq |z - a|$, $\forall z \in M$. Pour $\varepsilon > 0$ assez petit, le cercle de rayon ε de centre z_0 contient au moins un demi-cercle entièrement inclus dans le complémentaire de M. La valeur moyenne de v sur ce cercle est strictement inférieure à $m = v(z_0)$, ce qui contredit la propriété de la moyenne. On a donc $m = 0$. En raisonnant de même avec $-v$, on obtient que v est nulle sur $D(a,r)$ et donc que u coïncide avec h. $\square$

Soit ϕ une fonction définie et continue sur le cercle unité. Posons

$$c_n = \int_0^1 \phi(e^{2\pi i t}) e^{-2\pi i n t}\, dt, \text{ pour } n \in \mathbb{Z}.$$

Si ϕ vérifie la condition : $c_n = 0$ pour tout $n \leq -1$, (i.e. si $f \in H^1$, sous-espace de $L^1(\mathbb{T}^1)$ formé des fonctions telles que $c_n(f) = 0, \forall n < 0$), la fonction Φ définie par

$$\Phi(z) = \sum_0^\infty z^n c_n$$

est holomorphe dans D. La série converge uniformément sur tout disque compact contenu dans D, et on a

$$\lim_{r \to 1^-} \Phi(re^{it}) = \phi(e^{it}),$$

pour tout $t \in \mathbb{R}$, d'après la propriété d'identité approchée de la famille (P_r) définie plus haut.

Chaîne de Markov associée à la propriété de moyenne

Considérons un domaine borné D ouvert dans $\mathbb{R}^2$ de frontière ∂D. Notons $\overline{D}$ sa fermeture, $\delta(x)$ la distance d'un point x de D à ∂D.

Soit α une fonction borélienne sur D vérifiant $\alpha(x) \leq \delta(x)$, $\forall x \in D$. On définit un opérateur markovien P_α (noté aussi simplement P) sur D par

$$P_\alpha f(x) = \frac{1}{m(B(x,\alpha(x)))} \int_{B(x,\alpha(x))} f(y)\, dm(y),$$

où m est la mesure de Lebesgue 2-dimensionnelle. On peut définir P_α sur la frontière, en posant $P_\alpha f(x) = f(x)$, pour $x \in \partial D$.

Notons (X_n) la chaîne de Markov sur D définie par le noyau P et par $X_0 = x \in D$. La loi de X_{n+1} sachant que $X_n = y$ est la loi uniforme sur la boule $B(x,\alpha(x))$.

L'étude de cette chaîne a été développée en particulier par Feller, dans le cas où D est un disque et $\alpha = \delta$, le disque $B(x,\alpha(x))$ étant alors le disque tangent au bord, de centre x.

On suppose dans la suite que α est une fonction continue sur D, ce qui assure que l'opérateur de moyenne P_α associé préserve les fonctions continues.

Si f est une fonction harmonique au sens du Laplacien, elle vérifie la propriété de la moyenne, on a donc $P_\alpha f = f$. La suite $(f(X_n))_{n\geq 0}$ est alors une martingale. Si f est bornée, cette martingale bornée est convergente. C'est en particulier le cas pour la fonction $f(x) = x$ (toute fonction affine est harmonique). La chaîne (X_n) est donc une martingale bornée convergente.

Lemme 7.2.3 *La chaîne (X_n) est une martingale bornée qui converge $\mathbb{P}_x$-presque sûrement vers une limite $X_\infty \in \partial D$.*

Le lemme résulte du lemme général suivant :

Lemme 7.2.4 *Si une chaîne de Markov (X_n) est une martingale bornée, correspondant à un opérateur markovien P préservant l'espace des fonctions continues, elle converge $\mathbb{P}_x$-presque sûrement vers une limite X_∞, X_∞ étant une v.a. à valeur dans l'ensemble des points fixes de P.*

Preuve : Comme (X_n) est bornée, elle converge $\mathbb{P}_x$-presque sûrement vers une limite X_∞.

Soit $P : f \to \int f(y)\ P(x, dy)$ l'opérateur markovien définissant la chaîne. Comme (X_n) est une martingale de carré intégrable, on a $\sum_n \mathbb{E}_x[\|X_{n+1} - X_n\|^2] < \infty$.

Soit f une fonction continue à support compact dans $\mathbb{R}^2$. On a :

$$\begin{aligned}\mathbb{E}_x[|\int f(y - X_n)\ P(X_n, dy) - f(0)|] &= \mathbb{E}_x[|\mathbb{E}[f(X_{n+1} - X_n)|\mathcal{F}_n] - f(0)|] \\ &= \mathbb{E}_x[|f(X_{n+1} - X_n)] - f(0)| \to 0.\end{aligned}$$

Pour x fixé, il existe une suite $\phi(n)$ telle que

$$\lim_n \int f(y - X_{\phi(n)})\ P(X_{\phi(n)}, dy) = f(0).$$

La fonction $x \to \int f(y - x)\ P(x, dy)$ étant continue, on obtient, en passant à la limite :

$$\int f(y - X_\infty)\ P(X_\infty, dy) = f(0),$$

ce qui implique que $P(X_\infty, dy) = \delta_{X_\infty}(dy)$. $\square$

Si f est une fonction continue sur $\overline{D}$, la convergence $\lim_n X_n = X_\infty$ implique $\lim_n f(X_n) = f(X_\infty)$ et donc

$$\lim_n P^n f(x) = \lim_n \mathbb{E}_x[f(X_n)] = \mathbb{E}_x[f(X_\infty)].$$

7.3 Cas d'un domaine convexe, exemples

Lemme 7.3.1 *Si le domaine est convexe, pour toute fonction f continue sur $\overline{D}$, la suite $(P^n f)$ est équicontinue.*

Preuve : La fonction δ (distance au bord) est une fonction concave (borne inférieure de fonctions affines). Elle vérifie donc $P(\delta) \leq \delta$.

La suite $(P^n\delta)_{n\geq 0}$ est une suite décroissante de fonctions continues sur $\overline{D}$. Comme $\lim_n P^n\delta(x) = \mathbb{E}_x[\delta(X_\infty)] = 0$, on en déduit, par le lemme de Dini, la convergence uniforme vers 0 de la suite $(P^n\delta)_{n\geq 0}$.

Soit f une fonction continue sur $\overline{D}$ nulle au bord. Montrons que la suite de fonctions $(P^n f(x))$ converge uniformément vers 0. Désignons par $C_\eta = \{x \in D : \delta(x) \leq \eta\}$ une "couronne" d'épaisseur $\eta > 0$, voisinage du bord, et par C_η^c son complémentaire dans D. Soient $\varepsilon > 0$ et $\eta > 0$ tels que $|\int_{C_\eta} f(y)\ P^n(x, dy) \leq \sup_{y \in C_\eta} |f(y)| < \varepsilon$. On a :

$$\begin{aligned}
P^n f(x) &= \int_{C_\eta} f(y)\ P^n(x, dy) + \int_{C_\eta^c} f(y)\ P^n(x, dy) \\
&\leq \varepsilon + + \|f\|_\infty P^n(x, C_\eta^c) \\
&\leq \varepsilon + \|f\|_\infty \mathbb{P}_x[\{\delta(X_n) \geq \eta\}] \\
&\leq \varepsilon + + \frac{\|f\|_\infty}{\eta} P^n \delta(x) \leq 2\varepsilon,
\end{aligned}$$

pour n assez grand, d'après la convergence uniforme de $(P^n(\delta))$ vers 0.

□

Dans ce cas, pour toute fonction f sur $\overline{D}$, la fonction $\mathbb{E}_x[f(X_\infty)]$, qui est égale à $\lim_n P^n f(x)$, est continue.

Exemples :

1) Cas du disque

Dans le cas d'un disque, on retrouve la formule de prolongement, à l'aide du noyau de Poisson, d'une fonction continue sur le bord en une fonction harmonique à l'intérieur.

2) Cas du disque "pointé". Prenons pour D le dique ouvert de rayon 1, privé de son centre.

La fonction $h(x) = -\log(x_1^2 + x_2^2)$ est harmonique dans D. On sait que

$$-\log \|X_n\| \to -\log \|X_\infty\|.$$

La limite est $\mathbb{E}_x$-intégrable, pour tout $x \in D$, d'après le lemme de Fatou, puisque $\mathbb{E}_x[-\log \|X_n\|] \leq -\log \|x\|$; d'où $\mathbb{E}_x[\lim_n \|X_n\|] \leq -\log \|x\|$. On a donc $\mathbb{E}_x[X_\infty = 0] = 0$.

non équicontinuité de la suite $(P^n f)$

3) Cas d'une couronne. Prenons comme domaine ouvert D la couronne ouverte $D = \{x : r < \|x\| < 1\}$.

La fonc h étant harmonique, on obtient :

$$h(x) = \mathbb{E}_x[h(X_\infty)] = 0.\mathbb{P}_x(\|X_\infty\| = 1) + \log r^2 \mathbb{P}_x(\|X_\infty\| = r);$$

d'où :

$$\mathbb{P}_x(\|X_\infty\| = r) = \frac{\log \|x\|}{\log r}, \quad \mathbb{P}_x(\|X_\infty\| = 1) = \frac{1 - \log \|x\|}{\log r}.$$

Expression du "noyau" de Poisson.

Ici, il n'y a pas équicontinuité de la suite $(P^n f)$.

7.4 Convergence vers le bord : loi de grands nombres et TCL

Notons $\phi(x) = -\log \delta(x)$.

Lemme 7.4.1 *Les fonctions $\phi - P\phi$, $\phi^2 - (P\phi)^2$ sont prolongeables au bord.*

Preuve : ? ? $\square$

Soit $\psi = P\phi - \phi$. On a donc $P^n\phi = \phi + \sum_0^{n-1} P^k\psi$. La suite $Z_k = \phi(X_k) - P\phi(X_{k-1}$ est une suite d'accroissements de martingale, et on a :

$$\phi(X_n) = \phi(X_0) + \sum_0^n [\phi(X_k) - P\phi(X_{k-1})] + \sum_0^{n-1} [P\phi - \phi)(X_k).$$

La martingale vérifie la loi des grands nombres et d'autre part, on a $\psi(X_k) \to \psi(X_\infty)$. On a ainsi pour $(\phi(X_n))$ la loi des grands nombres :

$$\lim_n \frac{1}{n}\phi(X_n) = \psi(X_\infty).$$

Nous allons maintenant établir un TCL dans cette convergence :

$$\frac{1}{\sqrt{n}}(\phi(X_n) - n\psi(X_\infty)) \xrightarrow{d} N(O, \sigma^2).$$

La loi des grands nombres montre qu'on a une convergence exponentiellement rapide de (X_n) vers la frontière. Comme ψ est lipschitzienne, la série $\sum |\psi(X_n) - \psi(X_\infty)|$ converge, ce qui permet de négliger le terme résiduel...

Il rest à appliquer à la partie martingale de l'expression la méthode de Brown (théorème ? ?). Le seul point à vérifier est la condition de Lindeberg.

Chapitre 8

Systèmes dynamiques, théorèmes ergodiques

Considérons un espace mesuré $(E, \mathcal{E}, \mu)$ et une transformation mesurable et surjective τ de E dans lui-même, laissant μ-invariante ($\mu(\tau^{-1}A) = \mu(A),\ \forall A \in \mathcal{E}$). Notons T l'opérateur de composition par la transformation τ (i.e. $Tf(x) = f(\tau x), \forall x \in E$).

Cet opérateur positif opère sur l'espace des fonctions $\mathcal{E}$-mesurables sur E, à valeurs réelles ou complexes et, d'après l'invariance de la mesure μ, définit une isométrie (en général non surjective) de chaque espace $L^p(E, \mathcal{E}, \mu)$, $p \geq 1$.

Plus généralement, on peut considérer un opérateur positif qui est de norme 1 au sens des normes L^1 et L^∞.

Etant donnée une fonction f $\mathcal{E}$-mesurable sur E, ses **sommes ergodiques** (appelées aussi sommes de Birkhoff) et ses **moyennes ergodiques** sont définies respectivement par

$$S_n f(x) := \sum_{k=0}^{n-1} T^k f(x), x \in E,\ n \geq 1, \tag{8.1}$$

$$M_n f(x) := \frac{1}{n}\, S_n f(x), x \in E,\ n \geq 1. \tag{8.2}$$

Alors que l'itération de T présente en général un comportement très irrégulier, les théorèmes ergodiques permettent de décrire, sous des hypothèses générales, le comportement asymptotique (μ-presque sûr ou en norme) des *moyennes ergodiques*.

Le théorème ergodique, sous sa forme ponctuelle (convergence presque partout des moyennes ergodiques) a été obtenu par Birkhof en 1931 pour des opérateurs de composition par une transformation préservant une mesure finie, puis étendu par Hopf à des opérateurs plus généraux de $L^1(E, \mathcal{E}, \mu)$. Parallèlement, Von Neumann a montré un théorème de convergence en norme, applicable à un opérateur de contraction dans un espace de Hilbert, résultat généralisé par Kakutani et Yoshida pour les espaces de Banach réflexifs.

8.1 Théorèmes ergodiques en moyenne

Une forme générale de théorème ergodique (en moyenne) peut être établie pour des opérateurs de contraction sur un espace de Banach, sous des conditions sur cet espace. Nous commençons par le premier résultat de ce type, le théorème de Von Neumann, pour les contractions des espaces de Hilbert.

• Le théorème de Von Neumann

Désignons par $(H, \langle\ ,\ \rangle)$ un espace de Hilbert et par $\|\ \|$ la norme sur H associée à son produit scalaire $\langle\ ,\ \rangle$. Soit T une application linéaire de H dans H de norme au plus 1 :

$$\|T\| = \sup\{\|Tf\| : \|f\| = 1\} = \sup\{\langle Tf, g\rangle : \|f\| = \|g\| = 1\} \leq 1.$$

Nous notons T^* l'adjoint de T défini par $\langle Tf, g\rangle = \langle f, T^*g\rangle$.

Le théorème de Von Neumann décrit le comportement des moyennes des itérés de T. Donnons d'abord un résultat sur les vecteurs T-invariants, qui reprend le lemme 3.3.5.

Lemme 8.1.1 *Si T est une contraction dans un espace de Hilbert, T et son adjoint T^* ont les mêmes vecteurs invariants. Si $Tf = \lambda f$, avec $|\lambda| = 1$, alors $T^*f = \overline{\lambda} f$.*

Preuve : Pour une contraction dans un espace de Hilbert, la relation $Tf = f$ est équivalente à $\langle Tf, f\rangle = \|f\|^2$, (cas d'égalité dans l'inégalité de Cauchy-Schwarz).

Les opérateurs T et T^* ayant la même norme, T^* est aussi une contraction et les opérateurs ont les mêmes vecteurs invariants.

Ce résultat appliqué à $\overline{\lambda}T$, où λ est une valeur propre de module 1 de T, montre la deuxième assertion. □

Théorème 8.1.2 *(Von Neumann) Soit T un opérateur de contraction sur un espace de Hilbert $(H, \langle\ ,\ \rangle)$. Pour tout $f \in H$, la suite des moyennes $(M_n f)_{n\geq 1} = (\frac{1}{n}\sum_0^{n-1} T^k f)_{n\geq 1}$ converge en norme vers la projection orthogonale Πf de f sur le sous-espace formé des vecteurs T-invariants.*

Preuve : Notons $H = V \oplus V^\perp$ la décomposition de H en la somme directe du sous-espace V formé des vecteurs T-invariants et du sous-espace $V^\perp$ orthogonal de V. Cette décomposition est invariante par T et par T^*. Soit Π le projecteur orthogonal sur V.

Les vecteurs de la forme $h - Th,\ h \in H$, forment un sous-espace W_0 dense de $V^\perp$. En effet, le lemme 8.1.1 implique que ces vecteurs sont orthogonaux à V et que tout vecteur orthogonal à W_0 est invariant par T^* donc par T.

Soient $f \in H$ et $(h_p)_{p\geq 0}$ une suite d'éléments de H telle que

$$\lim_{p\to+\infty} \|h_p - Th_p - (f - \Pi f)\| = 0.$$

Nous avons, pour tous entiers $p \geq 0$, $n \geq 1$,

$$\begin{aligned}
\|M_n f - \Pi f\| &= \|M_n(f - \Pi f)\| \\
&\leq \|M_n(h_p - Th_p)\| + \|h_p - Th_p - (f - \Pi f)\| \\
&\leq \|\frac{1}{n}\ (h_p - T^n h_p)\| + \|h_p - Th_p - (f - \Pi f)\| \\
&\leq \frac{2}{n}\ \|h_p\| + \|h_p - Th_p - (f - \Pi f)\|\ ;
\end{aligned}$$

et par suite,

$$\limsup_n \|M_n f - \Pi f\| \leq \|h_p - Th_p - (f - \Pi f)\|.$$

D'où le résultat, en faisant tendre p vers $+\infty$. $\square$

• **Le théorème ergodique de Kakutani-Yoshida**

Soit maintenant B un espace de Banach. Pour f dans B et ϕ dans son dual B^*, nous notons $\langle f, \phi\rangle$ la valeur de la forme linéaire ϕ sur f.

Si T est une application linéaire continue de B dans B, l'application duale T^* de T opérant sur l'espace dual B^* de B est définie par

$$\langle Tf, \phi\rangle = \langle f, T^*\phi\rangle, \ f \in B, \phi \in B^*.$$

L'argument de la preuve du théorème de Von Neuman s'étend aux espaces de Banach **réflexifs**, i.e. égaux à leur bidual.

Lemme 8.1.3 *Si T est une contraction d'un Banach réflexif B, alors l'ensemble $\{g + h - Th : g,\ h \in B, Tg = g\}$ est dense dans B.*

Preuve : Soit f un élément quelconque dans B. Comme T est une contraction, les moyennes $M_n f = \frac{1}{n}\sum_0^{n-1} T^k f$ sont bornées en norme. Comme B est réflexif, on obtient en utilisant le théorème de Hahn-Banach que tout sous-espace fermé de B est réflexif. Appliquons cette propriété au sous-espace fermé B_0 engendré par les vecteurs $\{T^k f, k \geq 0\}$, qui est séparable. Par compacité faible de la boule unité de B_0 considéré comme le dual de son dual, nous pouvons extraire de la suite des moyennes $(M_n f)_{n\geq 1}$ une sous-suite $(M_{N_j} f)_{j\geq 1}$ qui converge faiblement vers une limite $\widetilde{f} \in B_0$. Nous avons, pour tout $\psi \in B^*$, par "cancellation" :

$$\lim_j [\langle \frac{1}{N_j} \sum_0^{N_j - 1} T^k f, T^*\psi\rangle - \langle \frac{1}{N_j} \sum_0^{N_j - 1} T^k f, \psi\rangle] = 0,$$

d'où : $\langle T\widetilde{f}, \psi\rangle = \langle \widetilde{f}, \psi\rangle, \forall \psi \in B^*$. Ceci implique l'invariance de $\widetilde{f}$.

Soit maintenant $\phi \in B^*$ tel que

$$\langle g + h - Th, \phi\rangle = 0, \ \forall h \in B, \ \forall g \ T\text{-invariante}.$$

Montrons que $\phi = 0$, ce qui établira le résultat de densité annoncé.

La relation $\langle h - Th, \phi\rangle = 0$, $\forall h \in B$ montre que $T^*\phi = \phi$, ce qui implique :

$$\langle \frac{1}{N} \sum_0^{N-1} T^k f, \phi\rangle = \langle f, \phi\rangle;$$

d'où $\langle f, \phi\rangle = \langle \widetilde{f}, \phi\rangle$.

Comme on a aussi $\langle \widetilde{f}, \phi\rangle = 0$, $\widetilde{f}$ étant T-invariant, il en résulte : $\langle f, \phi\rangle = 0, \forall f \in B$, d'où $\phi = 0$. $\square$

Théorème 8.1.4 *Soit T une contraction sur un espace de Banach réflexif B. Il existe un opérateur Π de projection sur le sous-espace des vecteurs T-invariants tel que, pour tout $f \in B$, la suite des moyennes $(M_n f)_{n\geq 1}$ convergent en norme vers Πf.*

Preuve : Montrons que, pour tout $f \in B$, la suite des moyennes $(M_n f)_{n\geq 1}$ forme une suite de Cauchy. Soit $\varepsilon > 0$. D'après le lemme 8.1.3, il existe $h, g \in B$, avec $Tg = g$, tels que $\|f - (g + Th - h)\| < \varepsilon$. Comme T est une contraction, il en résulte :

$$\|M_n f - M_m f\| \leq 2\varepsilon + (\frac{2}{n} + \frac{2}{m})\|h\|;$$

ce qui établit la propriété.

Notons $\Pi f := \lim_n M_n f$. L'opérateur Π est projecteur de B dans B : nous avons $T\Pi = \Pi T = \Pi$, ce qui entraîne : $\Pi^2 f = \lim_n M_n(\Pi f) = \Pi f$.

□

La réflexivité des espaces L^p, pour $1 < p < \infty$ permet d'appliquer le théorème 8.1.4. Pour les espaces L^p, une preuve du théorème ergodique peut également être basée sur une propriété de stricte convexité (cf. lemme 8.1.5 ci-dessous). Cette preuve présente l'avantage de caractériser la limite de la suite des moyennes $(M_n f)_{n\geq 1}$ par une propriété de minimisation de la norme.

Lemme 8.1.5 *Pour tout $p \in]1, \infty[$, il existe une constante C_p telle que, pour $f, g \in L^p$ vérifiant $\|f\|_p \leq 1$, $\|g\|_p \leq 1$, on ait :*

$$\|f - g\|_p^p \leq C_p[\|f\|_p^p + \|g\|_p^p - 2\|\frac{f+g}{2}\|_p^p]^{\inf(1,p/2)}. \tag{8.3}$$

Voir Dunford-Schwarz [DunSch58] pour une preuve du lemme.

Théorème 8.1.6 *Soit $p \in \,]1, \infty[$. Si T est une contraction de $L^p(E, \mathcal{E}, \mu)$, pour toute fonction $f \in L^p$ la suite des moyennes ergodiques $(M_n f)_{n\geq 1}$ converge en norme $\| \ \|_p$ vers une limite Πf, caractérisée comme le vecteur de norme minimale dans le convexe fermé $C(f)$ de $L^p(E, \mathcal{E}, \mu)$ engendré par l'orbite $\{T^n f, n \geq 0\}$ de f sous l'action de T.*

Preuve : Montrons que la suite des moyennes $(M_n f)$ est une suite de Cauchy.

Soit $\alpha = \inf\{\|h\|_p, h \in C(f)\}$. Soit $\varepsilon > 0$. Il existe un barycentre f_r de la forme $f_r = \sum_0^r a_j T^j f$, pour des $a_j \geq 0$ de sommes égale à 1, tel que $\|f_r\|_p < \alpha + \varepsilon$. Nous avons, pour $n \geq r$,

$$S_n f_r - S_n f = T^n \sum_{j=1}^{r} a_j S_j f - \sum_{j=1}^{r} a_j S_j f,$$

d'où :

$$\begin{aligned}
\alpha &\leq \|M_n f\|_p \leq \|M_n f_r\|_p + \frac{2}{n}\sum_{j=1}^{r} a_j \|S_j f\|_p \\
&\leq \|f_r\|_p + \frac{2}{n}\sum_{j=1}^{r} a_j \|S_j f\|_p.
\end{aligned}$$

Pour L assez grand, nous avons $\|M_n f\|_p \leq \alpha + 2\varepsilon$, pour $n > L$ et, d'après le lemme 8.1.5, en posant $\gamma = \inf(1, p/2)$, pour $n, m > L$,

$$\begin{aligned} \|M_n f - M_m f\|_p^p &\leq C_p[\|M_n f\|_p^p + \|M_m f\|_p^p - 2\|\frac{M_n f + M_m f}{2}\|_p^p]^\gamma \\ &\leq C_p[2(\alpha + 2\varepsilon)^p - 2\alpha^p]^\gamma \leq C'_p[2\varepsilon]^\gamma. \end{aligned}$$

La suite des moyennes $(M_n f)_{n\geq 1}$ converge donc en norme vers une limite, notée Πf. Nous avons $\|\Pi f\|_p = \alpha$ et l'inégalité (8.3) montre qu'une telle fonction h vérifiant $\|h\|_p = \alpha$ est unique. Il est clair que Πf appartient au sous-espace des vecteurs T-invariants de B et que Π est un projecteur sur ce sous-espace, projecteur positif si T est un opérateur positif. □

8.2 Contractions positives de L^1

Soit $(E, \mathcal{E}, \mu)$ un espace mesuré, μ étant une mesure positive finie ou σ-finie. Nous désignons par $\mathcal{C}^+$ le cône des fonctions mesurables définies sur $(E, \mathcal{E})$ à valeurs dans $([0, +\infty]$ et par $\overline{\mathcal{C}^+}$ le cône des classes modulo μ de fonctions de $\mathcal{C}^+$.

Si N est un noyau positif sur $(E, \mathcal{E})$, alors N opère sur le cône $\mathcal{C}^+$ et sur le cône des mesures positives sur $(E, \mathcal{E})$. Lorsque la mesure positive μN est absolument continue par rapport à μ, le noyau N opère sur le cône $\overline{\mathcal{C}^+}$.

Pour $1 \leq p < \infty$, nous désignons par $L^p(E, \mathcal{E}, \mu)$, ou plus simplement par $L^p(\mu)$, l'espace des (classes d'équivalence de) fonctions f de E dans $\mathbb{C}$ qui sont $\mathcal{E}$-mesurables et telles que $\|f\|_p = (\int |f|^p \mu(dx))^{\frac{1}{p}} < \infty$. Cet espace est muni de la norme $\| \ \|_p$.

Considérons un opérateur linéaire T sur le cône $\overline{\mathcal{C}^+}$. Pour tout $p \in [1, +\infty]$, notons $L^p_+(E, \mathcal{E}, \mu)$ le cône $L^p(E, \mathcal{E}, \mu) \cap \overline{\mathcal{C}^+}$. On obtient facilement le lemme suivant :

Lemme 8.2.1 *Toute contraction de $L^p_+(E, \mathcal{E}, \mu)$ se prolonge de façon unique en une contraction de $L^p(E, \mathcal{E}, \mu)$.*

Considérons maintenant un opérateur linéaire T positif de $L^1(E, \mathcal{E}, \mu)$ dans lui-même, qui est une contraction pour la norme $\| \ \|_1$ et la norme $\| \ \|_\infty$. On suppose donc que T vérifie les conditions

$$f \geq 0 \ \Rightarrow \ Tf \geq 0, \tag{8.4}$$

$$\int_E |Tf| \, d\mu \ \leq \ \int_E |f| \, d\mu, \tag{8.5}$$

$$|f| \leq M \ \Rightarrow \ |Tf| \leq M. \tag{8.6}$$

Notons qu'une classe importante d'opérateurs vérifiant ces conditions est celle des opérateurs de composition par une transformation préservant une mesure, étudiée au chapitre suivant.

Dans le cas général, la condition (8.6) est vérifiée par tout opérateur positif T opérant sur les fonctions $\mathcal{E}$-mesurables tel que $T1 \leq 1$. En fait nous allons voir que T peut être étendu en une contraction de L^∞ et (résultat d'interpolation) de tous les espaces L^p, $1 \leq p < \infty$.

Proposition 8.2.2 *Si T est une contraction de $L^1(\mu)$ vérifiant (8.6), alors, pour toute fonction $f \in L^1(\mu) \cap L^\infty(\mu)$, Tf appartient à $L^p(\mu)$, pour tout $p \in [1, \infty[$, et on a :*

$$\int_E |Tf|^p \, d\mu \leq \int_E |f|^p \, d\mu. \tag{8.7}$$

L'opérateur T se prolonge en une contraction de chacun des espaces $L^p(\mu)$, $1 \leq p \leq \infty$.

Preuve : Soit $p \in]1, \infty[$. Soit $f \in L^1(\mu) \cap L^\infty(\mu)$, condition qui assure que f est dans $L^p(\mu)$. L'opérateur T étant positif, nous pouvons supposer $f \geq 0$.

Pour tout réel positif α, nous avons

$$(Tf - \alpha)^+ \leq T(f - \alpha)^+. \tag{8.8}$$

En intégrant cette inégalité, nous obtenons :

$$\int_E (Tf - \alpha) \ 1_{(Tf > \alpha)} \ d\mu \leq \int_E (f - \alpha) \ 1_{(f > \alpha)} \ d\mu. \tag{8.9}$$

En multipliant les deux membres de (8.9) par α^{p-2} et en intégrant en α, nous obtenons après application du théorème de Fubini :

$$\begin{aligned}
(\frac{1}{p-1} - \frac{1}{p}) \int_E (Tf)^p \, d\mu &= \int_0^\infty \alpha^{p-2} (\int_E (Tf - \alpha) \ 1_{(Tf > \alpha)} \ d\mu) d\alpha \\
&\leq \int_0^\infty \alpha^{p-2} (\int_E (f - \alpha) \ 1_{(f > \alpha)} \ d\mu) d\alpha \\
&= (\frac{1}{p-1} - \frac{1}{p}) \int_E f^p \, d\mu < +\infty.
\end{aligned}$$

Montrons comment on peut prolonger T, défini sur $L^1(\mu) \cap L^\infty(\mu)$, en une contraction de $L^\infty(\mu)$. Le raisonnement est analogue, à partir de l'inégalité (8.7) pour les espaces L^p, $1 < p < \infty$.

Soit f une fonction dans $L^\infty(\mu)$ telle que $0 \leq f \leq C$. Si (f_n) est une suite croissante de fonctions positives dans $L^1(\mu) \cap L^\infty(\mu)$, telle que $\lim_n f_n = f$, p.p., alors (Tf_n) est une suite croissante majoré par C. Posons $Tf = \lim_n Tf_n$ et montrons que, pour toute suite croissante (g_n) de fonctions positives dans $L^1(\mu) \cap L^\infty(\mu)$, telle que $\lim_n g_n = f$, p.p., on a : $\lim_n Tf_n = \lim Tg_n$.

Pour $m \geq 1$ fixé, soit, pour tout $n \geq 1$, $\phi_n = \inf(f_m, g_n)$. La suite (ϕ_n) est croissante, majorée par f_m et elle converge en norme $\| \ \|_1$ vers f_m. Comme T est une contraction dans L^1, il en résulte : $\lim T\phi_n = T \lim_n \phi_n = Tf_m$. On a donc $\lim_n Tg_n \geq \lim_n T\phi_n = Tf_m$; d'où $\lim_n Tg_n \geq \lim_m Tf_m$ et donc égalité des limites en permutant les deux suites.

□

En particulier, on voit que l'on peut remplacer la condition (8.6) par la condition $T1 \leq 1$.

Nous considérerons donc dans ce qui suit une *contraction positive T de $L^1(\mu)$ et $L^\infty(\mu)$* et son action sur les différents espaces L^p qui, d'après la proposition 8.2.2 est une contraction. Ceci implique la proposition ci-dessous (convergence en moyenne dans les espaces L^p $(1 < p < \infty)$). Nous établirons ensuite un théorème ergodique ponctuel.

Proposition 8.2.3 *Si T est une contraction positive de $L^1(\mu)$, $L^\infty(\mu)$, pour tout $p \in]1, \infty[$ et tout $f \in L^p$, la suite des moyennes ergodiques $(M_n f)_{n \geq 1}$ converge en norme $\| \ \|_p$ vers $\Pi f \in L^p$, où Π est un projecteur dans L^p.*

Si $\mu(E) < \infty$, pour tout $f \in L^1$, la suite les moyennes ergodiques $(M_n f)_{n \geq 1}$ converge en norme $\| \ \|_1$ vers une limite $\Pi f \in L^1$.

Preuve : Il reste à vérifier la dernière assertion. Si $\mu(E) < \infty$, pour tout $\varepsilon > 0$, il existe $g \in L^2(E, \mathcal{E}, \mu) \subset L^1(E, \mathcal{E}, \mu)$ telle que $\|f - g\|_1 < \varepsilon$. On a donc :

$$\begin{aligned}\|M_n f - M_m f\|_1 &\leq 2\varepsilon + \|M_n g - M_m g\|_1 \\ &\leq 2\varepsilon + \mu(E)^{\frac{1}{2}} \left(\|M_n g - M_m g\|_2\right)^{\frac{1}{2}} \leq 3\varepsilon,\end{aligned}$$

pour n et m assez grand. Ceci montre la convergence en norme $\| \ \|_1$ de la suite $(M_n f)_{n \geq 1}$.
□

Remarque 8.2.4

La deuxième assertion du corollaire est fausse en mesure infinie, comme le montre l'exemple simple de l'opérateur de translation par 1 sur $\ell^1(\mathbb{N})$ et de la fonction $f = 1_{\{0\}}$. Sous les hypothèses du corollaire, on peut cependant définir le projecteur Π sur les éléments de L^1.

En effet, pour $f \in L^1 \cap L^2$, nous avons $\lim_n \|\frac{1}{n} S_n f - \Pi f\|_2 = 0$ et donc, pour tout ensemble A de mesure finie, $\lim_n \int_A \frac{1}{n} S_n f \ d\mu = \int_A \Pi f \ d\mu$.

Si $f \geq 0$, nous avons $\int_A \frac{1}{n} S_n f \ d\mu \leq \int_E \frac{1}{n} S_n f \ d\mu = \int_E f \ d\mu$, d'où $\int_A \Pi f \ d\mu \leq \int_E f \ d\mu$, et donc, la mesure étant σ-finie, $\int_E \Pi f \ d\mu \leq \int_E f \ d\mu$.

Pour $f \in L^1 \cap L^2$, on a donc : $\int_E |\Pi f| \ d\mu \leq \int_E \Pi |f| \ d\mu \leq \int_E |f| \ d\mu$. On peut alors étendre par densité, l'opérateur Π à L^1 tout entier, avec $\|\Pi f\|_1 \leq \|f\|_1$, $\forall f \in L^1(\mu)$.

8.3 Théorème ergodique ponctuel pour une contraction positive de L^1, L^∞

Le théorème de Von Neuman et ses extensions concernent le comportement en norme des moyennes ergodiques. Le théorème de Birkhoff et ses extensions permettent de préciser le comportement *ponctuel* de ces moyennes ergodiques : il fournit une forme très générale de loi des grands nombres.

Soit T une *contraction positive de* $(L^1(\mu), L^\infty(\mu))$. Etant donnée une fonction f mesurable sur E, pour tout entier $n \geq 1$, nous notons :

$$S_n f(x) := \sum_{k=0}^{n-1} T^k f(x), \quad S_n^* f(x) := \sup_{1 \leq k \leq n} S_k f(x).$$

Lemme 8.3.1 *Si f^+ est μ-intégrable, pour tout entier $n \geq 1$, on a :*

$$\int_{\{S_n^* f \geq 0\}} f \ d\mu \geq 0, \tag{8.10}$$

et pour tout un réel positif α :

$$\mu(\{S_n^*(f - \alpha) \geq 0\}) \leq \frac{1}{\alpha} \int_{\{S_n^*(f-\alpha) \geq 0\}} f \ d\mu. \tag{8.11}$$

Preuve : La fonction $(S_n^* f)^+$, comme f^+, est μ-intégrable. L'inégalité $f + T S_n^* f \geq f + T S_k f = S_{k+1} f$, $\forall k = 1, \cdots, n$, implique :

$$f + T S_n^* f \geq \sup_{2 \leq k \leq n+1} S_k f,$$

d'où :

$$\forall n \geq 1,\ f + T(S_n^* f)^+ \geq S_{n+1}^* f \geq S_n^* f. \tag{8.12}$$

On a donc, en intégrant sur $\{S_n^* f \geq 0\}$, pour tout $n \geq 1$,

$$\begin{aligned}\int_{\{S_n^* f \geq 0\}} f\ d\mu &\geq \int_{\{S_n^* f \geq 0\}} S_n^* f\ d\mu - \int_{\{S_n^* f \geq 0\}} T(S_n^* f)^+\ d\mu \\ &\geq \int_E (S_n^* f)^+\ d\mu - \int_E T(S_n^* f)^+\ d\mu.\end{aligned}$$

Comme T est une contraction de L^1, la différence dans l'inégalité précédente est ≥ 0, ce qui implique (8.10).

Pour $\alpha \geq 0$, en observant que $(f-\alpha)^+ \leq f^+$ et que $T1 \leq 1$, nous pouvons appliquer ce qui précède à $f - \alpha$ pour obtenir

$$\int_{\{S_n^*(f-\alpha) \geq 0\}} (f - \alpha)\ d\mu \geq 0.$$

Notons que cette inégalité implique que, sous l'hypothèse f^+ intégrable, les ensembles $\{S_n^*(f-\alpha) \geq 0\}$ sont de mesure finie pour $\alpha > 0$. On peut alors écrire l'inégalité précédente sous la forme (8.11). $\square$

Posons $f^* := \sup_{n \geq 1} M_n |f|$. L'ensemble $\{f^* \geq \alpha\}$ est la réunion croissante des ensembles $\{S_n^*(|f| - \alpha) \geq 0\}$. Par passage à la limite pour n tendant vers l'infini dans l'inégalité (8.11) appliquée à $|f|$, nous déduisons du lemme *l'inégalité maximale* pour les moyennes ergodiques :

Lemme 8.3.2 *(Inégalité maximale) Pour toute fonction $f \in L^1(\mu)$ et tout $\alpha > 0$, nous avons :*

$$\mu(\{f^* \geq \alpha\}) \leq \frac{1}{\alpha} \int_{\{f^* \geq \alpha\}} |f|\ d\mu \leq \frac{1}{\alpha} \|f\|_1.$$

Lemme 8.3.3 *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré. Pour toute fonction mesurable g sur E et toute fonction borélienne positive u sur $\mathbb{R}^+$, nous avons :*

$$\begin{aligned}\int_{\mathbb{R}^+} \mu(\{x : |g(x)| > \alpha\})\ u(\alpha)\ d\alpha &= \int_{\mathbb{R}^+} \int_E 1_{\{x : |g(x)| > \alpha\}}\ u(\alpha)\ d\mu(x)\ d\alpha \\ &= \int_E (\int_0^{|g(x)|} u(\alpha)\ d\alpha)\ d\mu(x).\end{aligned}$$

Proposition 8.3.4 *Soient ϕ et ψ deux fonctions mesurables positives sur un espace mesuré $(E, \mathcal{E}, \mu)$, vérifiant, pour tout $\alpha > 0$ les inégalités :*

$$\mu(\{\psi > \alpha\}) < \infty, \tag{8.13}$$

$$\mu(\{\psi > \alpha\}) \leq \frac{1}{\alpha} \int_{\{\psi > \alpha\}} \phi\ d\mu. \tag{8.14}$$

Si la fonction ϕ est dans L^p, pour $1 < p < \infty$, la fonction ψ est également dans L^p et vérifie :

$$\int_E \psi^p\ d\mu \leq (\frac{p}{p-1})^p \int_E \phi^p\ d\mu. \tag{8.15}$$

Preuve : Supposons d'abord ψ dans L^p. D'après le lemme 8.3.3 appliqué avec $u(\alpha) = \alpha^{p-1}$, nous avons la majoration :

$$\begin{aligned} \int_E \Big(\int_0^\psi \alpha^{p-1}\, d\alpha\Big)\, d\mu &\leq \int_0^\infty \alpha^{p-2}\Big(\int_{\{\psi>\alpha\}} \phi\, d\mu\Big)\, d\alpha \\ &\leq \int \phi\, \Big(\int_0^\infty 1_{\{\psi>\alpha\}}\, \alpha^{p-2}\, d\alpha\Big)\, d\mu; \end{aligned}$$

D'où :

$$\begin{aligned} \frac{1}{p}\|\psi\|_p^p &= \frac{1}{p}\int \psi^p\, d\mu \leq \frac{1}{p-1}\int_E \phi\, \psi^{p-1}\, d\mu \\ &\leq \frac{1}{p-1}\|\phi\|_p\|\psi^{p-1}\|_q = \frac{1}{p-1}\|\phi\|_p\|\psi\|_p^{p/q}; \end{aligned}$$

c'est-à-dire :

$$\|\psi\|_p \leq \frac{p}{p-1}\|\phi\|_p.$$

Si ψ n'est pas dans L^p, on s'y ramène en tronquant la fonction. Soient ε et C deux réels > 0. Posons

$$\psi_\varepsilon^C = (\inf(\psi, C) - \varepsilon)^+.$$

La fonction ψ_ε^C est non nulle sur l'ensemble $\{\psi > \varepsilon\}$. D'après l'hypothèse (8.13), cet ensemble est de mesure finie et donc ψ_ε^C qui est bornée est dans L^p. D'autre part, la fonction ψ_ε^C vérifie encore des inégalités du type précédent. En effet, pour $\alpha > C$, l'ensemble $\{\psi_\varepsilon^C > \alpha\}$ est vide, et pour $\alpha \leq C$, nous avons :

$$\begin{aligned} \mu(\{\psi_\varepsilon^C > \alpha\}) &= \mu(\{\psi > \alpha + \varepsilon\} \leq \frac{1}{\alpha+\varepsilon}\int_{\{\psi>\alpha+\varepsilon\}} \phi\, d\mu \\ &\leq \frac{1}{\alpha}\int_{\{\psi_\varepsilon^C>\alpha\}} \phi\, d\mu. \end{aligned}$$

Il en résulte que ψ_ε^C est dans L^p et vérifie :

$$\int_E |\psi_\varepsilon^C|^p\, d\mu \leq \Big(\frac{p}{p-1}\Big)^p \int_E |\phi|^p\, d\mu.$$

La conclusion, pour ψ, s'obtient en faisant tendre ε vers 0 et C vers $+\infty$. $\square$

En appliquant le lemme 8.3.2 et la proposition 8.3.4 à f et f^*, on obtient :

Proposition 8.3.5 *Soit $p > 1$. Si f est dans L^p, la fonction maximale f^* est dans L^p et vérifie :*

$$\int (f^*)^p\, d\mu \leq \Big(\frac{p}{p-1}\Big)^p \int |f|^p\, d\mu.$$

Rappelons qu'on a noté Π le projecteur, défini pour chaque $L^p(\mu)$, $1 \leq p > \infty$, sur les sous-espace des fonctions T invariantes. Le lemme maximal est la clef du théorème ergodique ponctuel donné ici pour un opérateur ? ? (Hopf, Dunford-Schwartz).

Théorème 8.3.6 *(théorème ergodique) : Soit T une contraction positive de $L^1(\mu)$, $L^\infty(\mu)$. Pour tout $p \in [1, \infty[$, pour toute fonction f de $L^p(\mu)$, la suite $(M_n f)_{n\geq 1}$ converge μ-presque sûrement vers Πf.*

Pour $1 < p < \infty$, pour toute fonction f de $L^p(\mu)$, la convergence a lieu également en norme $\| \|_p$. Si la mesure μ est finie, pour toute fonction f de $L^1(\mu)$, la convergence a lieu en norme $\| \|_1$.

Preuve : Pour une fonction f mesurable, notons $ls(f) = \limsup_n M_n f$. Nous avons $|ls(f)| \leq ls(|f|) \leq f^*$ et donc, d'après la proposition 8.3.5, si f est dans L^p, $ls(f)$ est aussi dans L^p et vérifie :

$$\int |ls(f)|^p \, d\mu \leq (\frac{p}{p-1})^p \int |f|^p \, d\mu. \tag{8.16}$$

Soit $f \in L^p$. En remplaçant f par $f - \Pi f$, nous nous ramenons au cas où $\Pi f = 0$. Nous devons montrer que $ls(f) = 0$ p.p.. La relation

$$M_n(M_m f) - M_n f = \frac{1}{n}[T^n \frac{1}{m} \sum_0^{m-1} S_j f - \frac{1}{m} \sum_0^{m-1} S_j f],$$

montre que $ls(f - M_m f) = 0$, p.p.. Pour f dans L^p, $p > 1$, nous avons la convergence en norme $\| \|_p$ de la suite $(M_n f)$ vers $\Pi f = 0$. D'après (8.16) nous avons donc :

$$\int |ls(f)|^p \, d\mu = \int |ls(M_m f)|^p \, d\mu \leq (\frac{p}{p-1})^p \int |M_m f|^p \, d\mu.$$

Comme $\lim_m \int |M_m f|^p \, d\mu = 0$, il en résulte que $ls(f)$ est nulle p.p..

Soit maintenant f dans L^1. Nous pouvons utiliser la projection Πf définie dans la remarque 8.2.4 et supposer que $\Pi f = 0$. Soit $\varepsilon > 0$. Il existe $g \in L^1 \cap L^2$ tels que $\|f - g\|_1 \leq \varepsilon$. On a donc aussi $\Pi(f - g)\|_1 < \varepsilon$ et d'autre part, d'après ce qui précède, $ls(g - \Pi g) = 0$, p.p..

D'après le lemme maximal 8.3.2, nous avons, pour tout $\alpha > 0$:

$$\begin{aligned} \mu[\{|ls(f - \Pi f)| > \alpha\}] &\leq \mu[\{|ls(f - g)| > \alpha/2\}] + \mu[\{|\Pi(g - f)| > \alpha/2\}] \\ &\leq \frac{2}{\alpha} \int |f - g| \, d\mu + \frac{2}{\alpha} \int |\Pi(g - f)| \, d\mu \leq \frac{4\varepsilon}{\alpha}. \end{aligned}$$

Comme ε et α sont arbitraires, $ls(f)$ est nulle p.p.. □

8.4 Systèmes dynamiques mesurés

Dans toute la suite nous désignons par $(E, \mathcal{E}, \mu, \tau)$ un **système dynamique** (au sens de la mesure), c'est-à-dire la donnée d'un espace E muni d'une tribu $\mathcal{E}$, d'une mesure positive σ-finie μ sur $\mathcal{E}$ et d'une transformation mesurable et surjective τ de E dans lui-même, laissant μ **invariante** :

$$\mu(\tau^{-1} A) = \mu(A), \ \forall A \in \mathcal{E}.$$

Si la transformation τ est bijective et si τ^{-1} est mesurable, on dira que le système est **inversible**.

L'espace E est souvent un espace métrique et la tribu $\mathcal{E}$ la tribu borélienne de E ou une sous-tribu de cette tribu.

• **Opérateurs associés à une transformation ponctuelle**

L'étude des propriétés des itérées d'une transformation τ conduit à introduire l'opérateur de composition par τ dans divers espaces fonctionnels.

On note T l'opérateur défini sur l'espace des fonctions mesurables sur E, à valeurs réelles ou complexes, par *composition* par la transformation τ :

$$f \to Tf, \text{ avec } Tf(x) = f(\tau x),\ x \in E.$$

D'après l'invariance de la mesure μ, la restriction de l'opérateur T à chaque espace $L^p(E, \mathcal{E}, \mu)$, $p \geq 1$, est une isométrie, en général non surjective.

Rappelons la notation des sommes ergodiques et des moyennes ergodiques :

$$S_n f(x) = \sum_{k=0}^{n-1} T^k f(x),\ M_n f(x) = \frac{1}{n}\ S_n f(x), x \in E,\ n \geq 1.$$

Si $\mathcal{F}$ est une sous-tribu de $\mathcal{E}$, pour $1 \leq p < \infty$, on désigne par $L^p(E, \mathcal{F}, \mu)$ le sous-espace de $L^p(\mu)$ des (classes modulo μ de) fonctions f qui sont $\mathcal{F}$-mesurables. Toute fonction f, qui est $\tau^{-1}\mathcal{F}$-mesurable, s'écrit $g \circ \tau$ pour une fonction $\mathcal{F}$-mesurable g. De plus nous avons, pour $f \in L^1(\mu)$:

$$\mathbb{E}[f \circ \tau | \tau^{-1}\mathcal{F}] = \mathbb{E}[f|\mathcal{F}] \circ \tau. \tag{8.17}$$

Une sous-tribu $\mathcal{F}$ de $\mathcal{E}$ est dite **invariante** (globalement) par τ si $\tau^{-1}\mathcal{F} \subset \mathcal{F}$. La transformation τ est dans ce cas une application mesurable de l'espace mesurable $(E, \mathcal{F})$ dans lui-même ; la tribu $\mathcal{E}$ est elle-même invariante, d'après l'hypothèse de mesurabilité de τ.

Si $\mathcal{F}$ est τ-invariante, L'opérateur T applique $L^p(\mathcal{F}, \mu)$ dans lui-même. La restriction de T à ces sous-espaces est une isométrie, qui n'est surjective que si, pour tout $A \in \mathcal{F}$, il existe $B \in \mathcal{F}$ tel que $\tau^{-1}B = A$, μ-p.s. ; les tribus complétées de $\mathcal{F}$ et $\tau^{-1}\mathcal{F}$ par les ensembles μ-négligeables coïncident alors.

Un rôle important est également joué par l'opérateur P **adjoint** de T dans les différents espaces $L^p(\mu)$, en particulier dans $L^2(\mu)$ et dans $L^1(\mu)$. Rappelons qu'il est défini, pour $1 \leq p < \infty$, par :

$$\int_E Pf\ \overline{g}\ d\mu = \int_E f\ T\overline{g}\ d\mu,\ \ f \in L^p(\mu),\ g \in L^q(\mu), \tag{8.18}$$

où $q \in]1, +\infty]$ vérifie $\frac{1}{p} + \frac{1}{q} = 1$.

Si nécessaire nous préciserons le lien entre P et μ en disant que P est l'**opérateur μ-adjoint** de T. L'invariance de μ par τ se traduit par $P1 = 1$.

L'adjoint P est un opérateur positif : $f \geq 0 \to Pf \geq 0$. Comme $\int Pf\, \mu = \int fT1\, d\mu = \int f\, d\mu$, on a :

$$\int |Pf|\, d\mu \leq \int P|f|\, d\mu = \int |f|\, d\mu \tag{8.19}$$

et P est une contraction de $L^1(\mu)$.

Ceci montre que P est une contraction de $L^p(\mu)$, pour $1 \leq p \leq \infty$.

Si le système dynamique $(E, \mathcal{E}, \mu, \tau)$ est inversible, l'opérateur P est simplement l'opérateur de composition par la transformation inverse τ^{-1}. Dans le cas général, l'opérateur P adjoint de T est un **inverse à gauche** de T :

$$PTf = f, \forall f \in L^p(\mu). \tag{8.20}$$

Plus généralement, pour tout u dans $L^\infty(\mu)$, on a

$$P(uTf) = P(u)f, \ \forall f \in L^p(\mu). \tag{8.21}$$

En effet, on a, $\forall f \in L^p(\mu), \forall g \in L^q(\mu)$,

$$\begin{aligned}
\int P(uTf)\, \overline{g}\, d\mu &= \int u\, Tf\, T\overline{g}\, d\mu \\
&= \int u\, T(f\overline{g})\, d\mu \\
&= \int Pu\, f\overline{g}\, d\mu.
\end{aligned}$$

Par contre P n'est un inverse à droite que si l'opérateur T est surjectif sur $L^p(\mu)$, ce qui correspond pratiquement au cas où la transformation τ possède un inverse τ^{-1} mesurable sur E.

En effet, pour f dans $L^p(\mu)$, $f - TPf$ est orthogonale à l'image de $L^q(\mu)$ par T, d'après les relations :

$$\begin{aligned}
\int TPf\, T\overline{g}\, d\mu &= \int Pf\, \overline{g}\, d\mu \\
&= \int f\, T\overline{g}\, d\mu, \forall f \in L^p(\mu), \forall g \in L^q(\mu).
\end{aligned}$$

Ceci montre que la restriction de l'opérateur TP à $L^2(\mu)$ est le projecteur orthogonal sur $L^2(E, \tau^{-1}\mathcal{E}, \mu)$.

Lorsque μ est une mesure de probabilité, l'opérateur TP est l'espérance conditionnelle par rapport à la sous-tribu $\tau^{-1}\mathcal{E}$ et on peut écrire formellement :

$$Pf = \mathbb{E}[f|\tau^{-1}\mathcal{E}] \circ \tau^{-1}. \tag{8.22}$$

• **Facteurs, isomorphisme, extension naturelle**

Définitions 8.4.1

A la structure de système dynamique est attachée une notion de facteur et d'isomorphisme (ou de conjugaison) entre systèmes dynamiques.

Un système $(E', \mathcal{E}', \mu', \tau')$ est un **facteur** du système $(E, \mathcal{E}, \mu, \tau)$ s'il existe une surjection mesurable φ d'un sous-ensemble mesurable de μ-mesure pleine E_0 de E sur un sous-ensemble E'_0 de E', qui applique μ sur μ' et vérifie la relation

$$\varphi \circ \tau = \tau' \circ \varphi.$$

On a alors le diagramme commutatif :

$$\begin{array}{ccc}
(E, \mathcal{E}, \mu) & \xrightarrow{\tau} & (E, \mathcal{E}, \mu) \\
\varphi \big\downarrow & & \varphi \big\downarrow \\
(E', \mathcal{E}', \mu') & \xrightarrow{\tau'} & (E', \mathcal{E}', \mu').
\end{array}$$

Si l'application φ est une bijection de E_0 sur E'_0, mesurable ainsi que son inverse, on dit que les systèmes dynamiques $(E, \mathcal{E}, \mu, \tau)$ et $(E', \mathcal{E}', \mu', \tau')$ sont **isomorphes** ou **conjugués**.

Etant donné un facteur $(E', \mathcal{E}', \mu', \tau')$, la sous-tribu $\mathcal{F} = \varphi^{-1}\mathcal{E}'$ de $\mathcal{E}$ est invariante par τ. Inversement, si $\mathcal{F}$ est une sous-tribu invariante, on peut construire, sous des hypothèses très générales sur $(E, \mathcal{E}, \mu)$, (cf. Rohklin [Roh67]) un système dynamique $(E', \mathcal{E}', \mu', \tau')$ facteur du système $(E, \mathcal{E}, \mu, \tau)$ par une application φ, tel que la tribu $\mathcal{F}$ s'identifie à l'image réciproque de $\mathcal{E}'$ par φ.

Une situation que nous examinerons par la suite est celle où la transformation τ est en un sens "fortement non inversible", ce qui se traduit par des propriétés spectrales remarquables de l'opérateur adjoint.

De façon générale, il est utile de faire apparaître l'opérateur P dans l'étude des propriétés asymptotiques du système dynamique. En effet, cet opérateur, contrairement à l'opérateur T, appartient souvent à une classe d'opérateurs ayant de bonnes propriétés spectrales dans certains espaces fonctionnels. Dans d'autres questions, il est au contraire commode de supposer le système dynamique inversible. Nous allons voir que cela est toujours possible, au prix d'une extension.

Définition 8.4.2

Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique. Un système dynamique inversible $(\widetilde{E}, \widetilde{\mathcal{E}}, \widetilde{\mu}, \widetilde{\tau})$ est une **extension naturelle** de $(E, \mathcal{E}, \mu, \tau)$, si $(E, \mathcal{E}, \mu, \tau)$ en est un facteur et si tout système inversible dont $(E, \mathcal{E}, \mu, \tau)$ est facteur est un facteur de $(\widetilde{E}, \widetilde{\mathcal{E}}, \widetilde{\mu}, \widetilde{\tau})$.

L'extension naturelle est unique à isomorphisme près. Elle peut être souvent réalisée de façon concrète. Son existence, dans le cas général, est établie par la proposition suivante :

Proposition 8.4.3 *Tout système dynamique $(E, \mathcal{E}, \mu, \tau)$ de mesure finie est facteur d'un système dynamique $(\widetilde{E}, \widetilde{\mathcal{E}}, \widetilde{\mu}, \widetilde{\tau})$ inversible.*

Preuve : Donnons l'idée formelle de la preuve. On forme l'espace $\widetilde{E}$ des suites $\widetilde{x} = (x_0, x_1, \ldots)$ à valeurs dans E telles que $\tau x_n = x_{n-1}, \forall n \geq 1$. On munit cet espace de la tribu $\widetilde{\mathcal{E}}$ rendant mesurables les applications coordonnées $X_n(\widetilde{x}) = x_n$, $n \geq 1$, et de la mesure $\widetilde{\mu}$ canoniquement déduite de μ par :

$$\widetilde{\mu}(X_0 \in A_0, \ldots, X_n \in A_n) = \mu(A_n),$$

pour $A_1 = \tau^{-1}A_0, \ldots, A_n = \tau^{-1}A_{n-1}$.

La transformation $\widetilde{\tau}$ est définie par $\widetilde{\tau}(x_0, x_1, \ldots) = (\tau x_0, \tau x_1, \ldots)$. Le décalage $\theta : (x_0, x_1, \ldots) \to (x_1, x_2, \ldots)$ sur $\widetilde{E}$ est l'inverse de $\widetilde{\tau}$. L'application $\pi : \widetilde{x} = (x_n) \to x_0$ fait de $(E, \mathcal{E}, \mu, \tau)$ un facteur de $(\widetilde{E}, \widetilde{\mathcal{E}}, \widetilde{\mu}, \widetilde{\tau})$. □

• Tribu des invariants

On pose :

$$\mathcal{J} = \{A \in \mathcal{E} : \tau^{-1}A = A\} \text{ et } \widetilde{\mathcal{J}} = \{A \in \mathcal{E} : \tau^{-1}A = A,\ \mu - \text{p.s.}\}.$$

On définit ainsi deux sous-tribus de $\mathcal{E}$ appelées respectivement **tribu des événements τ-invariants** et **tribu des événements μ-presque sûrement τ-invariants**.

Une fonction f sur E est dite τ-invariante [resp. τ-sur-invariante, τ-sous-invariante] si $f \circ \tau = f$ [resp. $f \circ \tau \leq f$, $f \circ \tau \geq f$]. Une fonction f sur E est dite τ-invariante [resp. τ-sur-invariante, τ-sous-invariante] μ-presque sûrement si, μ-p.s., $f \circ \tau = f$ [resp. $f \circ \tau \leq f$, $f \circ \tau \geq f$].

On montrera en exercice le résultat suivant :

Lemme 8.4.4 *Une fonction f sur $(E, \mathcal{E})$ est $\mathcal{J}$-mesurable si et seulement si elle est τ-invariante. Une fonction f sur $(E, \mathcal{E})$ est $\widetilde{\mathcal{J}}$-mesurable si et seulement si elle est τ-invariante μ-presque sûrement.*

Lemme 8.4.5 *Soit f une fonction $\mathcal{E}$-mesurable sur E.*

1) Si f est τ-surinvariante ou τ-sous-invariante, μ-presque sûrement, alors elle est τ-invariante, μ-presque sûrement.

2) Si f est μ-presque sûrement τ-invariante, alors il existe une fonction τ-invariante g telle que $f = g$, μ-p.s.. En particulier, pour tout $A \in \widetilde{\mathcal{J}}$, il existe $B \in \mathcal{J}$ tel que $B = A$, μ-p.s.

Preuve : 1) Supposons que, μ-p.s., $f \circ \tau \leq f$. Si f est μ-intégrable alors de l'invariance de μ par τ il résulte que, μ-p.s., $f \circ \tau = f$. Dans le cas général, on remarque que pour tous réels a, b, $a < b$, la fonction $\sup\{\inf\{f, b\}, a\}$ est encore τ-surinvariante, μ-p.s.. La première assertion du lemme est alors claire.

2) Soit f une fonction $\mathcal{E}$-mesurable et μ-presque sûrement τ-invariante.

- Si f est bornée, la fonction g définie par

$$g = \liminf_{n \to +\infty} \left(\frac{f + f \circ \tau + \ldots + f \circ \tau^{n-1}}{n}\right),$$

répond à la question.

- Si f est positive, il existe une suite de fonctions étagées $(f_n)_{n \geq 0}$ qui croît vers f. Chaque fonction étagée f_n est bornée ; il existe donc une fonction g_n telle que $g_n \circ \tau = g_n$ et $g_n = f_n$, μ-p.s.. On obtient alors le résultat voulu en posant $g = \liminf_n g_n$.

- Dans le cas général, on applique le résultat précédent aux fonctions positives f^+ et f^-.

Enfin soit $A \in \mathcal{E}$ tel que $\tau^{-1}(A) = A$, μ-p.s.. Posons

$$g = \liminf_{n \to +\infty} \left(\frac{1_A + 1_A \circ \tau + \ldots + 1_A \circ \tau^{n-1}}{n}\right).$$

L'ensemble $B = \{g = 1\}$ vérifie alors les propriétés voulues. □

• Ergodicité d'un système dynamique

Les sytèmes dynamiques "irréductibles" sont ceux pour lesquels l'espace ne peut pas être "découpé", au sens de la mesure, en deux sous-ensembles mesurables invariants par la transformation τ non triviaux pour la mesure μ. Cette notion d'irréductibilité est précisée dans la définition suivante.

Définition 8.4.6

Un système dynamique $(E, \mathcal{E}, \mu, \tau)$ est dit **ergodique** si, pour tout $A \in \mathcal{J}$, $\mu(A) = 0$ ou $\mu(A^c) = 0$.

L'ergodicité est donc satisfaite si, et seulement si, la tribu $\mathcal{J}$ est la tribu triviale, aux ensembles négligeables près.

On montre (exercice) qu'un système de *mesure finie* est ergodique si, et seulement si, toute fonction mesurable invariante par τ est égale à une constante μ-presque partout, ou encore, de façon équivalente, si l'opérateur T restreint à l'orthogonal des constantes dans $L^2(\mu)$ n'a pas de vecteur invariant.

8.5 Exemples de systèmes dynamiques mesurés

Donnons trois exemples simples de systèmes dynamiques.

Rotation

On considère le cercle unité $\mathbb{T}$ identifié à $\mathbb{R}/\mathbb{Z} = [0, 1[$, muni de sa tribu des boréliens $\mathcal{E}$ et de la mesure de Lebesgue normalisée notée λ.

Soit ρ_α la rotation d'angle α sur le cercle unité. La transformation laisse la mesure de Lebesgue λ invariante et elle est inversible. C'est de plus une isométrie de l'espace.

Supposons α irrationnel. Si f est une fonction de $L^2(\lambda)$ d'intégrale nulle et invariante par ρ_α, ses coefficients de Fourier sont nuls et donc f est nulle λ-presque partout. Le système $(\mathbb{R}/\mathbb{Z}, \rho_\alpha, \lambda)$ est donc ergodique.

Par un calcul analogue, on obtient que λ est la seule mesure de probabilité invariante par ρ_α. En effet, si μ est une mesure de probabilité invariante sur $\mathbb{T}$, ses coefficients de Fourier vérifient $\widehat{\mu}(n) = e^{2\pi i n\alpha}\widehat{\mu}(n)$, d'où $\widehat{\mu}(n) = 0, \forall n \neq 0$.

Multiplication par 2 mod 1

On considère le même espace mesuré $(\mathbb{T}, \mathcal{E}, \lambda)$. On désigne par τ la transformation $x \to 2x$ mod 1, opérant sur $\mathbb{T}$. Toute fonction sur $\mathbb{T}$ s'identifie à une fonction sur $\mathbb{R}$ périodique et de période 1.

Montrons que le système dynamique $(\mathbb{T}, \mathcal{E}, \lambda, \tau)$ est ergodique. Soit f une fonction borélienne bornée τ-invariante. Pour tout réel x et tous entiers n et k, nous avons : $f(x) = f(2^n x) = f(2^n(x+k2^{-n})) = f(x + k2^{-n})$; ce qui montre que f admet tous les nombres dyadiques pour période.

La fonction φ sur $\mathbb{R}$ définie par $\varphi(t) = \int_0^1 |f(x+t)-f(x)|\,dx$ s'annule pour tout nombre dyadique et on sait qu'elle est continue (continuité de la translation dans $(L^1, \|\ \|_1)$; elle est donc nulle partout. En utilisant le théorème de Fubini, il vient alors

$$\int_0^1 |f(x) - \int_0^1 f(t)\,dt|\,dx \;=\; \int_0^1 |\int_0^1 (f(x) - f(x+t))\,dt\,|\,dx \leq \int_0^1 \varphi(t)\,dt \;=\; 0.$$

D'où le résultat.

Dans cet exemple, l'opérateur adjoint est défini par

$$Pf(x) = \frac{1}{2}\,[f(\frac{x}{2}) + f(\frac{x+1}{2})].$$

Pour illustrer le rôle de l'opérateur adjoint, donnons une autre preuve de l'ergodicité du système dynamique, basée sur l'itération de P. On a :

$$P^n f(x) = \frac{1}{2^n}\sum_{k=0}^{2^n-1} f(\frac{x}{2^n} + \frac{k}{2^n}).$$

Pour une fonction f continue, la suite $(P^n f)_{n\geq 1}$ converge uniformément vers $\lambda(f)$ (elle se comporte comme une suite de sommes de Riemann). Comme P est une contraction de $L^2(\lambda)$, de la densité des fonctions continues dans $L^2(\lambda)$, il résulte que la suite $(P^n f)_{n\geq 1}$ converge en norme L^2 vers $\lambda(f)$, pour tout $f \in L^2(\lambda)$. En notant $\langle f, g\rangle$ le produit scalaire $\int f\overline{g}\,dx$ de deux fonctions de $L^2(\lambda)$, on a donc, pour toute fonction $f \in L^2(\lambda)$ d'intégrale nulle :

$$\lim_n |\langle T^n f, f\rangle| = \lim_n |\langle f, P^n f\rangle| = 0.$$

Cette relation montre que toute fonction invariante dans $L^2(\lambda)$, d'intégrale nulle, est nulle. Le système est donc ergodique.

Remarques : 1) Nous avons en fait montré au passage une condition de mélange plus forte que l'ergodicité.

2) Notons que pour f lipschitzienne d'intégrale nulle, nous avons une "vitesse" de décroissance des corrélations :

$$|\langle T^n f, f\rangle| = |\langle f, P^n f\rangle| \leq C 2^{-n},$$

où C est une constante.

3) La transformation $x \to 2x \bmod 1$ est non inversible. Une réalisation de son extension naturelle est la transformation (dite du boulanger) sur le carré $[0,1[\times[0,1[$ muni de la mesure de Lebesgue définie par :

$$(x,y) \to \begin{cases} (2x, \frac{y}{2}), \text{ pour } 0 \leq x \leq \frac{1}{2}, \\ (2x-1, \frac{y}{2}+\frac{1}{2}), \text{ pour } \frac{1}{2} \leq x < 1. \end{cases}$$

On vérifiera en exercice que la transformation du boulanger est ergodique pour la mesure produit sur le carré.

bullet **Décalage et mesure produit, schéma de Bernoulli**

Schéma de Bernoulli unilatère

Considérons maintenant un système "symbolique" défini sur l'espace $\Omega^+ = I^{\mathbb{N}}$, l'espace produit (unilatère) d'une famille de $\mathbb{N}$ copies d'un alphabet I fini. Un point ω de Ω^+ est donc une suite $\omega = (\omega_0, \omega_1, \ldots)$ à valeurs dans l'alphabet I.

On note θ la transformation définie par *décalage* (ou *shift*) sur les coordonnées, $\theta(\omega)_n = \omega_{n+1}$, $n \geq 0$, soit :

$$\theta(\omega_0, \omega_1, \ldots) = (\omega_1, \omega_2, \ldots).$$

L'espace Ω^+ est muni de la topologie produit, qui peut être définie à l'aide de la distance $d(\omega, \omega') = \sum_{n=0}^{\infty} 2^{-n} \delta(\omega_n, \omega'_n)$, δ désignant le symbole de Kronecker. Nous notons $\mathcal{E}$ la tribu borélienne sur Ω^+.

Pour tout vecteur de probabilité $p = (p_1, \ldots, p_{|I|})$ sur I, on définit une mesure de probabilité produit $\mathbb{P}$ sur Ω^+ ("mesure de Bernoulli"), en posant

$$\mathbb{P}(X_0 = a_0, \ldots, X_n = a_n) = p_{a_0} \ldots p_{a_n},$$

et en prolongeant $\mathbb{P}$ à la tribu borélienne de Ω^+. Il est clair que $\mathbb{P}$ est invariante par l'action du décalage θ. L'opérateur adjoint de la composition par le décalage est donné par

$$Pf(\omega) = \sum_{a \in I} p(a) f(a\omega),$$

où, pour $\omega = (\omega_0, \omega_1, \ldots)$, $a\omega$ désigne la suite $(a, \omega_0, \omega_1, \ldots)$.

Il est facile de vérifier que, pour une telle mesure produit $\mathbb{P}$, le système $(\Omega^+, \mathcal{E}, \mathbb{P}, \theta)$ est ergodique. Une méthode consiste à prouver, d'abord pour les fonctions ne dépendant que d'un nombre fini de coordonnées, puis par densité pour toutes les fonctions de $L^2(\mathbb{P})$, la propriété de mélange (plus forte que l'ergodicité) :

$$\lim_{n\to\infty} \langle f \circ \theta^n, g\rangle = (\int f \, d\mathbb{P}) \, (\int g \, d\mathbb{P}), \ \forall f, g \in L^2(\mathbb{P}).$$

Schéma de Bernoulli bilatère

L'*extension naturelle* du système $(\Omega^+, \mathcal{E}, \mathbb{P}, \theta)$ est le système $(\Omega, \mathcal{E}, \mathbb{P}, \theta)$, où Ω est l'espace produit $I^{\mathbb{Z}}$ dont les éléments sont les suites bilatères $\omega = (\omega_n)_{n\in\mathbb{Z}}$, muni de la probabilité produit (notée encore $\mathbb{P}$) définie à l'aide du vecteur p et sur lequel opère le décalage θ, maintenant inversible.

8.6 Théorème ergodique pour les systèmes dynamiques

Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique. Les théorème ergodiques obtenus précédemment s'appliquent aux opérateurs de composition par τ. C'est en particulier le cas du lemme maximal. Nous redonnerons cependant une preuve des théorèmes limites liée aux propriétés particulières des opérateurs de composition. Un fait important, dans le cas des opérateurs de composition par τ est que la limite des moyennes ergodiques s'exprime à l'aide de la tribu des invariants.

Le paragraphe suivant abordera un théorème limite spécifique aux systèmes dynamiques, le théorème sous-additif de Kigman, qui est une extension importante du théorème de Birkhoff.

En appliquant le théorème de Von Neumann 8.1.2 à l'opérateur T de composition par τ opérant dans $L^2(\mu)$ on obtient, pour f dans $L^2(\mu)$:

$$\lim_{n\to\infty} \|M_n f - \Pi f\|_2 = 0,$$

où Π désigne le projecteur orthogonal sur le sous-espace des fonctions τ-invariantes.

Le projecteur Π est une contraction positive de $L^2(\mu)$. Lorsque μ est une mesure de probabilité, son prolongement à $L^1(\mu)$ coïncide avec l'espérance conditionnelle par rapport à la tribu des événements τ-invariants. On a donc $\Pi f = \mathbb{E}[f|\mathcal{J}]$.

Proposition 8.6.1 *Soit f une fonction de $L^1(\mu)$. Si μ est une mesure de probabilité, alors la suite des moyennes $(M_n f)_{n\geq 1}$ converge dans $L^1(\mu)$ vers $\mathbb{E}[f|\mathcal{J}]$.*

<u>Preuve</u> : Pour f dans $L^2(\mu)$, d'après le théorème de Von Neumann la convergence a lieu dans $L^2(\mu)$ et donc, a fortiori, dans $L^1(\mu)$.

Soit $f \in L^1(\mu)$. Comme $L^2(\mu)$ est dense dans $L^1(\mu)$, il existe une suite de fonctions $(g_p)_{p\geq 0}$ de $L^2(\mu)$ qui converge vers f en norme L^1. Nous avons alors :

$$\begin{aligned} \|M_n f - \mathbb{E}[f|\mathcal{J}]\|_1 &\leq \|M_n f - M_n g_p\|_1 + \|M_n g_p - \mathbb{E}[g_p|\mathcal{J}]\|_1 + \|\mathbb{E}[g_p - f|\mathcal{J}]\|_1 \\ &\leq 2\ \|f - g_p\|_1 + \|M_n g_p - \mathbb{E}[g_p|\mathcal{J}]\|_1\ ; \end{aligned}$$

et par suite $\limsup_n \|M_n f - \mathbb{E}[f|\mathcal{J}]\|_1 \leq 2\ \|f - g_p\|_1$. D'où le résultat, en faisant tendre p vers $+\infty$. □

Si le système $(E, \mathcal{E}, \mu, \tau)$ est ergodique, pour f dans $L^1(\mu)$, l'espérance conditionnelle $\mathbb{E}[f|\mathcal{J}]$ se réduit à l'intégrale $\int f\, d\mu$ et donc, dans ce cas, les moyennes ergodiques de f convergent en norme $\|\ \|_1$ vers $\mu(f)$.

• Théorèmes ergodiques ponctuels

Dans tout ce paragraphe nous supposons que $(E, \mathcal{E}, \mu, \tau)$ est un système dynamique avec μ mesure **de probabilité** invariante par τ.

Le théorème de Birkhoff permet de préciser le comportement **ponctuel** des moyennes ergodiques, en fournissant une forme générale de la loi des grands nombres, basée sur la stationnarité. Ce théorème est un cas particulier du théorème de Hopf démontré plus haut.

Rappelons d'abord l'inégalité maximale pour les moyennes ergodiques. Etant donnée une fonction f mesurable sur E, pour tout entier $n \geq 1$, posons $S_n^* f = \sup_{1 \leq k \leq n} S_k f$.

Posons $f^* = \sup_{n \geq 1} M_n |f|$. Le lemme maximal 8.3.2 s'applique aux moyennes ergodiques pour l'opérateur de composition par une transformation préservant la mesure :

Lemme 8.6.2 *Si f^+ est μ-intégrable, pour tout entier $n \geq 1$, on a :*

$$\int_{\{S_n^* f \geq 0\}} f \quad d\mu \geq 0. \tag{8.23}$$

Lemme 8.6.3 *(maximal) Pour toute fonction $f \in L^1(\mu)$ et tout $\alpha > 0$, nous avons :*

$$\mu(\{f^* \geq \alpha\}) \leq \frac{1}{\alpha} \int_{\{f^* \geq \alpha\}} |f| \; d\mu \leq \frac{1}{\alpha} \|f\|_1. \tag{8.24}$$

En particulier, il en résulte que $\sup_{n \geq 1} M_n f < +\infty$, μ-p.s..

On a donc $\overline{f} = \limsup_{n \to +\infty} M_n f < +\infty$, μ-p.s.. En changeant f en $-f$, on obtient que, μ-p.s.,

$$\inf_{n \geq 1} M_n f > -\infty \quad \text{et} \quad \underline{f} = \liminf_{n \to +\infty} M_n f > -\infty.$$

Le théorème ergodique a été obtenu plus haut pour le cas plus général d'une contraction positive de (L^1, L^∞). Redonnons une variante de la preuve dans le cas d'une transformation ponctuelle.

Théorème 8.6.4 *(Birkhoff) : Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique, avec $\mu(E) = 1$. Pour toute fonction f de $L^1(\mu)$, la suite $(M_n f)_{n \geq 1}$ converge, μ-presque sûrement et dans $L^1(\mu)$ vers $\mathbb{E}[f|\mathcal{J}]$.*

<u>Preuve</u> : Montrons que les fonctions $\underline{f}^+$ et $\overline{f}^-$ sont μ-intégrables. Nous avons

$$\liminf_n M_n f \leq \liminf_n |M_n f| \text{ et par suite } \underline{f}^+ \leq \liminf_n |M_n f|.$$

D'après le lemme de Fatou,

$$\int_E \liminf_n |M_n f| \; d\mu \leq \liminf_n \int_E |M_n f| \; d\mu \leq \|f\|_1 \;;$$

d'où l'intégrabilité de $\underline{f}^+$. Celle de $\overline{f}^-$ s'en déduit en changeant f en $-f$.

Soient A et ε deux réels > 0. Considérons la fonction $\varphi = \max(\underline{f}, -A) + \varepsilon$. Les fonctions $\overline{f}$, $\underline{f}$ et φ sont μ-p.s. invariantes par τ. En outre, la fonction φ est intégrable. Par définition de la limite inférieure, on a

$$\bigcup_{k \geq 1} \{S_k(\varphi - f) > 0\} = \bigcup_{k \geq 1} \{\varphi > M_k f\} = E.$$

Le lemme 8.6.2, appliqué à la fonction $\varphi - f$, implique alors :

$$\int_E f \; d\mu \leq \int_E \varphi \; d\mu.$$

En faisant tendre ε vers 0 puis A vers $+\infty$, il vient alors : $\int_E f\ d\mu \leq \int_E \underline{f}\ d\mu$.

En changeant, dans ce qui précède, f en $-f$, on obtient

$$\int_E \overline{f}\ d\mu \leq \int_E f\ d\mu \leq \int_E \underline{f}\ d\mu.$$

Il en résulte que $\overline{f}$ et $\underline{f}$ sont égales μ-presque partout.

La suite $(M_n f)_{n\geq 1}$ a donc une limite $\widetilde{f}$ qui est intégrable, de même intégrale que f. Montrons que la convergence des moyennes ergodiques a lieu au sens de la norme $\|\ \|_1$. On peut supposer f positive. D'après le théorème de convergence dominée, la suite $(\min(M_n f, \widetilde{f}))_{n\geq 1}$ converge vers $\widetilde{f}$ dans $L^1(\mu)$. En intégrant la relation

$$|M_n f - \widetilde{f}| = M_n f + \widetilde{f} - 2\min(M_n f, \widetilde{f}),$$

on obtient :

$$\int |M_n f - \widetilde{f}|\ d\mu = 2\int (\widetilde{f} - \min(M_n f, \widetilde{f})\ d\mu \overset{n\to+\infty}{\longrightarrow} 0.$$

Il reste à montrer que $\widetilde{f} = \mathbb{E}[f|\mathcal{J}]$. Pour tout $A \in \mathcal{J}$, nous avons

$$\int_E f\ 1_A\ d\mu = \int_E M_n f\ 1_A\ d\mu,\ \forall n \geq 1,$$

et par suite $\int_E f\ 1_A\ d\mu = \int_E \widetilde{f}\ 1_A\ d\mu$; d'où le résultat.

□

Exemples

Sur le cercle $\mathbb{T}^1$ muni de la mesure de Lebesgue λ, pour toute fonction f λ-intégrable, on a, pour α irrationnel, la convergence au sens de la norme $\|\ \|_1$ et λ-p.p. :

$$\lim_n \frac{1}{n}\sum_0^{n-1} f(x + k\alpha) \to \int f\ d\lambda,$$

De même on a, pour tout entier $a > 1$, en norme $\|\ \|_1$ et λ-p.p.

$$\lim_n \frac{1}{n}\sum_0^{n-1} f(a^k x) \to \int f\ d\lambda.$$

Notons au passage le lemme suivant qui est souvent utile dans l'établissement de théorèmes limites.

Lemme 8.6.5 *Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique. Pour toute fonction f de $L^p(\mu)$, $p \geq 1$, la suite $(\dfrac{f\circ\tau^n}{n})_{n\geq 0}$ converge μ-p.s. vers zéro.*

Preuve : Nous avons, pour tout $\varepsilon > 0$, d'après la majoration $\sum_{n\geq 1} 1_{\{|f|>\varepsilon n\}} \leq \frac{|f|}{\varepsilon}$,

$$\begin{aligned}\sum_{n\geq 1} \mu[\{|f\circ\tau^n|^p > \varepsilon n\}] &= \sum_{n\geq 1} \mu[\{|f|^p > \varepsilon n\}] \\ &\leq \int_E \sum_{n\geq 1} 1_{\{|f|^p>\varepsilon n\}}\ d\mu \leq \int_E \frac{|f|^p}{\varepsilon}\ d\mu < +\infty.\end{aligned}$$

D'où le résultat. On note que le résultat est valable pour μ finie ou non. □

• **Théorèmes ergodiques pour semi-groupe à un paramètre**

Considérons $(\tau_t)_{t\in\mathbb{R}^+}$, un semi-groupe à un paramètre de transformations $x \to \tau_t x$ opérant sur un espace mesuré $(E, \mathcal{E}, \mu)$ et préservant la mesure μ.

Nous supposerons pour simplifier que E est un espace métrique et que le semi-groupe vérifie l'hypothèse de continuité suivante : $\lim_{t\to 0} \tau_t x = x$.

Notons T_t, $t \in \mathbb{R}$, les opérateurs de composition par τ_t et R_δ les opérateurs de moyenne définis par :

$$R_\delta f = \frac{1}{\delta}\int_0^\delta T_t f(x)\ dt, f \in L^1(\mu).$$

Soit $f \in L^1$. On montre facilement, en utilisant l'hypothèse de continuité, la convergence en norme L^1 :

$$\lim_{h\to 0} \|R_h f - f\|_1 = 0.$$

Nous allons montrer la convergence ponctuelle :

$$\lim_{\delta\to 0} R_\delta f(x) = f(x), \text{ pour presque tout } x \in \mathbb{R}. \tag{8.25}$$

Dans le cas particulier de l'action de $\mathbb{R}$ sur lui-même par translation, ce résultat est le théorème de dérivation de Lebesgue : si $f \in L^1(\mathbb{R})$, on a :

$$\lim_{h\to 0} \frac{1}{h}\int_0^h f(x+t)\ dt = f(x), \text{ pour presque tout } x \in \mathbb{R}.$$

En notant plus simplement $F_\delta = R_\delta f$, définissons les fonctions maximales

$$\begin{aligned} F_\delta^* &= \sup_{n\geq 1} \frac{1}{n\delta}\int_0^{n\delta} T_t f(x)\ dt = \sup_{n\geq 1}\frac{1}{n}\sum_{k=0}^{n-1} T_{k\delta}F_\delta(x), \\ F^* &= \sup_{\delta>0} R_\delta f = \sup_{\delta>0}\frac{1}{\delta}\int_0^\delta T_t f(x)\ dt. \end{aligned}$$

On note que la fonction maximale rend compte simultanément du comportement de l'opérateur de moyenne aux petites comme aux grandes échelles. Le lemme maximal qui suit montre que le sous-espace de L^1 des fonctions vérifiant (8.25) est fermé. En utilisant l'hypothèse de continuité, on en déduira en exercice que la convergence (8.25) est vérifiée pour toute fonction f de L^1.

Proposition 8.6.6 *(Inégalité maximale) La fonction maximale F^* vérifie l'inégalité :*

$$\mu(\{F^* > \alpha\}) \leq \frac{1}{\alpha}\int_{\{F^*>\alpha\}} f\ d\mu,\ \forall \alpha > 0. \tag{8.26}$$

Preuve : D'après le lemme maximal, la fonction F_δ^*, qui est une fonction maximale de moyennes ergodiques discrètes, vérifie :

$$\mu(\{F_\delta^* > \alpha\}) \leq \frac{1}{\alpha}\int_{\{F_\delta^*>\alpha\}} F_\delta\ d\mu,\ \forall \alpha > 0.$$

Considérons la suite définie, pour $k \geq 0$, par $\delta_k = \dfrac{1}{!k}$.

Les moyennes d'ordre n de $F_{\delta_{k-1}}$ sont alors des moyennes d'ordre nk de F_{δ_k}. Ceci implique que $F^*_{\delta_{k-1}} \leq F^*_{\delta_k}$ et donc, pour tout $\alpha > 0$, la suite des ensembles $\{F^*_{\delta_k} > \alpha\}$ est croissante avec k.

Il est clair que $F^* \geq F^*_{\delta_k}$, $\forall k \geq 1$. Pour tout $\delta > 0$, nous avons $\lim_k \delta_k [\frac{\delta}{\delta_k}] = \delta$. Par suite, en posant $n_k = [\frac{\delta}{\delta_k}]$, nous obtenons :

$$\lim_k \frac{1}{n_k \delta_k} \int_0^{n_k \delta_k} T_t f(x)\, dt = \frac{1}{\delta} \int_0^{\delta} T_t f(x)\, dt.$$

Il en résulte que

$$\{F^* > \alpha\} = \bigcup_k \{F^*_{\delta_k} > \alpha\}.$$

On en déduit le lemme maximal pour les opérateurs R_δ :

$$\begin{aligned} \mu(\{F^* > \alpha\}) &= \lim_{k \to \infty} \uparrow \mu(\{F^*_{\delta_k} > \alpha\}) \leq \frac{1}{\alpha} \lim_k \int_{\{F^*_{\delta_k} > \alpha\}} F_{\delta_k}\, d\mu \\ &\leq \frac{1}{\alpha} \lim_k \int_{\{F^* > \alpha\}} F_{\delta_k}\, d\mu \leq \frac{1}{\alpha} \int_{\{F^* > \alpha\}} f\, d\mu, \end{aligned}$$

car $F_{\delta h}$ converge en norme L^1 vers f.

□

8.7 Théorème ergodique sous-additif

Le théorème de Birkhoff s'applique à la suite des sommes ergodiques $(S_n f)_{n \geq 1}$, sommes qui, d'après leur définition, vérifient la propriété d'additivité :

$$S_{n+m} f(x) = S_n f(x) + S_m(\tau^n x), \forall n, m \in \mathbb{N}^*.$$

Le théorème de Kingman étend le théorème de Birkhoff aux suites de fonctions qui vérifient la propriété plus faible de sous-additivité suivante :

Définition 8.7.1

Etant donné un système dynamique $(E, \mathcal{E}, \mu, \tau)$, une suite $(F_n)_{n \geq 1}$ de fonctions mesurables sur E est dite τ-**sous-additive**, si elle vérifie, pour μ-presque tout $x \in E$, l'inégalité :

$$F_{n+m}(x) \leq F_n(x) + F_m(\tau^n x), \ \forall n, m \in \mathbb{N}^*. \tag{8.27}$$

Soit $(F_n)_{n \geq 1}$ une suite τ-sous-additive de fonctions telle que $\int_E F_1^+ d\mu < +\infty$.

Pour tout entier $n \geq 1$, l'inégalité $F_n \leq S_n F_1 \leq S_n F_1^+$ montre que $F_n^+ \leq S_n F_1^+$ et donc

$$\frac{1}{n} \int_E F_n^+ d\mu \leq \int_E F_1^+ d\mu < +\infty.$$

La suite des intégrales $u_n := \int_E F_n\, d\mu, n \geq 1$, à valeurs dans $[-\infty, +\infty[$ est sous-additive (i.e. vérifie $\forall n, p \in \mathbb{N}^*,\ u_{n+p} \leq u_n + u_p$). La suite des moyennes $(\frac{1}{n} \int_E F_n\, d\mu)_{n \geq 1}$ converge donc vers

$$\gamma = \inf_{p \geq 1} \frac{1}{p} \int F_p\, d\mu \in [-\infty, \infty[.$$

Théorème 8.7.2 *(Théorème ergodique sous-additif) (Kingman [King68]) Pour toute suite τ-sous-additive $(F_n)_{n\geq 1}$ de fonctions vérifiant $\int_E F_1^+ d\mu < +\infty$, la suite $(\frac{1}{n}F_n)_{n\geq 1}$ converge μ-presque partout, vers une limite $\widetilde{F}$, telle que*

$$\int_E \widetilde{F}^+ \, d\mu < +\infty \ \mathit{et} \ \int_E \widetilde{F} \, d\mu = \gamma = \lim_{n\to\infty} \frac{1}{n} \int_E F_n \, d\mu \in [-\infty, +\infty[.$$

Si le système dynamique est ergodique, alors ? ? $\widetilde{F} = \gamma = \lim_n \dfrac{F_n}{n}$.

Démontrons d'abord un lemme étendant aux suites τ-sous-additives une forme de lemme maximal.

Lemme 8.7.3 *Soit $(F_n)_{n\geq 1}$ une suite τ-sous-additive de fonctions vérifiant $\int_E F_1^+ d\mu < +\infty$. Supposons qu'il existe une application $\mathcal{E}$-mesurable m de E dans $\mathbb{N}^*$ vérifiant en tout point x, $F_{m(x)}(x) \leq 0$. Alors*

$$\limsup_n \frac{1}{n} F_n \leq 0 \quad \mathit{et} \quad \lim_n \frac{1}{n} \int F_n = \gamma \leq 0.$$

<u>Preuve</u> : Soit $x \in E$. Considérons la suite d'entiers naturels non nuls $(m_n(x))_{n\geq 1}$ définie par :

$$m_1(x) = m(x) \quad \text{et} \quad \forall n \geq 2,\ m_n(x) = m(\tau^{m_1(x)+\ldots+m_{n-1}(x)} x).$$

Nous avons :

$$F_{m(x)}(x) \leq 0 \quad \text{et} \quad \forall n \geq 2,\ F_{m_n(x)}(\tau^{m_1(x)+\ldots+m_{n-1}(x)} x) \leq 0.$$

Pour tout entier naturel $n > m_1(x)$, appelons $r = r(n, x)$ l'entier naturel non nul tel que

$$m_1(x) + \ldots + m_r(x) \leq n < m_1(x) + \ldots + m_{r+1}(x).$$

Nous avons alors en appliquant (8.27) :

$$\begin{aligned}
F_n(x) &\leq F_{m_1(x)+\ldots+m_r(x)} + F_{n-(m_1(x)+\ldots+m_r(x))}(\tau^{m_1(x)+\ldots+m_r(x)} x) \\
&\leq F_{n-(m_1(x)+\ldots+m_r(x))}(\tau^{m_1(x)+\ldots+m_r(x)} x) \\
&\leq F_1(\tau^{m_1(x)+\ldots+m_r(x)} x) + \ldots + F_1(\tau^{n-1} x) \\
&\leq F_1^+(\tau^{m_1(x)+\ldots+m_r(x)} x) + \ldots + F_1^+(\tau^{n-1} x) \\
&\leq S_n F_1^+(x) - S_{m_1(x)+\ldots+m_r(x)} F_1^+(x).
\end{aligned}$$

Supposons d'abord que m soit majorée par un entier L : $\forall x \in E,\ m(x) \leq L$. On a alors :

$$n - (m_1(x) + \ldots + m_r(x)) \leq m_{r+1}(x) \leq L,\ m_1(x) + \ldots + m_r(x) \geq n - L$$

et par suite

$$F_n(x) \leq S_n F_1^+(x) - S_{n-L} F_1^+(x).$$

D'après le théorème de Birkhoff, le second membre de cette inégalité, divisé par n, converge μ-presque partout et dans $L^1(\mu)$ vers zéro. Le résultat annoncé s'en déduit immédiatement.

Dans le cas général, considérons un entier $L > 1$ et soit $A_L = \{m > L\}$. Posons

$$\varphi_L = F_1^+ 1_{A_L}, \ \text{et} \ G_n = F_n - S_n \varphi_L.$$

A partir de la fonction m, nous définissons la fonction $\widetilde{m}$ par

$$\widetilde{m}(x) = \begin{cases} m(x) & \text{si } x \notin A_L \\ 1 & \text{sinon} \end{cases}.$$

Nous avons $G_{\widetilde{m}(x)} = F_{m(x)}(x)$, si $x \notin A_L$, $= F_1(x) - F_1^+(x)$, sinon.

La suite $(G_n)_{n\geq 1}$ est encore τ-sous-additive et, par construction, on a $\widetilde{m}(x) \leq L$ et $G_{\widetilde{m}(x)}(x) \leq 0$, pour tout $x \in E$. On en déduit que $\limsup_n \frac{G_n}{n} \leq 0$ et $\lim_n \int_E \frac{G_n}{n} d\mu \leq 0$ et par suite :

$$\limsup_n \frac{F_n}{n} \leq \limsup_n \frac{G_n}{n} + \lim_n \frac{S_n \varphi_L}{n} \leq \mathbb{E}[1_{A_L} F_1^+ | \mathcal{J}].$$

Du théorème de convergence dominée (pour les espérances conditionnelles) il résulte que $\mathbb{E}[1_{A_L} F_1^+ | \mathcal{J}]$ tend vers 0, quand L tend vers $+\infty$; d'où l'inégalité

$$\limsup_n \frac{F_n}{n} \leq 0.$$

En intégrant l'inégalité et en faisant tendre L vers l'infini, on obtient la majoration :

$$\lim_n \frac{1}{n} \int F_n \ \ d\mu \leq 0.$$

□

Preuve du théorème 8.7.2 Soient $\overline{F} = \limsup_{n\to+\infty} \frac{1}{n} F_n$, $\underline{F} = \liminf_{n\to+\infty} \frac{1}{n} F_n$.

D'après la sous-additivité, on a : $F_{n+1} \leq F_1 + T F_n$, ce qui implique que $\overline{F}$ et $\underline{F}$ sont τ-sous-invariantes et donc τ-invariantes, d'après le lemme 8.4.5. D'autre part des inégalités, $\forall n \geq 1,\ F_n \leq S_n F_1^+$ et du théorème de Birkhoff, il résulte que les fonctions $\underline{F}$ et $\overline{F}$ sont majorées par $\mathbb{E}[F_1^+ | \mathcal{J}]$. En particulier on a : $\int \overline{F} \ d\mu \leq \int F_1^+ \ d\mu < +\infty$

En raisonnant comme dans la preuve du théorème de Birkhoff, soient A un réel et $\varepsilon > 0$; posons $\varphi = \varphi_{A,\varepsilon} = \sup(\underline{F}, A) + \varepsilon$. La fonction φ est *invariante* et intégrable ($A + \varepsilon \leq \varphi \leq \mathbb{E}[F_1^+ | \mathcal{J}] + A + \varepsilon$).

La suite $H_n = F_n - n\varphi$ est τ-sous-additive et vérifie les hypothèses du lemme précédent 8.7.3. En effet, en observant que $\liminf_{n\to+\infty} \frac{H_n}{n} \leq -\varepsilon$, on prend comme fonction m du lemme $m := \min\{k \geq 1 : H_k \leq 0\}$. En appliquant ce lemme, on obtient : $\overline{F} \leq \varphi_{A,\varepsilon}$, d'où, en faisant tendre ε vers 0 et A vers $-\infty$, $\overline{F} \leq \underline{F}$.

Ceci montre la convergence presque partout de la suite $(\frac{F_n}{n})_{n\geq 1}$ vers une limite que nous noterons $\widetilde{F}$.

D'après le lemme 8.7.3, on a pour les intégrales l'inégalité :

$$\lim_n \frac{1}{n} \int_E F_n \ d\mu \leq \int_E \varphi_{A,\varepsilon} \ d\mu$$

et par suite, en faisant tendre ε vers 0 et A vers $-\infty$,

$$\lim_n \frac{1}{n} \int_E F_n \ d\mu \leq \int_E \widetilde{F} \ d\mu.$$

Il reste à montrer que cette inégalité est une égalité. De la sous-additivité, on déduit, pour n et r entiers ≥ 1 :

$$\frac{1}{nr} F_{nr} \leq \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{r} F_r(\tau^{kr} x).$$

En appliquant le théorème ergodique à τ^r et à la fonction $\frac{F_r}{r}$, on obtient que la fonction $\widetilde{F}$ est majorée par l'espérance de $\frac{F_r}{r}$ par rapport à la tribu des ensembles τ^r-invariants. En passant aux intégrales, on en déduit la majoration :

$$\int_E \widetilde{F}\, d\mu \leq \int \frac{F_r}{r}\, d\mu.$$

En faisant tendre r vers $+\infty$, on obtient alors le résultat recherché. $\square$

Exemple d'application.

Soient $(E, \mathcal{E}, \mu, \tau)$ un système dynamique de mesure finie et M une application $\mathcal{E}$-mesurable de E dans l'algèbre $\mathcal{M}(d, \mathbb{K})$ des matrices carrées d'ordre d à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Pour tout entier naturel n, considérons le produit

$$M_n(x) = M(x)M(\tau x)\ldots M(\tau^{n-1}x).$$

Posons $F_n(x) = \log \|M_n(x)\|$. La suite $(F_n)_{n\geq 1}$ est τ-sous-additive. Du théorème ergodique sous-additif, on déduit :

Proposition 8.7.4 *Sous l'hypothèse* $\int_E \log^+ \|M\|\, d\mu < +\infty$*, la suite* $(\dfrac{1}{n} \log \|M_n(x)\|)_{n\geq 1}$ *converge* μ*-presque partout vers une fonction* $\widetilde{F}$ *vérifiant*

$$\int_E \widetilde{F}^+\, d\mu < +\infty \quad \textit{et} \quad \int_E \widetilde{F}\, d\mu = \inf_{n\geq 1} \frac{1}{n} \int_E \log \|M_n\|\, d\mu.$$

Si le système est ergodique la suite $(\frac{1}{n} \log \|M_n(x)\|)_{n\geq 1}$ *converge* μ*-presque partout vers la constante* $\inf_{n\geq 1} \frac{1}{n} \int_E \log \|M_n\|\, d\mu \in [-\infty, +\infty[$.

8.8 Continuité de "type faible", principe de Banach

Nous terminons ce chapitre par deux résultats sur les inégalités de "type faible" telle que celle qui est donnée par le lemme maximal.

Un résultat d'interpolation

Soit $(E, \mathcal{E}, \mu)$ un espace mesuré. Soit $p \in [1, \infty[$. On dit qu'une application sous-linéaire T appliquant $L^p(\mu)$ dans les classes de fonctions mesurables est de type L^p-faible, s'il existe une constante K telle que :

$$m(\{x : |Tf(x)| > \alpha\}) \leq \frac{K}{\alpha^p} \int |f|^p\, dx,\ \forall \alpha > 0.$$

Il est clair, par l'inégalité de Tchébychev, que cette propriété est vérifiée si T est un opérateur borné sur $L^p(\mu)$, avec la constante $K = \|T\|_p^p$; mais elle est plus faible. La preuve du théorème ergodique (lemme maximal) fournit des exemples d'applications sous-linéaires qui sont de type L^1-faible.

Le lemme 8.3.3 permet de montrer de façon élémentaire le résultat d'interpolation suivant :

Proposition 8.8.1 *Soit $(E, \mathcal{E}, \mu)$ un espace mesuré. Soient p_1, p_2, $1 \leq p_1 < p_2$, et T un opérateur linéaire défini sur un sous-espace $\mathcal{V}$ de $L^{p_1}(\mu) \cap L^{p_2}(\mu)$ tel que si f est dans $\mathcal{V}$, pour tout $\alpha > 0$, f_α, définie par $f_\alpha := f$, si $|f| \leq \alpha, = 0$, si $|f| > \alpha$, est dans $\mathcal{V}$. Si tout f dans $\mathcal{V}$ vérifie :*

$$\mu(\{x : |Tf(x)| > \alpha\}) \leq \frac{C_1}{\alpha^{p_1}} \int |f|^{p_1}\, dx, \tag{8.28}$$

$$\mu(\{x : |Tf(x)| > \alpha\}) \leq \frac{C_2}{\alpha^{p_2}} \int |f|^{p_2}\, dx, \tag{8.29}$$

alors, pour tout $f \in \mathcal{V}$, pour tout $p \in]p_1, p_2[$, on a :

$$\int |Tf|^p\, d\mu(x) \leq 2^p \left(\frac{C_1 p}{p - p_1} + \frac{C_2 p}{p_2 - p}\right) \int |f|^p\, d\mu(x). \tag{8.30}$$

Preuve : En appliquant les inégalités (8.28) et (8.29) à $f^\alpha := f - f^\alpha$ et f_α, on obtient :

$$\begin{aligned} \mu(\{x : |Tf(x)| > 2\alpha\}) &\leq \mu(\{x : |Tf^\alpha(x)| > \alpha\}) + \mu(\{x : |Tf_\alpha(x)| > \alpha\}) \\ &\leq \frac{C_1}{\alpha^{p_1}} \int |f^\alpha|^{p_1}\, dx + \frac{C_2}{\alpha^{p_2}} \int |f_\alpha|^{p_2}\, dx. \end{aligned}$$

Appliquons maintenant le lemme 8.3.3 à $g = \frac{1}{2} Tf$ et à la fonction $u(\alpha) = \alpha^{p-1}$. En multipliant l'inégalité précédente par α^{p-1} en intégrant en α de 0 à ∞, puis en appliquant le théorème de Fubini, nous obtenons :

$$\begin{aligned} \int_E \left(\int_0^{|Tf(x)|/2} \alpha^{p-1}\, d\alpha\right) d\mu(x) &\leq C_1 \int_E |f|^{p_1} \left(\int_0^{|f|} \alpha^{p-p_1-1}\, d\alpha\right) d\mu(x) \\ &+ C_2 \int_E |f|^{p_2} \left(\int_{|f|}^\infty \alpha^{p_2-p-1}\, d\alpha\right) d\mu(x). \end{aligned}$$

Le résultat s'en déduit immédiatement.

□

Comme corollaire de la proposition, on obtient que, si T est un opérateur borné sur $L^2[0,1]$ et tel que, pour un treillis dense dans $L^1[0,1]$ de fonctions f de $L^1[0,1]$, Tf soit défini et vérifie la propriété de type faible L^1, alors, pour tout p, $1 < p < 2$, T se prolonge en un opérateur borné de $L^p[0,1]$ dans lui-même (cf. proposition 8.1.5).

Le principe de Banach

Dans plusieurs questions d'analyse (par exemple dans le théorème ergodique), étant donnés un espace mesuré $(E, \mathcal{E}, \mu)$ et une famille d'applications linéaires (A_n), on est conduit à étudier une application sous-linéaire A^* associant à toute fonction $f \in L^p(E, \mathcal{E}, \mu)$ une fonction mesurable $A^* f$ (fonction "maximale").

L'étude de ces applications passe souvent par l'établissement d'une inégalité de *type faible* du type suivant.

$$m(\{x : |A^* f(x)| > \alpha\}) \leq \frac{K}{\alpha^p} \int |f|^p\, dx, \ \forall \alpha > 0. \tag{8.31}$$

Le rôle de ces inégalités est mis en évidence à travers le "principe de Banach", qui est une application du *théorème de Baire*.

Théorème 8.8.2 *Soit $(E, \mathcal{E}, \mu)$ une espace mesuré. Soit (A_n) une suite d'applications linéaires définies sur l'espace de Banach $B = L^p(E, \mathcal{E}, \mu)$, pour $p \in [1, \infty[$, à valeurs dans les fonctions mesurables, telle que $A^*f(x) = \sup_n |A_n f(x)|$ soit finie μ-presque-partout. Alors, il existe une fonction $C(\alpha)$ définie pour $\alpha > 0$ vérifiant $\lim\limits_{\alpha \to 0} \downarrow C(\alpha) = 0$ telle que :*

$$m(\{x : |A^*f(x)| > \alpha \|f\|_p\}) \leq C(\alpha), \ \forall \alpha > 0.$$

<u>Preuve</u> : Notons $A_n^* f(x) = \sup_{1 \leq k \leq n} |A_k f(x)|$. Soit $\varepsilon > 0$. Pour tout $f \in B$, il existe n tel que $\mu(\{x : A^* f(x) > n\}) \leq \varepsilon$. On a :

$$B \ = \ \bigcup_{n=1}^{\infty} \{f : \mu(\{x : A^* f(x) > n\}) \leq \varepsilon\} = \bigcup_{n=1}^{\infty} \bigcap_{k=1}^{\infty} \{f : \mu(\{x : A_k^* f(x) > n\}) \leq \varepsilon\}.$$

Comme les applications A_k^* sont continues en mesure, les ensembles $\{f : \mu(\{x : A_k^* f(x) > n\}) \leq \varepsilon\}$ sont fermées dans B, donc aussi leur intersection. L'espace $B = L^p(E, \mathcal{E}, \mu)$ étant complet, le théorème de Baire implique que l'un des ensembles de la réunion précédente contient une boule : il existe $f_0 \in B$ et $\delta > 0$ tels que $\mu(\{x : A^* f(x) > n\}) \leq \varepsilon\}$, pour toute f vérifiant $\|f - f_0\| \leq \delta$.

On a donc $\mu(\{x : A^*(f_0 + \delta g)(x) > n\}) \leq \varepsilon\}$, $\forall g : \|g\|_p \geq 1$. L'inégalité de sous-additivité de A^* implique :

$$A^* g \leq \frac{1}{\delta} A^*(f + \delta g) + \frac{1}{\delta} A^* f_0;$$

d'où, pour toute fonction g telle que $\|g\|_p \leq 1$,

$$\begin{aligned} & \mu(\{x : A^*(g)(x) > \frac{2n}{\delta}\}) \\ \leq \ & \mu(\{x : A^*(f_0 + \delta g)(x) > n\}) + \mu(\{x : A^*(f_0 + \delta g)(x) > n\}) \leq 2\varepsilon\}. \end{aligned}$$

Posons

$$C(\alpha) := \sup_{\|g\|_p \leq 1} \mu(\{x : A^* g(x) > \alpha\}).$$

Il est clair que C est une fonction décroissante et d'après ce qui précède, on a : $\lim_{\alpha \to \infty} C(\alpha) = 0$. $\square$

Chapitre 9

Processus stationnaires

9.1 Stationnarité et représentation des processus stationnaires

Définition 9.1.1 Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Un processus $(X_n)_{n\in\mathbb{Z}}$ défini sur cet espace, à valeurs réelles, est *stationnaire (au sens strict)* si, pour tout $p \geq 1$, la loi du p-uple $(X_n, \ldots, X_{n+p-1})$ est indépendante de $n \in \mathbb{Z}$.

Pour tout $p \geq 1$ et pour toute fonction ϕ borélienne sur $\mathbb{R}^p$, telle que $\mathbb{E}[|\phi(X_0, ..., X_{p-1})|] < \infty$, l'espérance $\mathbb{E}[\phi(X_n, ..., X_{n+p})]$ est alors indépendante de n.

Systèmes dynamiques / Processus stationnaires

Etant donné un système dynamique $(E, \mathcal{E}, \mu, \tau)$, l'invariance de la mesure μ sous l'action de τ montre que, pour toute fonction f mesurable sur E, la suite $(X_n = f \circ \tau^n)_{n\geq 0}$ forme un processus stationnaire sur l'espace probabilisé $(E, \mathcal{E}, \mu)$.

Inversement, étant donné un processus $(X_n)_{n\in\mathbb{Z}}$ défini sur un espace probabilisé $(B, \mathcal{B}, \nu)$, vérifiant la condition de stationnarité (9.1.1), on peut construire un système dynamique (un espace produit muni du décalage θ et d'une mesure de probablité θ-invariante), permettant de représenter (X_n) sous la forme $(X_n) = (X_0 \circ \theta^n)$.

Pour le voir, désignons par $(\Omega^+, \mathcal{F})$ l'espace produit $\mathbb{R}^{\mathbb{N}}$ muni de la tribu borélienne et par $(Y_n)_{n\geq 0}$ les applications coordonnées de Ω^+. Considérons l'application mesurable η, de $(B, \mathcal{B})$ dans $(\Omega^+, \mathcal{F}^+)$, définie par $x \in B \to\ \eta(x) = (X_n(x))_{n\geq 0} \in \Omega^+$. On a $X_n = Y_n \circ \eta, \forall n \geq 0$. Appelons $\mathbb{P}$ l'image de ν par η. Du théorème des classes monotones, il résulte que la mesure de probabilité $\mathbb{P}$ est invariante par le décalage sur Ω^+. Nous obtenons ainsi un système dynamique $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \theta)$ et le processus des coordonnées $(Y_n) = (Y_0 \circ \theta^n)$ sous la probabilité $\mathbb{P}$ est équivalent à (X_n).

De cette façon, nous avons associé un système dynamique à tout processus stationnaire et nous l'avons réalisé comme le processus des applications coordonnées de Ω^+.

On peut ainsi utiliser les deux points de vue, "systèmes dynamiques" et "processus stationnaires". Notons que les terminologies employées peuvent être un peu différentes. Par exemple, pour les sytèmes dynamiques, il est question de convergence ponctuelle, alors que dans la terminologie des processus, on parle de convergence presque sûre.

Extension bilatère du décalage

Soit $(B, \mathcal{B})$ un espace mesurable. Considérons les espaces mesurables produits $(\Omega, \mathcal{F}) = \otimes^{\mathbb{Z}} (B, \mathcal{B})$ et $(\Omega^+, \mathcal{F}^+) = \otimes^{\mathbb{N}} (B, \mathcal{B})$. Soit $\mathbb{P}$ une mesure de probabilité sur $(\Omega^+, \mathcal{F}^+)$ invariante par le décalage

θ.

Pour tout $n \in \mathbb{Z}$, appelons Π_n la projection de Ω sur l'espace produit $\otimes^{[n,+\infty[} B$ identifié à Ω^+. Désignons par $\mathbb{Q}$ la mesure de probabilité sur Ω telle que $\forall n \in \mathbb{Z},\ \Pi_n\mathbb{Q} = \mathbb{P}$, construite à l'aide du théorème de prolongement de Kolmogorov. Du théorème des classes monotones, il résulte que la probabilité $\mathbb{Q}$ est invariante par le décalage θ.

Nous avons donc construit un système dynamique inversible *bilatère* $(\Omega, \mathcal{F}, \mathbb{Q}, \theta)$ qui est une réalisation de **l'extension naturelle** de $(\Omega^+, \mathcal{F}^+, \mathbb{P}, \theta)$.

Grâce à cette construction, tout processus stationnaire indexé par $\mathbb{N}$ peut être considéré comme la restriction à $\mathbb{N}$ d'un processus stationnaire indexé par $\mathbb{Z}$.

Définition 9.1.2 Un processus stationnaire $(X_n)_{n\geq 0}$ est dit **ergodique** si le décalage unilatère associé est ergodique.

On note que, si $(E, \mathcal{E}, \mu, \tau)$ est un système ergodique, pour toute fonction f, le processus $(f \circ \tau^n)$ est ergodique. En effet le décalage unilatère associé au processus $(f \circ \tau^n)$ est isomorphe à un facteur de $(E, \mathcal{E}, \mu, \tau)$.

Considérons un processus aléatoire stationnaire au sens strict. Soit ϕ borélienne sur $\mathbb{R}^p$, telle que $\mathbb{E}[|\phi(X_0, ..., X_{p-1})|] < \infty$. Le théorème ergodique assure la convergence p.s. et en norme $\|\|_1$ des moyennes ergodiques vers une limite :

$$\frac{1}{N}\sum_{n=0}^{N-1} \phi(X_n, ..., X_{n+p-1}) \to \widetilde{\phi}(\omega), \text{ avec } \mathbb{E}[\widetilde{\phi}] = \mathbb{E}[\phi(X_0, ..., X_{p-1})]. \tag{9.1}$$

L'ergodicité du processus est équivalente à la propriété que, pour tout ϕ, la limite $\widetilde{\phi}$ est une constante, qui est alors égale à $\mathbb{E}[\phi(X_0, \cdots, X_{p-1})]$. Ainsi, la loi des grands nombres est valide et les moyennes temporelles convergent vers l'espérance (moyenne spatiale dans les systèmes dynamiques réalisés sur un espace, comme c'est le cas en mécanique).

Dans le cadre des applications, l'ergodicité permet la construction d'estimateurs. Si l'on recherche une estimation de quantités globales attachées au processus telles que $\mathbb{E}[\phi(X_0, \cdots, X_{p-1})]$, d'après la convergence (9.1) vers $\mathbb{E}[\phi(X_0, \cdots, X_{p-1})]$, il y a "assez d'information" sous l'hypothèse d'ergodicité, dans (presque) toute trajectoire ω du processus pour permettre l'estimation d'une telle quantité à partir de l'échantillon observé $(X_0(\omega), \cdots, X_{n-1}(\omega), \cdots)$. Se pose alors la question des fluctuations dans la convergence et cette question est réglée par le théorème de la limite centrale quand ce théorème est valide.

? ? ?Notation : Pour un processus (X_n) indexé par $\mathbb{Z}$ on note $(\mathcal{F}_n^X)_{n\geq 0}$ la *filtration naturelle croissante* associée : $\mathcal{F}_n^X = \sigma(X_k, k \leq n)$ et $\mathcal{G}_n^X = \sigma(X_k, k \geq n))$ la *filtration naturelle décroissante*. Quand le processus est indexé par $\mathbb{N}$, on se limite aux indices ≥ 0.

Nous avons, pour tout $n \geq 0$,

$$\begin{aligned}
\mathbb{E}_\nu[X_{n+1}|\mathcal{F}_n^X] &= \mathbb{E}_{\mathbb{Q}}[Y_{n+1}|\mathcal{F}_n^Y] \circ \eta, \\
\mathbb{E}_\nu[X_n|\mathcal{G}_{n+1}^X] &= \mathbb{E}_{\mathbb{P}}[Y_n|\mathcal{G}_{n+1}^Y] \circ \eta, \\
\mathbb{E}_{\mathbb{Q}}[Y_{n+1}|\mathcal{F}_n^Y] &= \mathbb{E}_{\mathbb{Q}}[Y_0|\mathcal{F}_{-1}^Y] \circ \theta^{n+1}, \\
\mathbb{E}_{\mathbb{P}}[Y_n|\mathcal{G}_{n+1}^Y] &= \mathbb{E}_{\mathbb{P}}[Y_0|\mathcal{G}_1^Y] \circ \theta^n.
\end{aligned}$$

• Processus stationnaires du second ordre

L'étude spectrale des processus stationnaires porte sur les propriétés hilbertiennes de ces processus : dans ce cadre, on ne retient que la stationnarité "au second ordre". Ce point de vue est insuffisant dans les problèmes d'estimation, quand l'étude des lois limites statistiques est nécessaire, et on est alors conduit à se placer dans le cadre des processus strictement stationnaires (i.e. au sens de la définition 9.1.1). Notons que l'on peut cependant obtenir certains résultats ponctuels dans un cadre hilbertien (par exemple une loi forte des grands nombres dans le cas orthogonal).

Définition 9.1.3 Un **processus du second ordre** sur un espace $(\Omega, \mathcal{E}, \mathbb{P})$ (indexé par $\mathbb{Z}$) est une suite $X = (X_n)_{n\in\mathbb{Z}}$ de v.a. réelles ou complexes de carré intégrable. Il est dit **stationnaire au sens large** (on dit aussi stationnaire du second ordre), si l'espérance $\mathbb{E}[X_n]$ et les produits scalaires $\mathbb{E}[X_{n+p}\overline{X_n}]$, pour tout $p \in \mathbb{Z}$, sont indépendants de $n \in \mathbb{Z}$.

On note γ_X la fonction de **covariance** du processus $X = (X_n)_{n\in\mathbb{Z}}$ définie sur $\mathbb{Z}$ par

$$\gamma_X(n) := cov(X_n, X_0) = \mathbb{E}[X_n\overline{X_0}] - \mathbb{E}[X_n]\mathbb{E}[\overline{X_0}],\ n \in \mathbb{Z}.$$

En retranchant $\mathbb{E}(X_n)$, on se ramène à un processus *centré* ($\mathbb{E}(X_n) = 0, \forall n$), dont la covariance s'écrit simplement : $\gamma_X(n) = \mathbb{E}[X_n\overline{X_0}]$.

Exemples 9.1.4

a) On appelle *bruit blanc* centré de variance σ^2, tout processus stationnaire $W = (W_n)$ formé d'une suite de v.a. de carré intégrable deux à deux orthogonales, centrées et de même variance σ^2 :

$$\mathbb{E}[W_n\overline{W_m}] = 0,\ \text{si } n \neq m, = \sigma^2,\ \text{si } n = m.$$

Une suite de v.a. centrées i.i.d. de carré intégrable est un exemple de bruit blanc. L'orthogonalité est une propriété beaucoup plus faible que l'indépendance.

b) Soit $\{Y_k\}$ une famille de v.a. centrées deux à deux orthogonales et de variance $\sigma_k^2 = \mathbb{E}[|Y_k|^2]$ telles que $\sum_k \sigma_k^2 < \infty$, et soient u_k des fréquences fixées.

On définit un processus $(X_n)_{n\in\mathbb{Z}}$ en posant $X_n = \sum_k Y_k e^{2\pi i n u_k}$, $n \in \mathbb{Z}$. Ce processus est stationnaire de fonction de covariance

$$\gamma_X(n) = \sum_k \sigma_k^2 \exp(2\pi i n u_k), n \in \mathbb{Z}.$$

c) *Processus gaussien stationnaire* (à faire ? ?)

Rappelons qu'une v.a.r. Y est de loi gaussienne (ou normale) $\mathcal{N}(m, \sigma^2)$ si sa loi a une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$ de la forme :

$$\frac{1}{\sqrt{2\pi}}\exp(-\frac{(y-m)^2}{2\sigma^2}).$$

On a donc, pour une variable gaussienne,

$$\mathbb{P}(Y \in A) = \frac{1}{\sqrt{2\pi}}\int_A \exp(-\frac{(y-m)^2}{2\sigma^2})\ dy,\ \forall A \text{ borélien de } \mathbb{R}.$$

Sa fonction caractéristique est $\mathbb{E}[\exp(itY)] = \exp(itm)\exp(-\frac{1}{2}\sigma^2 t^2)$.

Une v.a. $X = (X_1, \cdots, X_r)$ à valeurs dans $\mathbb{R}^d$ suit une loi gausienne si sa loi est de la forme

$$\frac{1}{(2\pi\ det(C))^{d/2}}\exp(-(\frac{1}{2}\langle Cx, x\rangle),$$

où C une matrice $r \times r$ symétrique et définie positive.

On a alors pour tout borélien A de $\mathbb{R}^r$:

$$P((X_{t_1}, ..., X_{t_r} \in A) = \frac{1}{(2\pi\ det(C))^{d/2}} \int_A \exp(-(\frac{1}{2}\langle Cx, x\rangle)\ dx_1 \cdots dx_r.$$

et la fonction caractéristique de X est ? ? ?$\mathbb{E}[\exp(itY)] = \exp(itm)\exp(-\frac{1}{2}\sigma^2 t^2)$.

Pour toute matrice C $r \times r$ symétrique et définie positive, on peut construire...

On a donc, pour tout vecteur $u = (u_1, \cdots, u_p, 0, 0, \cdots)$ à support fini (i.e. de composantes u_k nulle pour $k > p$, $\langle X, u\rangle$ est une v.a. gaussienne centrée de variance $\langle \Gamma_X, u\rangle$ et de fonction caractéristique

$$\mathbb{E}(\exp(it\langle X, u\rangle) = \exp(-\frac{1}{2}t^2\langle \Gamma_X, u\rangle).$$

(? ?) Proposition : *Soit* $\Gamma = (\gamma(n, m), n, m \in I\!N)$ *une matrice infinie symétrique et définie positive. Il existe un processus gaussien de matrice de covariance* Γ.

Preuve :

□

Les processus stationnaires formés de v.a. au sens strict (i.e. au sens de la définition donnée en (9.1.1)) de carré intégrables sont des processus du second ordre stationnaires au sens large. La réciproque, évidemment fausse dans le cas général, est vérifiée par la classe des processus gaussiens.

Comme on l'a remarqué plus haut, en considérant l'extension naturelle du système, on peut se ramener au cas où le processus est défini sur $\mathbb{Z}$.

• Représentation vectorielle des processus du second ordre

Si T est un opérateur unitaire opérant sur un sous-espace de $L^2(\Omega, \mathcal{A}, \mathbb{P})$, pour tout $X_0 \in L^2(\Omega, \mathcal{A}, \mathbb{P})$, le processus $X = (X_n)$ défini par $X_n = T^n X_0, n \in \mathbb{Z}$, est un processus du second ordre stationnaire au sens large (si les X_n sont centrés) de covariance $\gamma_X(n) = \langle T^n X_0, X_0\rangle$, $n \in \mathbb{Z}$. Nous allons voir que tout processus stationnaire du second ordre peut être représenté sous cette forme.

Notons que, pour un processus $(T^n f)$ engendré par une fonction sous l'action d'un système dynamique, cet opérateur est la composition par la transformation du système. En particulier, pour un processus stationnaire représenté comme la suite des applications coordonnées d'un espace produit muni d'une mesure stationnaire, T est l'opérateur de composition par le décalage θ.

Notations : Soit $X = (X_n)_{n\in\mathbb{Z}}$ un processus de carré intégrable, stationnaire au sens large, défini sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$. On note $\mathcal{H}(X)$ le sous-espace vectoriel fermé engendré dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$ par les v.a. X_k :

$$\mathcal{H}(X) = \overline{sp}\{X_k, k \in \mathbb{Z}\}.$$

La stationnarité du processus permet de construire un opérateur T de décalage sur $\mathcal{H}(X)$.

Pour cela, on définit d'abord T sur le sous-espace engendré par les combinaisons linéaires finies des v.a. X_n par

$$T(\sum_k \alpha_k X_k) = \sum_k \alpha_k X_{k+1}. \tag{9.2}$$

La stationnarité du processus implique que T est une isométrie sur ce sous-espace :

$$\| \sum_k \alpha_k X_{k+1}\|^2 = \sum_{k,j} \alpha_k \overline{\alpha_j} \langle X_{k+1}, X_{j+1}\rangle = \| \sum_k \alpha_k X_k\|^2,$$

prolongeable en un opérateur isométrique surjectif (donc unitaire) de $\mathcal{H}(X)$ sur $\mathcal{H}(X)$. La suite (X_n) est ainsi représentée comme la suite des images de X_0 par les itérés de l'opérateur T : $X_n = T^n X_0$.

9.2 Mesure spectrale d'un processus stationnaire

Considérons un processus stationnaire X. Pour toute suite finie $\alpha_{-N}, ..., \alpha_N$ dans $\mathbb{C}$, la norme de la v.a. $\sum_{-N}^N \alpha_k X_k$ s'exprime à l'aide de la fonction de covariance γ_X du processus sous la forme :

$$\mathbb{E}|\sum \alpha_k X_k|^2 = \sum \alpha_k \overline{\alpha_j} \mathbb{E}(X_k \overline{X_j}) = \sum_{-N \leq k,j \leq N} \alpha_k \overline{\alpha_j} \gamma_X(k-j).$$

Ceci montre, en particulier, que la somme $\sum_{-N \leq k,j \leq N} \alpha_k \overline{\alpha_j} \gamma_X(k-j)$ est un réel ≥ 0. La fonction de covariance d'un processus stationnaire est donc une fonction sur $\mathbb{Z}$ (une suite) qui est définie positive au sens de la définition suivante :

Définition 9.2.1 Une suite à valeurs complexes $\gamma = (\gamma(n), n \in \mathbb{Z})$ est dite *définie positive* si, pour toute suite finie $\alpha_{-N}, ..., \alpha_N$ dans $\mathbb{C}$, la somme

$$\sum_{-N \leq k,j \leq N} \alpha_k \overline{\alpha_j} \gamma(k-j) \text{ est un réel } \geq 0.$$

Il existe une notion générale de fonction définie positive sur un groupe abélien et une caractérisation de ces fonctions à l'aide du groupe dual (théorème de Bochner). Le théorème suivant caractérise les fonction définies positives dans le cas particulier qui nous intéresse ici, le cas du groupe $\mathbb{Z}$.

Théorème 9.2.2 *(Bochner-Herglotz) Pour toute suite γ définie positive sur $\mathbb{Z}$, il existe une (unique) mesure positive μ sur le cercle telle que*

$$\gamma(n) = \int e^{2\pi i n u}\, d\mu(u), \forall n \in \mathbb{Z}.$$

Preuve : La mesure μ est telle que la suite de ses coefficients de Fourier est égale (au changement près de n en $-n$) à la suite γ. Elle est donc déterminée par la donnée de γ.

1) La condition est suffisante. En effet, la suite des coefficients de Fourier d'une mesure ≥ 0 est définie positive : on a, pour toute suite (a_n) à support fini de coefficients complexes,

$$\sum_{k,j} a_k \overline{a_j} \gamma(k-j) = \int |\sum a_k \exp(2\pi i k u)|^2\, d\mu(u) = \int |P(u)|^2\, d\mu(u),$$

où P est le polynome trigonométrique $P(u) = \sum_k a_k \exp(2\pi iku)$.

2) Montrons la réciproque. Soit γ une suite définie positive. Posons

$$\begin{aligned} f_n(u) &= \frac{1}{n} \sum_{r,s=1}^{n} \gamma(r-s) \exp(2\pi i(s-r)u) \\ &= \sum_{|m|<n} (1 - \frac{|m|}{n}) \gamma(m) \exp(-2\pi imu) \end{aligned}$$

Comme γ est définie positive, f_n est une fonction à valeurs réelles positives. Soit μ_n la mesure de densité f_n. Les mesures μ_n sont positives et de masse bornée indépendante de n.

Calculons les coefficients de Fourier des μ_n :

$$\begin{aligned} \int \exp(2\pi iku)\ d\mu_n(u) &= \sum_{|m|<n} (1 - \frac{|m|}{n})\gamma(m) \int \exp(2\pi i(k-m)u)\ du \\ &= (1 - \frac{|k|}{n})\gamma(k),\ \text{si } |k| < n, = 0,\ \text{sinon.} \end{aligned}$$

La suite (μ_n) appartient au compact (pour la topologie faible) formé des mesures positives de masse 1. On peut donc en extraire une sous-suite (μ_{n_j}) qui converge au sens faible vers une limite μ. On a, par définition même de la convergence faible, pour toute fonction $\varphi \in \mathcal{C}(\mathbb{T})$,

$$\lim_j \int \varphi(u)\ d\mu_{n_j}(u) = \int \varphi(u)\ d\mu(u).$$

Cette convergence a lieu, en particulier, sur les exponentielles. D'après le calcul précdent, on a :

$$\gamma(k) = \lim_j \int \exp(2\pi iku)\ d\mu_{n_j}(u) = \int \exp(2\pi iku)\ d\mu(u).$$

Ceci montre que la suite des coefficients de Fourier de la mesure μ est égale à la suite γ (ou plus exactement à la suite $(\gamma(-n), n \in \mathbb{Z})$). On a donc établi l'existence de la mesure positive sur le cercle annoncée par le théorème. Comme les coefficients de Fourier déterminent la mesure, cette mesure μ est unique.

Cette dernière remarque prouve qu'en fait la suite (μ_n) converge vers μ, puisque cette suite appartient au compact (pour la topologie faible) formé des mesures positives de masse 1 et qu'elle possède une seule valeur d'adhérence, la mesure μ. $\square$

Définition 9.2.3 On appelle *mesure spectrale* d'un processus stationnaire centré du second ordre $X = (X_n)$ la mesure ν_X définie sur le cercle dont les coefficients de Fourier sont donnés par :

$$\widehat{\nu}_X(n) = cov(X_n, X_0) = \mathbb{E}[X_n \overline{X_0}] = \gamma_X(n) = \int_0^1 e^{2\pi inu}\ d\nu_X(u). \tag{9.3}$$

C'est donc la mesure sur le cercle dont la suite des coefficients de Fourier est égale (au changement près de n en $-n$) à la covariance du processus X. Son existence et son unicité sont assurées par le théorème de Bochner-Herglotz (9.2.2).

La mesure ν_X est symétrique (invariante par changement de u en $-u$), si le processus X est à valeurs réelles. Sa masse est égale à $\mathbb{E}[|X_0|^2]$. Le support de ν_X est parfois appelé "spectre" de X.

Si la mesure spectrale ν_X de $X = (X_n)$ a une densité f_X par rapport à la mesure de Lebesgue, $d\nu_X(u) = f_X(u)du$, on dit que X a une *densité spectrale*.

Exemples 9.2.4 a) Bruit blanc de variance σ^2 : sa mesure spectrale est la mesure de Lebesgue multipliée par la constante σ^2. Le terme "bruit blanc" vient de ce que le spectre du processus contient toutes les fréquences $u \in \mathbb{T}$.

b) Dans l'exemple b) de (9.1.4) la mesure spectrale du processus $(X_n)_{n\in\mathbb{Z}}$ est $\nu_X = \sum_k \sigma_k^2 \delta_{u_k}$. C'est une mesure discrète, somme des mesures de Dirac de masse σ_k^2 aux points u_k. Elle est donc singulière par rapport à la mesure de Lebesgue.

c) Toute mesure positive sur le cercle est la mesure spectrale d'un processus stationnaire. En effet, pour toute mesure positive ν sur le cercle, on peut construire un processus stationnaire *gaussien* tel que $\nu_X = \nu$ (cf. c) de 9.1.4).

Il existe donc des processus stationnaires dont la mesure spectrale est sans masse ponctuelle, mais singulière par rapport à la mesure de Lebesgue.

• **Existence d'une densité**

On peut donner des conditions nécessaires et des conditions suffisantes d'existence d'une densité spectrale. D'après le lemme de Riemann-Lebesgue, une condition *nécessaire* pour qu'un processus stationnaire possède une densité spectrale est que $\lim_{|n|\to\infty} \gamma_X(n) = 0$: la covariance du processus tend vers 0, ce qui exprime une *décorrélation asymptotique* des variables aléatoires formant le processus (propriété de mélange fort en terme de système dynamique).

Pour qu'un processus stationnaire $X = (X_n)$ possède une densité spectrale f_X de *carré intégrable*, il faut et il suffit que l'on ait : $\sum_{n\in\mathbb{Z}} |\gamma_X(n)|^2 < \infty$. La densité f_X est alors donnée par la série de Fourier convergeant dans $L^2(\mathbb{T})$:

$$f_X(u) = \sum_{n\in\mathbb{Z}} \gamma_X(-n) e^{2\pi i n u}.$$

Une hypothèse plus forte est la convergence $\sum_{n\in\mathbb{Z}} |\gamma_X(n)| < \infty$, qui assure la convergence normale de la série de Fourier. La densité spectrale est dans ce cas une fonction continue. De façon générale, la vitesse de décorrélation du processus est liée à l'existence et la régularité de la densité spectrale.

• **Représentation isométrique du processus $X = (X_n)$ dans $L^2(\mathbb{T}, \nu_X)$**

Notons $e_n(u) = e^{2\pi i n u}, n \in \mathbb{Z}$. Soit ν une mesure positive. La suite $(e_n)_{n\in\mathbb{Z}}$ forme un processus du second ordre, si l'on considère les fonctions e_n comme éléments de l'espace $L^2(\mathbb{T}, d\nu)$ des fonctions de carré intégrable sur le cercle unité muni de la mesure ν (suite de v.a. orthogonales non indépendantes si ν est la mesure uniforme).

Cette classe fournit une représentation (à une isométrie près) de tous les processus stationnaires du second ordre.

Théorème 9.2.5 *Il existe une isométrie surjective $\mathcal{I}$ de $\mathcal{H}(X)$ sur l'espace $L^2(\mathbb{T}, d\nu_X)$ telle que $\mathcal{I}X_n = e_n$, pour tout $n \in \mathbb{Z}$.*

Preuve : Pour construire l'isométrie, nous procédons, comme précédemment, en définissant d'abord $\mathcal{I}$ sur le sous-espace formé des combinaisons linéaires finies des v.a. X_n :

$$\mathcal{I}(\sum_k \alpha_k X_k) = \sum_k \alpha_k e_k.$$

Cette application est une isométrie :

$$\|\mathcal{I}(\sum_k \alpha_k X_k)\|^2_{L^2(\nu_X)} = \int_0^1 |\sum_k \alpha_k e^{2\pi iku}|^2 \, d\nu_X(u)$$
$$= \int_0^1 \sum_{k,j} \alpha_k \overline{\alpha_j} e^{2\pi i(k-j)u} \, d\nu_X(u) = \sum_{k,j} \alpha_k \overline{\alpha_j} \gamma_X(k-j) = \| \sum_k \alpha_k X_k \|^2_{L^2(\mathbb{P})},$$

qui se prolonge en une isométrie de $\mathcal{H}(X)$ sur le sous-espace fermé engendré dans $L^2(\nu_X)$ par les polynomes trigonométriques, i.e. $L^2(\nu_X)$ tout entier.

L'opérateur de décalage T défini sur $\mathcal{H}(X)$ correspond dans cette représentation à l'opérateur de multiplication par e_1, opérateur unitaire sur $L^2(\nu_X)$. Du point de vue de la structure hilbertienne, le processus $X = (X_n)$ est, via l'isométrie $\mathcal{I}$, isométrique au processus (e_n) défini sur $(\mathbb{T}, \nu_X)$. $\square$

9.3 Décorrélation, variance, cobords

Donnons d'abord un lemme général, utile pour obtenir une réduction d'un processus stationnaire, dans un cadre hilbertien ou de filtration (représentation de martingale).

Lemme 9.3.1 *a) Soit B un opérateur borné sur un espace de Hilbert $\mathcal{B}$ (ou plus généralement sur un espace de Banach réflexif). Soit $f \in \mathcal{B}$ tel que $\sup_{n\geq 1} \| \sum_{k=0}^{n-1} B^k f\| < +\infty$. Alors il existe un élément g de $\mathcal{B}$ tel que $f = g - Bg$.*

b) Si T est un opérateur tel que $BT = Id$, cette propriété est équivalente à l'existence de h et g_1 dans $\mathcal{B}$ tels que $f = g_1 + Th - h$ et $Bg_1 = 0$.

Preuve : a) Pour tout entier $k \geq 0$, posons $\Sigma_k f = \sum_{j=0}^{k-1} B^j f$; nous avons $B\Sigma_k f = \Sigma_k f - f + B^k f$. On en déduit, pour tout entier $n \geq 1$,

$$B[\frac{1}{n}\sum_{k=1}^{n} \Sigma_k f] = \frac{1}{n}\sum_{k=1}^{n} \Sigma_k f \; - \; f \; + \; \frac{1}{n}\sum_{k=1}^{n} B^k f. \tag{9.4}$$

D'après nos hypothèses, la suite $(\frac{1}{n}\sum_{k=1}^{n} B^k f)$ converge en norme vers 0 et la suite $(\frac{1}{n}\sum_{k=1}^{n} \Sigma_k f)_{n\geq 1}$ est bornée dans $\mathcal{B}$. Si $\mathcal{B}$ est séparable, nous pouvons donc extraire de cette dernière suite une sous-suite qui converge faiblement vers un élément g de $\mathcal{B}$. De la relation (9.4), il résulte alors que l'on a $Bg = g - f$.

Si $\mathcal{B}$ n'est pas séparable, on applique ce qui précède au sous-espace fermé de $\mathcal{B}$ invariant par B engendré par les éléments $B^n f$, $n \geq 0$.

Notons que la réciproque est clairement vérfiée, si $\sup \|B^n\| < \infty$.

b) Supposons qu'on ait $BT = Id$. Si f est tel que $f = g_1 + Th - h$, avec $Bg_1 = 0$, on a $Bf = h - Bh$, d'où $f = f + h - B(f + h)$.

Inversement, soit f tel qu'il existe g vérifiant l'équation : $f = g - Bg$. Alors, on peut écrire : $f = g_1 + TBg - Bg$, avec $g_1 = g - TBg$, et on a bien $Bg_1 = Bg - BTBg = 0$.

La démonstration s'applique sans changement à un espace de Banach réflexif $\mathcal{B}$. $\square$

• Variance d'un processus stationnaire

Soit $X = (X_n) = (T^n X_0)$ un processus stationnaire au sens large, centré, défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et formé de v. a. de carré intégrable. Considérons les sommes ergodiques

$$S_n X_0 := X_0 + X_1 + \cdots + X_{n-1}.$$

D'après le théorème de Von Neumann, la suite des moyennes $(\frac{1}{n} S_n X_0)$ converge en norme L^2 vers ΠX_0 la projection de X_0 sur le sous-espace des vecteurs invariants. S'il n'y a pas de vecteurs invariants non nuls dans le sous-espace fermé $\mathcal{H}(X)$ engendré par les v.a. X_k, il y a donc convergence en norme L^2 des moyennes ergodiques vers 0.

Notons que ceci équivaut au fait que la mesure spectrale μ_X du processus n'a pas de masse ponctuelle en 0. Nous pouvons en effet appliquer le théorème de convergence dominée de Lebesgue à la mesure μ_X et à la suite $(|\frac{1}{n}\sum_0^{n-1} e^{2\pi iku}|^2)$ qui est majorée par 1 en module et converge vers $1_{\{0\}}(u)$. L'égalité $\|\frac{1}{n} S_n X_0\|_2^2 = \int_0^1 |\frac{1}{n}\sum_0^{n-1} e^{2\pi iku}|^2 \, d\mu_X(u)$ permet alors de conclure.

Dans le cas d'un processus stationnaire au sens strict centré ergodique, cette condition est assurée et il y a convergence p.s.. Les problèmes d'estimation posent de façon naturelle la question d'une vitesse de convergence de ces moyennes ergodiques. Cette question apparaît également dans l'étude du TCL pour les sommes ergodiques.

La première question est l'étude de $\limsup_n \frac{1}{n}\|S_n\|_2^2$, qui définit la variance du processus. Plus précisément on cherche à montrer que la limite

$$\sigma^2(X_0) := \lim_n \frac{1}{n}\|S_n\|_2^2, \tag{9.5}$$

existe et est finie. Une autre question est celle de la non nullité de la limite. Ce deuxième point est lié, comme nous allons le voir, à la possibilité de représenter X_0 comme un *cobord* pour T.

D'après la stationnarité, on a :

$$\frac{1}{n}\|S_n X_0\|^2 = \|X_0\|^2 + 2\sum_{k=1}^{n-1}(1 - \frac{k}{n})\, \Re e\langle T^k X_0, X_0\rangle, n \geq 1.$$

Il en résulte que, sous la condition

$$\sum_{k=0}^{\infty} |\langle T^k X_0, X_0\rangle| < \infty, \tag{9.6}$$

la variance $\sigma^2(X_0)$ existe et est donnée par l'expression

$$\sigma^2(X_0) = \|X_0\|^2 + 2\sum_1^{\infty} \Re e\langle T^k X_0, X_0\rangle. \tag{9.7}$$

Sous la condition (9.6) la mesure spectrale a une densité ϕ_x continue donnée par :

$$\phi_x(u) = \sum_{k\in\mathbb{Z}} \langle T^k X_0, X_0\rangle \; e^{2\pi iku}. \tag{9.8}$$

On a alors

$$\sigma^2(X_0) = \phi_x(0).$$

On note que cette relation est vérifiée, dès que la mesure spectrale associée à X_0 a une densité ϕ_x continue à l'origine. On peut en effet exprimer $\|S_n X_0\|$ sous la forme

$$\|S_n X_0\|_2^2 = \int_0^1 [\frac{\sin \pi n u}{\sin \pi u}]^2 \ \phi_x(u) \ du,$$

et appliquer le fait que la suite $(\frac{1}{n}[\frac{\sin \pi n u}{\sin \pi u}]^2)_{n\geq 1}$ forme une "identité approchée" sur $[-\frac{1}{2}, \frac{1}{2}]$.

Considérons une isométrie T dans un espace de Hilbert $\mathcal{V}$. Nous dirons qu'un X_0 est un *cobord* pour T, s'il existe $Z_0 \in \mathcal{V}$ tel que $X_0 = TZ_0 - Z_0$.

La condition "variance finie" ($\sigma^2(X_0) < \infty$) pour X_0 reste vérifiée, si l'on ajoute à X_0 un cobord dans $L^2(\mu)$ (alors que la propriété de décorrélation avec vitesse peut ne pas subsister). Si X_0 est un cobord dans $L^2(\mu)$, la suite des normes $\|S_n X_0\|_2$ est bornée. On a une réciproque et si la corrélation tend suffisamment vite vers zéro, on peut caractériser les fonctions de variance nulle comme les cobords dans L^2.

Proposition 9.3.2 *a) La suite $(\|S_n X_0\|_2)_{n\in\mathbb{N}}$ est bornée si et seulement si X_0 est un cobord dans $L^2(\mu)$.*

b) Sous la condition de décorrélation

$$\sum_{k=0}^{\infty} k|\langle T^k X_0, X_0\rangle| < \infty \tag{9.9}$$

la suite $(\|S_n X_0\|^2 - n\sigma^2(X_0))_{n\geq 0}$ est bornée. De plus, sous cette condition, la variance $\sigma^2(X_0)$ est nulle si, et seulement si, X_0 est un cobord dans L^2.

Preuve : a) La première assertion est un corollaire du lemme 9.3.1 appliqué en prenant pour opérateur B l'opérateur T lui-même.

b) Sous l'hypothèse (9.9) nous avons la majoration :

$$\begin{aligned} |\|S_n\|^2 - n\sigma^2(X)| &= |n\|X_0\|^2 + 2\sum_{k=1}^{n}(n-k)\langle T^k X_0, X_0\rangle - n[\|X_0\|^2 + 2\sum_{k=1}^{\infty}\langle T^k X_0, X_0\rangle]| \\ &= 2|n\sum_{k=n+1}^{\infty}\langle T^k X_0, X_0\rangle - \sum_{k=1}^{n} k\ \langle T^k X_0, X_0\rangle| \leq 2\sum_{k=1}^{\infty} k|\langle T^k X_0, X_0\rangle| < \infty. \end{aligned}$$

Si de plus $\sigma^2(X) = 0$, la suite $(\|S_n\|_2)_{n\in\mathbb{N}}$ est bornée. D'après le lemme 9.3.2, ceci implique qu'il existe $Z_0 \in L^2$ tel que $X_0 = TZ_0 - Z_0$. □

• En terme de système dynamique

Ce qui vient d'être fait en terme de processus du second ordre peut être reformulé en terme de système dynamique. Considérons un système dynamique inversible $(E, \mathcal{A}, \mu, \tau)$.

Soit f une fonction à valeurs réelles dans $L^2(\mu)$ d'intégrale nulle. Notons $T^n f$ les composées $f \circ \tau^n$, $n \in \mathbb{N}$.

Définition 9.3.3 La **variance** $\sigma^2(f)$ est définie comme la limite (si cette limite existe et est finie) :

$$\sigma^2(f) = \lim_{n\to+\infty} \frac{1}{n}\|f + Tf + ... + T^{n-1}f\|_2^2. \tag{9.10}$$

Sous la condition

$$\sum_{k\in\mathbb{N}} |\langle T^k f, f\rangle| < \infty \tag{9.11}$$

la variance $\sigma^2(f)$ existe et est donnée par l'expression

$$\sigma^2(f) = \|f\|^2 + 2\sum_{k=1}^{\infty}\langle T^k f, f\rangle. \tag{9.12}$$

On dit qu'une fonction f est un **cobord** au sens de τ (dans la classe des fonctions mesurables), s'il existe h mesurable telle que $f = Th - h$. Si h est dans L^2, on dit que f est un cobord dans L^2.

D'après la proposition 9.3.2, sous l'hypothèse

$$\sum_{k\in\mathbb{N}} k|\langle T^k f, f\rangle| < \infty \tag{9.13}$$

f est un cobord dans L^2 si et seulement si $\sigma^2(f) = 0$.

Notons qu'une fonction peut être un cobord au sens des fonctions mesurables, mais vérifier : $\sigma^2(f) = 1$. Par exemple sur le tore $\mathbb{T}^2$, faisons opérer la transformation : $\tau : (x, y) \to (x, y{+}x \bmod 1)$. La fonction $f(x, y) = e^{2\pi i y}$ s'écrit $f = Th - h$, avec $h(x, y) = e^{2\pi i y}/(e^{2\pi i x} - 1)$, mais vérifie : $\sigma^2(f) = 1$.

Nous verrons par la suite qu'il existe des systèmes dynamiques pour lesquels toute fonction régulière, qui est un cobord au sens des fonctions mesurables, l'est aussi au sens L^2.

9.4 Filtration vectorielle

Définition 9.4.1 Nous appellerons "**T-filtration vectorielle décroissante**", dans un espace de Hilbert $\mathcal{V}$ muni d'une isométrie T, une suite décroissante $(\mathcal{V}_n)_{n\geq 0}$ de sous-espaces fermés de $\mathcal{V}$, telle que $T^n\mathcal{V} = \mathcal{V}_n$. On définit de même une T-filtration vectorielle croissante.

Soit P l'opérateur adjoint de T : $\langle Tf, g\rangle = \langle f, Pg\rangle, \forall f, g \in \mathcal{V}$. Si $\mathcal{W}$ est un sous-espace de $\mathcal{V}$, notons $\Pi_{\mathcal{W}}$ le projecteur orthogonal sur $\mathcal{W}$.

On note que $PT = Id$, car $\langle PTf, g\rangle = \langle Tf, Tg\rangle = \langle f, g\rangle$. D'autre part, on a $TP = \Pi_{T\mathcal{V}}$, et comme T est une isométrie bijective de $\mathcal{V}$ sur $T\mathcal{V}$, on peut écrire $P = T^{-1}\Pi_{T\mathcal{V}}$.

Plus généralement, on a $T^nP^n = \Pi_{T^n\mathcal{V}}$, car

$$\langle f - T^nP^nf, T^ng\rangle = \langle f, T^ng\rangle - \langle P^nf, g\rangle = 0.$$

Si $\bigcap_n \mathcal{V}_n = \{0\}$, on peut écrire f comme la somme de la série orthogonale :

$$f = (f - TPf) + (TPf - T^2P^2f) + \cdots + (T^nP^nf - T^{n+1}P^{n+1}f) + \cdots.,$$

où le n-ième terme de cette série, $T^nP^nf - T^{n+1}P^{n+1}f$, est dans le supplémentaire orthogonal de $T^{n+1}\mathcal{V} = \mathcal{V}_{n+1}$ dans $T^n\mathcal{V} = \mathcal{V}_n$.

Processus stationnaires réguliers

Considérons un processus stationnaire (X_n). Nous avons vu qu'il peut être représenté sous la forme $X_n = T^n X_0$. Une filtration vectorielle associée à (X_n) peut être construite de la façon suivante.

On note $\mathcal{H}_n(X)$ le sous-espace vectoriel fermé engendré dans $L^2(\Omega, \mathcal{F}, \mathbb{P})$ par les v.a. X_k, pour $k \leq n$:

$$\mathcal{H}_n(X) = \overline{sp}\{X_k, k \leq n\}.$$

Ce sous-espace s'interprète comme l'espace du passé (linéaire), contenant l'information apportée par le processus jusqu'au temps n. L'opérateur de décalage T fait passer d'un sous-espace $\mathcal{H}_n(X)$ au suivant :

$$T\mathcal{H}_n(X) = \mathcal{H}_{n+1}(X).$$

Les sous-espaces $\mathcal{H}_n(X)$ forment une suite croissante avec n. Leur réunion $\cup_n \mathcal{H}_n(X)$ est dense dans $\mathcal{H}(X)$. L'intersection de ces sous-espaces

$$\mathcal{H}_{-\infty}(X) = \bigcap_n \mathcal{H}_n(X) = \bigcap_n T^n \mathcal{H}_0(X),$$

est un sous-espace fermé de $\mathcal{H}(X)$, contenu dans tous les $\mathcal{H}_n(X)$ et strictement invariant par T : $T\mathcal{H}_{-\infty}(X) = \mathcal{H}_{-\infty}(X)$. On a obtenu ainsi une filtration vectorielle :

$$\mathcal{H}_{-\infty}(X) \subset \cdots \subset \mathcal{H}_n(X) \subset \cdots \subset \mathcal{H}(X). \tag{9.14}$$

Définition 9.4.2 Le processus $X = (X_n)$ est dit *singulier* si $\mathcal{H}_{-\infty}(X) = \mathcal{H}(X)$ (ce qui équivaut à $\mathcal{H}_n(X) = \mathcal{H}_{n+1}(X), \forall n \in \mathbb{Z}$). Il est dit *régulier*, si $\mathcal{H}_{-\infty}(X) = \{0\}$.

Dans le cas général, la codimension de $\mathcal{H}_{n-1}(X)$ dans $\mathcal{H}_n(X)$ est 0 ou 1. Dans le cas régulier, elle est égale à 1 (sinon on aurait l'égalité $\mathcal{H}_{n-1}(X) = \mathcal{H}_n(X)$ et tous les espaces serait égaux).

Plaçons-nous dans le cas régulier. Soit W_0 un vecteur unitaire dans l'orthogonal de $\mathcal{H}_1$ dans $\mathcal{H}_0$ et $W_n = T^{-n} W_0, \forall n \in \mathbb{Z}$. La suite (W_n) est un bruit blanc et forme une base orthonormée de l'espace $\mathcal{H}(X)$ engendré par le processus régulier X. En décomposant X_0 dans cette base adaptée, on obtient :

$$X_0 = \sum_{k=0}^{\infty} a_k W_{-k},$$

et par application de l'opérateur T :

$$X_n = \sum_{k=0}^{\infty} a_k W_{n-k}.$$

Ceci montre la première partie du théorème ci-dessous, dont la deuxième partie est laissée en exercice.

Théorème 9.4.3 *(Décomposition de Wold) Si X est régulier, il existe un bruit blanc W dans l'espace $\mathcal{H}(X)$ et une suite de carré sommable de coefficients $h = (h_k)_{k \geq 0}$ (un filtre) telle que*

$$X_n = \sum_{k=0}^{\infty} h_k W_{n-k}.$$

Tout processus stationnaire X du second ordre se décompose en la somme d'un processus régulier X^R et d'un processus singulier X^S mutuellement orthogonaux, qui engendrent des sous-espaces invariants par l'opérateur de décalage dans l'espace $\mathcal{H}(X)$ du processus X.

• Réduction dans une filtration vectorielle

Cas unilatère

Considérons une T-filtration vectorielle pour une isométrie T dans un espace de Hilbert $\mathcal{V}_0$:

$$\cdots \subset T^n\mathcal{V}_0 \subset \cdots \subset T\mathcal{V}_0 \subset \mathcal{V}.$$

Proposition 9.4.4 *Sous la condition*

$$\sum_{n\geq 0} \|\Pi_{T^n\mathcal{V}_0} f\| < \infty,$$

il existe g et h dans $\mathcal{H}$, avec $g \in \mathcal{V}_0 \ominus T\mathcal{V}_0$, tels que $f = g + Th - h$ et f est "homologue" à un vecteur g tel que les vecteurs $T^n g, n \geq 0$ soient deux à deux orthogonaux.

Preuve : En notant P l'opérateur adjoint de T dans $\mathcal{V}_0$, on a $\Pi_{T^n\mathcal{V}_0} f = T^n P^n f$. La condition

$$\sum_{n\geq 0} \|P^n f\| = \sum_{n\geq 0} \|\Pi_{T^n\mathcal{V}_0} f\| < \infty$$

permet d'appliquer le lemme 9.3.1 à f et à l'opérateur P. $\square$

Cas bilatère

Traitons maintenant le cas "bilatère". Nous supposons que l'opérateur T est isométrique et **surjectif**, autrement dit T est un opérateur unitaire dans un espace de Hilbert $\mathcal{H}$. Soit $\mathcal{V}_0$ un sous-espace de $\mathcal{H}$ tel que $\mathcal{V}_0 \subset T\mathcal{V}_0$. Considérons la filtration vectorielle "bilatère" pour T associée, i.e. la suite de sous-espaces vectoriels emboîtés :

$$\cdots \subset \mathcal{V}_{-n} \subset \cdots \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \cdots \subset \mathcal{V}_n \subset \cdots \subset \mathcal{H}, \tag{9.15}$$

telle que $T^n\mathcal{V}_0 = \mathcal{V}_n$, $\forall n \in \mathbb{Z}$.

On supposera que la réunion $\bigcup_n \mathcal{V}_n$ est dense dans $\mathcal{V}$ et que $\cap_n \mathcal{V}_n = \{0\}$.

Notons $\mathcal{W}_0 = \mathcal{V}_0^\perp$, $\mathcal{W}_n = T^n\mathcal{W}_0 = \mathcal{V}_{-n}^\perp$. La condition $T^{-1}\mathcal{V}_0 \subset \mathcal{V}_0$ équivaut à $\mathcal{W}_0 \subset T^{-1}\mathcal{W}_0$.

On a donc les inclusions :

$$\{0\} \subset \cdots \subset \mathcal{W}_n \subset \cdots \subset \mathcal{W}_1 \subset \mathcal{W}_0.$$

Notons P l'adjoint de la restriction de T^{-1} à $\mathcal{V}$ et Q l'adjoint de la restriction de T à $\mathcal{W}_0$. Donnons une condition suffisante pour qu'un vecteur de $\mathcal{H}$ soit égal, à un cobord près, à un vecteur de $\mathcal{V}_0$ orthogonal à $T\mathcal{V}_0$.

Théorème 9.4.5 *Etant donnée une filtration bilatère (9.15), sous les conditions*

$$\sum_{n\geq 0} \|\Pi_{T^{-n}\mathcal{V}_0} f\| < \infty \ \textit{et} \ \sum_{n\geq 0} \|f - \Pi_{T^n\mathcal{V}_0} f\| < \infty,$$

il existe g et h dans $\mathcal{H}$, avec $g \in \mathcal{V}_0 \ominus T^{-1}\mathcal{V}_0$, tels que

$$f = g + Th - h.$$

En particulier, sous ces conditions, f est "homologue" à un vecteur g tel que les vecteurs $T^n g$ soient deux à deux orthogonaux. et f vérifie la condition de variance finie et l'on a :

$$\Big|\|\sum_0^{n-1} T^k f\| - \sqrt{n}\|g\|\Big| \leq 2\|h\|.$$

Preuve : On décompose f en $f = f_1 + f_2$ avec $f_1 \in \mathcal{V}_0$ et $f_2 \in \mathcal{W}_0$.

On a $\Pi_{T^{-n}\mathcal{V}_0} f = \Pi_{T^{-n}\mathcal{V}_0} f_1 = T^{-n} P^n f_1$, pour $n \geq 0$, et donc la condition

$$\sum_{n\geq 0} \|P^n f_1\| = \sum_{n\geq 0} \|\Pi_{T^{-n}\mathcal{V}_0} f\| < \infty$$

permet d'appliquer le lemme 9.3.1 à f_1 et à l'opérateur P : il existe g_1 et h_1 tels que $f_1 = g_1 + T{-}1h_1 - h_1$, avec $g_1 \in \mathcal{V}_0 \ominus T^{-1}\mathcal{V}_0$.

De même, on a

$$f - \Pi_{T^n\mathcal{V}_0} f = \Pi_{\mathcal{W}_n} f_2 = T^n Q^n f_2.$$

La condition

$$\sum_{n\geq 0} \|Q^n f_2\| = \sum_{n\geq 0} \|f - \Pi_{T^n\mathcal{V}_0} f\| < \infty$$

permet d'appliquer le lemme 9.3.1 à f_2 : Il existe g_2 et h_2, avec $g_2 \in \mathcal{W}_0 \ominus T\mathcal{W}_0 = T\mathcal{V}_0 \ominus \mathcal{V}_0)$, tels que $f_2 = g_2 + Th_2 - h_2$.

On obtient la représentation de f annoncée en prenant $g = g_1 + T^{-1}g_2$ et $h = T^{-1}g_2 + h_2 - T^{-1}h_1$.

□

Chapitre 10

Spectre et entropie des systèmes dynamiques

Dans ce chapitre nous donnons des méthodes de construction de systèmes dynamiques et étudions quelques propriétés : le spectre, la propriété de K-système, puis l'entropie.

10.1 Construction de systèmes dynamiques

Nous allons décrire des procédés permettant de construire de nouveaux systèmes dynamiques à partir d'un ou de plusieurs systèmes donnés.

Considérons un système dynamique $(E, \mathcal{E}, \mu, \tau)$, où μ est une mesure de probabilité invariante par τ.

Système induit

Si A est un ensemble mesurable de mesure > 0, on définit le temps de premier retour dans A, pour $x \in A$, par $n_A(x) = \inf(n \geq 1, \tau^n x \in A)$, si x revient dans A, $= +\infty$ sinon. Le théorème de récurrence de Poincaré assure que $n_A(x) < \infty$, pour μ-presque tout $x \in A$:

Théorème 10.1.1 *L'ensemble $B := \{x \in A : \tau^n x \notin A, \forall n \geq 1\}$ est de μ-mesure nulle.*
Preuve : Les ensembles $\tau^k B$, $k \geq 1$, sont deux à deux disjoints. Ils ont même mesure et la mesure μ est finie, ce qui implique $\mu(B) = 0$. $\square$

La transformation induite τ_A sur A est définie sur A par

$$\tau_A(x) = \tau^{n_A} x.$$

Notons $A_k := \{x : n_A(x) = k\}$, pour $k \geq 1$. Si τ est ergodique sur X, l'espace X se décompose en "tours" de hauteur k. Il est la réunion disjointe :

$$X = \bigcup_{k \geq 1} \cup_{j=0}^{k-1} \tau^j A_k.$$

Soit f invariante par τ_A sur A. On définit $\tilde{f}$ sur X par $\tilde{f}(x) := f(y)$, si $x = \tau^j y$, $0 \leq j < n(x)$. Si f est τ_A-invariante sur la base, alors $\tilde{f}$ est τ-invariante, donc constante, de même que f. Ainsi l'ergodicité de τ entraîne celle de τ_A.

Flot spécial

La construction d'un flot spécial est la construction inverse de l'induction. On se donne une fonction φ (*fonction plafond*) telle que $\varphi(x) > 0, \forall x \in E$ et $\int_E \varphi \, d\mu = 1$.

Soit $F = \{(x,s) \in E \times \mathbb{R} \text{ tel que } 0 \leq s < \varphi(x)\}$. On met sur F la mesure trace de la mesure produit $\mu(dx) \otimes ds$. Pour $(x,s) \in F$ et $t > 0$, si $s + t \geq \varphi(x)$, soit $n(x, s+t)$ défini par

$$\sum_{k=0}^{n(x,s+t)} \varphi(\tau^k x) \leq s + t < \sum_{k=0}^{n(x,s+t)+1} \varphi(\tau^k x).$$

On définit le flot spécial $(\tilde{\tau}_t)_{t \geq 0}$ au-dessus de $(E, \mathcal{E}, \mu, \tau)$ par $\tilde{\tau}_t(x,s) = (x, s+t)$, si $0 \leq s+t < \varphi(x)$ et, pour $s + t \geq \varphi(x)$, par

$$\tilde{\tau}_t(x,s) = (\tau^{n(x,s+t)} x, s + t - \sum_{k=0}^{n(x,s+t)} \varphi(\tau^k x)).$$

Exercice : Si τ est ergodique sur la base, pour toute fonction plafond le flot spécial associé est ergodique.

Produit de systèmes

Considérons deux systèmes dynamiques $(E, \mathcal{E}, \mu, \tau)$ et $(E', \mathcal{E}', \mu', \tau')$, les mesures μ et μ', respectivement τ et τ' invariantes, n'étant pas nécessairement finies.

Le produit des deux systèmes est défini sur l'espace produit $E \times E'$ muni de la tribu produit des tribus $\mathcal{E}$ et $\mathcal{E}'$ et de la mesure produit $\mu \times \mu$ par la transformation

$$(x, x') \to (\tau x, \tau' x').$$

10.2 Propriétés spectrales d'un système dynamique

Considérons un système dynamique $(E, \mathcal{E}, \mu, \tau)$. On note T l'opérateur de composition par τ

Fonctions propres

Une fonction f de E dans $\mathbb{C}$ mesurable non nulle est une fonction **propre** pour T, si elle vérifie $Tf = \alpha f$, pour une constante α.

Si f est dans $L^2(\mu)$, en passant au module et en utilisant l'invariance de la mesure μ, on obtient que $|\alpha| = 1$. Ceci implique que le module de f est invariant, donc constant dans le cas où le système est ergodique.

Lorsque $(E, \mathcal{E}, \mu, \tau)$ est un système dynamique ergodique de mesure finie, toute fonction f propre est de module constant et donc nécessairement dans $L^2(\mu)$. L'ergodicité d'un système dynamique en mesure finie est ainsi caractérisée par le fait que l'opérateur T dans $L^2(\mu)$ possède 1 comme valeur propre simple. C'est donc, en mesure finie, un exemple de propriété spectrale, c'est-à-dire attachée à l'opérateur T opérant dans $L^2(\mu)$.

Les translations sur le cercle, ou plus généralement sur un groupe compact abélien (muni de la mesure de Haar normalisée λ, mesure invariante par les translations par les éléments du groupe), possèdent une base de fonctions propres dans $L^2(\lambda)$ (systèmes à spectre discret). A isomorphisme près, ce sont exactement les systèmes dynamiques ayant la propriété de spectre discret : en effet, on peut montrer que *si un système dynamique ergodique possède une base de fonctions propres, il est isomorphe à une translation sur un groupe compact abélien muni de la mesure de Haar.*

Si le système est ergodique, d'après le théorème de Von Neuman, on a

$$\lim_n \frac{1}{n} \sum_0^{n-1} \langle T^k f, g\rangle = \langle f, 1\rangle \ \langle 1, g\rangle, \forall f, g \in L^2(\mu). \tag{10.1}$$

En particulier, si f ou g sont orthogonales aux constantes, on a

$$\lim_n \frac{1}{n} \sum_0^{n-1} \langle T^k f, g\rangle = 0.$$

Appliquée à des fonctions indicatrices d'ensembles, la convergence (10.1) donne :

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} \mu(\tau^{-k} A \cap B) = \mu(A)\mu(B). \tag{10.2}$$

L'ergodicité implique donc que la corrélation $\langle T^n f, g\rangle$, pour des fonctions $f, g \in L^2(\mu)$ d'intégrale nulle, converge en moyenne vers 0. Nous allons introduire des propriétés qui renforcent ce résultat de convergence.

Définitions 10.2.1 Le système est dit à *spectre continu*, si l'opérateur T n'a pas d'autre fonction propre que les constantes.

On montre (exercice) que le système est à spectre continu si et seulement s'il vérifie la condition dite de *mélange faible* (intermédiaire entre ergodicité et mélange fort) :

$$\lim_{N\to\infty} \frac{1}{N} \sum_0^{N-1} |\langle T^n f, g\rangle - \mu(f)\mu(\overline{g})| = 0, \forall f, g \in L^2(\mu). \tag{10.3}$$

Cette propriété est vérifiée si, et seulement si,

$$\lim_n \frac{1}{n} \sum_0^{n-1} |\mu(\tau^{-n} A \cap B) - \mu(A)\mu(B)| = 0, \ \forall A, B \in \mathcal{E}. \tag{10.4}$$

Une autre caractérisation du mélange faible est donnée par la proposition suivante :

Proposition 10.2.2 *Un système dynamique vérifie la propriété de mélange faible si et seulement si le carré cartésien du système par lui-même est ergodique.*

<u>Preuve</u> : 1) Soient $f, g, f', g' \in L^2(\mu)$. D'après le lemme ci-dessous, sous la condition de mélange faible, il existe une suite de densité nulle en dehors de laquelle les suites $(\langle T^k f, g\rangle)$, $\langle T^k f', g'\rangle)$ convergent vers 0. On a donc :

$$\lim_n \frac{1}{n} \sum_0^{n-1} \langle T^k f, g\rangle \langle T^k f', g'\rangle \ = \ \langle f, 1\rangle \ \langle 1, g\rangle \langle f', 1\rangle \ \langle 1, g'\rangle, \ \forall f, g, f', g' \in L^2(\mu).$$

En approximant les fonctions sur l'espace produit par des combinaisons linéaires de fonctions de la forme $f(x)g(y)$, on obtient à partir de la relation précédente :

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1}\int F(\tau^k x, \tau^k y)\,\overline{G(x,y)}\; d\mu(x)\, d\mu(y)$$
$$= (\int F(x,y)\; d\mu(x)\, d\mu(y))\; (\int \overline{G(x,y)}\; d\mu(x)\, d\mu(y)), \forall F, G \in L^2(\mu\times\mu).$$

Ceci implique l'ergodicité du produit.

2) Inversement, si f est une fonction propre pour le système, $f\circ\tau = \lambda f$ (avec nécessairement $|\lambda| = 1$) la fonction $(x,y)\to f(x)\overline{f(y)}$ est invariante pour la transformation produit $\tau\times\tau$, donc constante, ce qui implique que f est constante. □

Lemme 10.2.3 *Les conditions suivantes pour une suite (a_n) positive bornée sont équivalentes :*
(1) $\lim_n \frac{1}{n}\sum_0^{n-1} a_k = 0$,
(2) $\lim_n \frac{1}{n}\sum_0^{n-1} a_k^2 = 0$,
(3) la suite (a_n) tend vers 0, en dehors d'un ensemble de densité nulle.

Preuve : Notons $M = \sup_k a_k$.

L'implication (1) ⇒ (2) résulte de l'inégalité de Cauchy-Schwarz :

$$\frac{1}{n}\sum_0^{n-1} a_k \leq (\frac{1}{n}\sum_0^{n-1} a_k^2)^{1/2},$$

et l'implication (2) ⇒ (1) de la majoration

$$\frac{1}{n}\sum_0^{n-1} a_k^2 \leq \frac{M}{n}\sum_0^{n-1} a_k.$$

Montrons que (3) implique (1). Soit $\mathcal{N}$ un ensemble d'entiers de densité nulle tel que

$$\lim_{n\to+\infty, n\notin\mathcal{N}} a_n = 0.$$

La majoration

$$\frac{1}{n}\sum_0^{n-1} a_k \leq \frac{M}{n}\sum_0^{n-1} 1_{\mathcal{N}}(k) + \frac{1}{n}\sum_{k\notin\mathcal{N}, 0\leq k<n} a_k$$

montre que la suite (a_n) converge en moyenne de Césaro vers 0.

Enfin, supposons qu'on ait (1). Soit (N_j) une suite strictement croissante telle que

$$\frac{1}{n}\sum_0^{n-1} a_k \leq \frac{1}{j^2}, \text{ pour } n \geq N_j.$$

Posons

$$\mathcal{N} = \bigcup_j ([N_j, N_{j+1}[\cap\{k : a_k > 1/j\}).$$

Montrons que l'ensemble $\mathcal{N}$ est de densité nulle. Soit n, avec $N_j \leq n < N_{j+1}$. Pour $k \leq n$ et $k \in \mathcal{N}$, on a, par construction de l'ensemble $\mathcal{N}$:

$$k \in [N_{j'}, N_{j'+1}[\cap\{\ell : a_\ell > 1/j'\},$$

pour un $j' \leq j$. D'où : $a_k j \geq a_k j' > 1$, ce qui permet d'écrire :

$$\frac{1}{n}\sum_{0}^{n-1} 1_{\mathcal{N}}(k) \leq \frac{j}{n}\sum_{0}^{n-1} a_k \leq \frac{j}{j^2} = \frac{1}{j} \to 0.$$

Enfin, on a $\lim_{k \notin \mathcal{N}} a_k = 0$, car, pour $k \in [N_j, N_{j+1}[$ et $k \notin \mathcal{N}$, on a $a_k < \frac{1}{j}$. □

Définitions 10.2.4 On dit que le système est *fortement mélangeant*, s'il y vérifie :

$$\lim_{n\to\infty} \langle T^n f, g\rangle = \langle f, 1\rangle\ \langle 1, g\rangle,\ \forall f, g \in L^2(\mu), \tag{10.5}$$

ou, de façon équivalente :

$$\lim_n \mu(\tau^{-n} A \cap B) = \mu(A)\mu(B),\ \forall A, B \in \mathcal{E}. \tag{10.6}$$

On dit que le système est à *spectre de Lebesgue*, s'il existe une famille $(f_i)_{i\in I}$ de fonctions telle que les fonctions $(T^n f_i, n \in \mathbb{Z}, i \in I)$ forment une base orthonormale de l'orthogonal des constantes dans $L^2(\mu)$. On a supposé dans cette définition τ inversible. Si le système n'est pas inversible, on restreint l'indice n à $\mathbb{N}$. Le cardinal de l'ensemble d'indices J est appelé multiplicité du spectre.

On vérifiera en exercice que tout système à spectre de Lebesgue vérifie la propriété de mélange fort.

Remarques : Les propriétés de décorrélation que nous venons de définir sont des propriétés spectrales : elles ne dépendent que de l'opérateur T associé à la transformation τ et fournissent des énoncés de convergence sans vitesse. Pour obtenir des vitesses de convergence vers 0 des corrélations, il est nécessaire de se restreindre à des sous-espaces de fonctions particulières (par exemple, dans l'exemple $x \to ax$ mod 1 , à l'espace des fonctions lipschitziennes).

Exemple 1

Soit ρ_ϕ la transformation extension d'une rotation ergodique par la fonction $\phi(x) = x$, c'est-à-dire :

$$\rho_\phi(x, y) \to (x + \alpha, y + x) \bmod 1.$$

Cette transformation possède un facteur, la rotation d'angle α, qui est à spectre discret. La transformation n'est donc pas à spectre continu.

Soit V_0 le sous-espace de $L^2(\mu)$ formé des fonctions ne dépendant que de x, et $V_0^\perp$ son orthogonal. Ces deux sous-espaces sont fermés dans $L^2(\mu)$ et T-invariants. On peut restreindre T à $V_0^\perp$.

Proposition 10.2.5 *La transformation est à spectre de Lebesgue sur l'orthogonal des fonctions ne dépendant que de x.*

Preuve : Soient $\psi_{p,q}(x, y) = \exp(2\pi i[px + qy])$. La famille $\{\psi_{p,q}, p, q \in \mathbb{Z}, q \neq 0\}$ forme une base du sous-espace $V_0^\perp$.

On a $T\psi_{p,q}(x, y) = \exp(2\pi i\alpha)\exp(2\pi i[(p + q)x + qy])$ et donc, pour $q \neq 0$, $T\psi_{p,q} \perp \psi_{p,q}$.

Soit J l'ensemble d'indices $J = \{(0, q), (1, q), ..., (q - 1, q), q = \pm 1, \pm 2, ...\}$.

La famille de fonctions $(\psi_{p,q}, (p, q) \in J)$ a la propriété que les images $(T^n\psi_{p,q}, n \in \mathbb{Z}, (p, q) \in J)$ sont deux à deux orthogonales et forment une base orthonormée de $V_0^\perp$. □

Exemple 2 : Schémas de Bernoulli

Considérons l'espace $\Omega^+ = I^{\mathbb{N}}$, l'espace produit (unilatère) d'une famille de $\mathbb{N}$ copies d'un alphabet I fini, muni de sa tribu borélienne $\mathcal{E}$ et du décalage θ ($\theta(\omega)_n = \omega_{n+1}$, $n \geq 0$).

Rappelons que, pour tout vecteur de probabilité $p = (p_1, \ldots, p_{|I|})$ sur I, on a défini une mesure de probabilité produit $\mathbb{P}_p = \mathbb{P}$ sur Ω^+ ("mesure de Bernoulli") invariante par l'action du décalage θ, en posant

$$\mathbb{P}_p(X_0 = a_0, \ldots, X_n = a_n) = p_{a_0} \ldots p_{a_n},$$

et en prolongeant $\mathbb{P}_p$ à la tribu borélienne de Ω^+.

Nous avons déjà établi (en 8.5) la propriété de mélange fort :

$$\lim_{n \to \infty} \langle f \circ \theta^n, g \rangle = (\int f \; d\mathbb{P})(\int \overline{g} \; d\mathbb{P}), \; \forall f, g \in L^2(\mathbb{P}).$$

Montrons la propriété de spectre de Lebesgue.

Soit V_0 l'orthogonal dans $L^2(\Omega^+, \mathcal{E}, \mathbb{P})$ du sous-espace $L^2(\Omega^+, \theta^{-1}\mathcal{E}, \mathbb{P})$ (formé des fonctions de la forme $f = g \circ \theta$, $g \in L^2(\Omega^+, \mathcal{E}, \mathbb{P})$). Notons $(f_j)_{j \in J}$ une base orthonormée de V_0. (On peut montrer que l'ensemble d'indices J est de cardinal dénombrable infini.)

La suite des sous-espaces $L^2(\Omega^+, \theta^{-N}\mathcal{E}, \mathbb{P})$ est décroissante d'intersection réduite aux constantes. On vérifie que les fonctions $(f_j \circ \theta^n, j \in J, 0 \leq n \leq N)$ forment une base orthonormée de l'orthogonal dans $L^2(\Omega^+, \mathcal{E}, \mathbb{P})$ du sous-espace $L^2(\Omega^+, \theta^{-N}\mathcal{E}, \mathbb{P})$ (formé des fonctions de la forme $f = g \circ \theta^N$, $g \in L^2(\Omega^+, \mathcal{E}, \mathbb{P})$. Il en résulte que la famille des fonctions $(f_j \circ \theta^n, j \in J, 0 \leq n < \infty)$ forme une base orthonormée de l'orthogonal des constantes.

On a ainsi montré que, pour tout vecteur de probabilité p sur I, le système dynamique $(\Omega, \mathbb{P}_p, \theta)$ (schéma de Bernoulli) est à spectre de Lebesgue dénombrable.

Ceci montre en particulier que tous les schémas de Bernoulli ont le même type spectral : les opérateurs T définis sur $L^2(\Omega^+, \mathcal{E}, \mathbb{P}_p)$ comme opérateurs de composition par θ sont tous conjugués entre eux. Les "invariants spectraux" ne permettent donc pas de montrer que les schémas de Bernoulli, construits par exemple sur les vecteurs $(\frac{1}{2}, \frac{1}{2})$ et sur $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, ne sont pas isomorphes. Pour montrer qu'ils ne sont pas isomorphes, il faut construire un autre invariant : l'**entropie** définie plus loin.

10.3 Systèmes exacts et K-systèmes

• Systèmes exacts

Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique.

La suite $(\tau^{-n}\mathcal{E})_{n \geq 0}$ est une suite décroissante de sous-tribus, qui est strictement décroissante si la transformation τ n'est pas inversible. Nous allons définir une classe de systèmes dynamiques pour lesquels τ est "fortement non inversible" et qui ont un comportement de type stochastique.

Définition 10.3.1 *(Rokhlin [Rokh67])* Le système $(E, \mathcal{E}, \mu, \tau)$ est dit **exact** si l'intersection $\bigcap_n \tau^{-n}\mathcal{E}$ est triviale (aux ensembles négligeables près).

Cette propriété est équivalente à

$$\lim_n \mathbb{E}(f | \tau^{-n}\mathcal{E}) = \mu(f), \forall f \in L^2(\mu). \tag{10.7}$$

Notons P l'opérateur adjoint de l'opérateur T (de composition par τ) dans $L^2(\mathcal{E}, \mu)$. L'opérateur P est un inverse à gauche de T : $PT = Id$ et plus généralement :

$$P^n(uT^n f) = P^n(u)f, \ \forall n \in \mathbb{N}, \ \forall f \in L^2(\mu), \forall u \in L^\infty(\mu). \tag{10.8}$$

En effet, l'opérateur T étant un opérateur de composition, nous avons :

$$\langle P^n(uT^n f), g\rangle = \langle uT^n f, T^n g\rangle = \int uT^n(fg)\, d\mu = \int P^n(u)fg\, d\mu.$$

Les fonctions de la forme $T^n g$, $g \in L^2$, sont les fonctions de L^2 qui sont $\tau^{-n}\mathcal{E}$ mesurables. Pour tout $n \geq 1$, l'opérateur $T^n P^n$, contraction idempotente et positive de $L^2(\mu)$, projection orthogonale sur le sous-espace $\mathcal{V}_n$ de $L^2(\mu)$ des fonctions mesurables par rapport à la sous-tribu $\tau^{-n}\mathcal{E}$, est l'espérance conditionnelle par rapport à cette tribu :

$$T^n P^n f = \mathbb{E}(f|\tau^{-n}\mathcal{E}). \tag{10.9}$$

La propriété de système exact s'interprète à l'aide de l'opérateur adjoint P. Elle est équivalente à la convergence en norme vers 0, pour chaque fonction f dans $L^2(\mu)$ d'intégrale nulle (convergence au sens fort) :

$$\lim_n \|T^n P^n f\|_2 = \lim_n \|P^n f\|_2 = 0, \forall f \in L^2(\mu), \text{ d'intégrale nulle.}$$

C'est une propriété plus forte que celle de mélange "fort" ($\lim_n \langle T^n f, g\rangle = 0$), qui n'assure que la convergence au sens faible (!) vers 0 de la suite $(T^n f)$.

La condition de convergence $P^n f \to 0$, dans $L^2(\mu)$, pour un système exact, ne comporte pas de vitesse. Pour obtenir des théorèmes limites tels que le théorème de la limite centrale, on est conduit à considérer des classes de fonctions pour lesquelles une condition de vitesse permet d'assurer la convergence du "potentiel" $\sum P^n f$.

Exemples :

1) Pour la transformation τ définie par $\tau x \to 2x \bmod 1$ sur $\mathbb{R}/\mathbb{Z}$ muni de la mesure de Lebesgue, l'opérateur adjoint P est donné par

$$P : g(x) \to \frac{1}{2}[g(x/2) + g(x/2 + 1/2)],$$

pour tout $n \geq 1$, l'opérateur $T^n P^n = \Pi_n$ est la projection orthogonale sur le sous-espace des fonctions f invariantes par le groupe dyadique d'ordre n, c'est-à-dire telles que $f(x + k2^{-n}) = f(x), \forall k = 0, 1, ..., 2^n - 1$.

Il est obtenu par moyennisation sur l'orbite du groupe dyadique d'ordre n.

Si $f \in L^2(\mathbb{T}^1)$ est une fonction "impaire" $(f(x + 1/2) = -f(x))$, elle vérifie $Pf = 0$.

2) Etant donné un alphabet fini I, considérons le décalage unilatère sur $I^{\mathbb{N}}$ muni d'une mesure produit. L'opérateur $\Pi_n = T^n P^n$ est ici la projection orthogonale sur le sous-espace des fonctions qui ne dépendent que des coordonnées d'indice $\geq n$.

Décomposition d'une fonction pour un système exact

Pour toute $f \in L^2(\mu)$, la série $\sum_{k=0}^{\infty} [\mathbb{E}(f|\tau^{-k}\mathcal{E}) - \mathbb{E}(f|\tau^{-(k+1)}\mathcal{E})]$ est une série orthogonale qui, dans le cas d'un système exact, converge en norme L^2 vers $f - \mu(f)$.

Choisissons une base orthonormée $(f_j)_{j\in J}$ de $L^2(\mathcal{E},\mu)\ominus L^2(\tau^{-1}\mathcal{E},\mu)$. On peut montrer que l'ensemble d'indices J indexant la base est infini dénombrable. Alors les fonctions $(T^n f_j)_{n\geq 0, j\in J}$ forment une base orthonormée de l'orthogonal des constantes dans $L^2(\mathcal{E},\mu)$. *Les systèmes exacts ont donc la propriété de spectre de Lebesgue.*

• K-systèmes

Supposons maintenant le système $(E,\mathcal{E},\mu,\tau)$ inversible. Rappelons que, pour toute sous-tribu $\mathcal{F}$ de $\mathcal{E}$, nous avons :

$$\mathbb{E}[f|\tau\mathcal{F}] = T^{-1}\mathbb{E}[Tf|\mathcal{F}] = \mathbb{E}[f\circ\tau|\mathcal{F}]\circ\tau^{-1}.$$

Définitions 10.3.2 On dit que le système $(E,\mathcal{E},\mu,\tau)$ est un *K-système*, s'il existe une sous-tribu $\mathcal{F}$ vérifiant les conditions suivantes :

i) $\mathcal{F}$ est croissante : $\mathcal{F}\subset\tau\mathcal{F}$,
ii) l'intersection $\bigcap_n \tau^{-n}\mathcal{F}$ est triviale,
iii) la réunion $\bigcup_n \tau^n\mathcal{F}$ engendre $\mathcal{E}$.

Posons $\mathcal{H}_0 = L^2(\mathcal{F},\mu_0)$ et $\mathcal{V}_n = T^n\mathcal{H}_0$. La condition *i*) équivaut au fait que la suite $(\mathcal{V}_n)$ est croissante, la condition *ii*) implique que l'intersection des sous-espaces $\mathcal{V}_n$ est réduite au sous-espace des constantes, enfin la condition *iii*) implique que la réunion des sous-espaces $\mathcal{V}_n$ est dense dans $L^2(\mathcal{E})$.

Nous dirons qu'une suite $(\mathcal{F}_n) = (\tau^{-n}\mathcal{F})_{n\in\mathbb{Z}}$ vérifiant les conditions *i*), *ii*) et *iii*) est une *filtration de K-système.*

Si l'on arrête la suite croissante $(\tau^n\mathcal{F})_{n\in\mathbb{Z}}$ à $n=0$, on retrouve la condition de système exact sur le "facteur" de τ^{-1} associé à la sous-tribu $\mathcal{F}$. Comme pour les systèmes exacts, on montre que les K-systèmes ont la propriété de spectre de Lebesgue.

Décomposition d'une fonction pour un K-système

Notons $\mathcal{F}_n = \tau^n\mathcal{F}$, $\mathcal{H}_n = L^2(E,\tau^{-n}\mathcal{F},\mu)$ et Π_n la projection orthogonale sur $\mathcal{H}_n$. On a $\mathcal{H}_n = T^n\mathcal{H}_0$ et $\Pi_n f = \mathbb{E}[f|\mathcal{F}_n]$.

L'opérateur unitaire T applique $\mathcal{H}_n$ sur $\mathcal{H}_{n-1}$. Les propriétés de K-système se traduisent vectoriellement par :

$$\lim_{n\to-\infty}\Pi_n f = \mu(f),\ \lim_{n\to\infty}\Pi_n f = f.$$

On obtient une "filtration vectorielle" de $L^2(\mathcal{E},\mu)$:

$$\{1\} = \bigcap_n L^2(\tau^{-n}\mathcal{F}) \subset ... \subset L^2(\tau^{-1}\mathcal{F}) \subset L^2(\mathcal{F}) \subset L^2(\tau\mathcal{F}) \subset ... \subset L^2(\mathcal{E})$$

et une décomposition de $L^2(\mathcal{E})$ en somme directe de sous-espaces deux à deux orthogonaux :

$$L^2(\mathcal{E}) = \{1\}\bigoplus_{n\in\mathbb{Z}}[L^2(\tau^{-n}\mathcal{F})\ominus L^2(\tau^{-(n+1)}\mathcal{F})]. \tag{10.10}$$

Toute fonction f dans $L^2(\mu)$ d'intégrale nulle se décompose en série orthogonale (réduite aux termes d'indices <0 si f est $\mathcal{F}$-mesurable) :

$$f = \sum_{n\in\mathbb{Z}}[\mathbb{E}(f|\tau^{-n}\mathcal{F}) - \mathbb{E}(f|\tau^{-(n+1)}\mathcal{F})]. \tag{10.11}$$

On montre que les K-système, comme les systèmes exacts, ont un spectre de Lebesgue.

10.4 Entropie d'une partition, entropie conditionnelle

On note u la fonction définie sur $[0,1]$ par $u(x) = -x\log x$, avec la convention $0\log 0 = 0$. Les logarithmes sont pris en base 2. La fonction u est strictement concave sur $[0,1]$ et sous-additive : $u(t+s) \leq u(t) + u(s)$, $\forall t, s \in [0,1]$.

Soit $p = (p_1, \cdots, p_\ell)$ un vecteur de probabilité de longueur ℓ. On appelle entropie de p le nombre

$$H(p) = -\sum_{i=1}^{\ell} p_i \log p_i = \sum_{i=1}^{\ell} u(p_i).$$

Il est clair que $H(p) = 0$ si et seulement s'il existe un indice $i_0 \in \{1, \cdots, \ell\}$ tel que $p_{i_0} = 1$ et $p_i = 0$, pour $i \neq i_0$. Par concavité de u, on a par ailleurs :

$$H(p) = \ell(\frac{1}{\ell}\sum_{i=1}^{\ell} u(p_i)) \leq \ell u(\frac{1}{\ell}\sum_{1}^{\ell} p_i) = \ell u(\frac{1}{\ell}) = \log \ell.$$

L'entropie du vecteur $p = (p_1, \cdots, p_\ell)$ est donc majorée par l'entropie du vecteur équiprobable $\{\frac{1}{\ell}, \cdots, \frac{1}{\ell}\}$ de même longueur.

La quantité $H(p)$ mesure l'"*incertitude*" attachée au vecteur p, les deux cas extrêmes étant :
$H(p) = 0$ (l'un des évènements $i_0 \in \{1, \cdots, \ell\}$ est certain) ;
$H(p) = \log \ell$ (les évènements $i \in \{1, \cdots, \ell\}$ ont une égale probabilité de se produire).

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, prenant un nombre fini ℓ de valeurs. Notons $\mathcal{R}(X)$ l'ensemble de ces valeurs. La loi de X est un vecteur de probabilité indexé par $\mathcal{R}(X)$ et on peut définir l'entropie de X comme étant l'entropie de ce vecteur, soit :

$$H(X) = -\sum_{a \in \mathcal{R}(X)} \mathbb{P}(X = a) \log \mathbb{P}(X = a).$$

Entropie d'une partition finie ou dénombrable

Considérons un espace probabilisé $(E, \mathcal{E}, \mu)$ et une partition $\mathcal{P} = \{A_1, \cdots, A_\ell\}$ de E en un nombre fini de sous-ensembles mesurables (deux à deux disjoints aux ensembles μ-négligeables près). La mesure des ensembles formant la partition $\mathcal{P}$ définit un vecteur de probabilité. L'entropie de la partition est par définition l'entropie de ce vecteur, soit :

$$H(\mathcal{P}) = -\sum_{A \in \mathcal{P}} \mu(A) \log \mu(A).$$

Les définitions précédentes s'étendent au cas dénombrable (vecteurs de probabilité, v.a. prenant un nombre dénombrable de valeurs, partition dénombrable), les sommes étant remplacées par des séries. Dans le cas dénombrable, l'entropie peut être infinie.

Entropie conditionnelle

Etant donnée une sous-tribu $\mathcal{B}$, notons $\mu(A|\mathcal{B})$ la mesure conditionnelle d'un ensemble mesurable A par rapport à $\mathcal{B}$, i.e. $\mathbb{E}[1_A|\mathcal{B}]$. L'entropie conditionnelle d'une partition $\mathcal{P}$ par rapport à $\mathcal{B}$ est définie par :

$$H(\mathcal{P}|\mathcal{B}) \quad = \quad -\int_E \sum_{A \in \mathcal{P}} \mu(A|\mathcal{B}) \log \mu(A|\mathcal{B}) \; d\mu$$

$$= \quad -\int_E \sum_{A\in\mathcal{P}} 1_A \log \mathbb{E}[1_A|\mathcal{B}] \; d\mu.$$

C'est donc l'intégrale de l'entropie de $\mathcal{P}$ calculée sous la mesure conditionnelle $\mu(\ |\mathcal{B})(.)$.

Proposition 10.4.1 *$H(\mathcal{P}|\mathcal{B}) = 0 \Longleftrightarrow \mathcal{P}$ est $\mathcal{B}$-mesurable.*

Preuve : La condition $H(\mathcal{P}|\mathcal{B}) = 0$ implique que, pour chaque élément A de $\mathcal{P}$, on a $\mathbb{E}[1_A|\mathcal{B}] = 1$, p.p. sur A.

Comme par ailleurs, on a : $0 \le \mathbb{E}[1_A|\mathcal{B}] \le 1$ et $\int \mathbb{E}[1_A|\mathcal{B}]\; d\mu = \mu(A)$, ceci implique $\mathbb{E}[1_A|\mathcal{B}] = 1_A$, μ-presque partout, et donc chaque élément A de $\mathcal{P}$ est $\mathcal{B}$-mesurable.

La réciproque est évidente. $\square$

Rappelons maintenant l'inégalité de Jensen :

Lemme 10.4.2 *Si ϕ est une fonction concave sur $[0,1]$, pour toute fonction mesurable f, telle que $0 \le f \le 1$ et toute tribu $\mathcal{B}$, on a :*

$$\mathbb{E}[\phi \circ f|\mathcal{B}](.) \le \phi(\mathbb{E}(f|\mathcal{B})(.)). \tag{10.12}$$

Si ϕ est strictement concave, l'égalité, dans l'inégalité précédente, est réalisée si et seulement si f est $\mathcal{B}$-mesurable.

Lemme 10.4.3 *Soient $\mathcal{B}$ et $\mathcal{C}$ deux tribus, avec $\mathcal{C} \subset \mathcal{B}$ et ϕ une fonction concave sur $[0,1]$. Pour toute fonction mesurable f à valeurs dans $[0,1]$, on a :*

$$\int \phi(\mathbb{E}[f|\mathcal{B}])\; d\mu \ \le \ \int \phi(\mathbb{E}[f|\mathcal{C}])\; d\mu. \tag{10.13}$$

Preuve : En appliquant l'inégalité de Jensen à la fonction concave ϕ et à $\mathbb{E}[f|\mathcal{B}]$, on obtient :

$$\mathbb{E}[\phi(\mathbb{E}[f|\mathcal{B}])|\mathcal{C}] \le \phi(\mathbb{E}[\mathbb{E}[f|\mathcal{B}]|\mathcal{C}]) = \phi(\mathbb{E}[f|\mathcal{C}]);$$

d'où le résultat en intégrant. $\square$

Rappelons que l'ordre sur l'ensemble des sous-tribus est celui de l'inclusion. A toute partition finie ou dénombrable $\mathcal{P}$ est associée la sous-tribu finie qu'elle engendre. L'ordre sur l'ensemble des partitions finies ou dénombrables est l'ordre d'inclusion des tribus associées. En d'autres termes, si $\mathcal{P}$ et $\mathcal{Q}$ sont deux partitions dénombrables de E, on écrira $\mathcal{P} \le \mathcal{Q}$, si les éléments de $\mathcal{P}$ sont des réunions d'éléments de $\mathcal{Q}$. On dit encore dans ce cas que $\mathcal{Q}$ est plus fine que $\mathcal{P}$ (ou que $\mathcal{P}$ est plus grossière que $\mathcal{Q}$).

Si $\mathcal{P}$ et $\mathcal{Q}$ sont deux partitions finies ou dénombrables, on peut former la partition notée $\mathcal{P} \vee \mathcal{Q}$ dont les éléments sont de la forme $C = A \cap B$, où A et B décrivent respectivement les éléments de $\mathcal{P}$ et de $\mathcal{Q}$. Au sens de l'ordre décrit précédemment, $\mathcal{P} \vee \mathcal{Q}$ est la borne supérieure du couple $(\mathcal{P}, \mathcal{Q})$. On a ainsi une structure de *treillis* sur l'ensemble des partitions dénombrables.

La proposition suivante rassemble les propriétés de l'entropie conditionnelle.

Proposition 10.4.4 *L'application $(\mathcal{P}, \mathcal{B}) \to H(\mathcal{P}, \mathcal{B})$ est croissante en $\mathcal{P}$ et décroissante en $\mathcal{B}$. En particulier, on a*

$$H(\mathcal{P}|\mathcal{B}) \leq H(\mathcal{P}),$$

pour toute sous-tribu $\mathcal{B}$, l'égalité n'ayant lieu que si $\mathcal{P}$ et $\mathcal{B}$ sont indépendantes (i.e. $\mu(A \cap B) = \mu(A)\ \mu(B),\ \forall A \in \mathcal{P}, B \in \mathcal{B}$).

Preuve : 1) Si la partition $\mathcal{P}$ est plus fine que la partition $\mathcal{Q}$, on a (avec $u(x) = -x \log_2 x$) :

$$u(\mu(A)) = u(\mu(\cup_{B \in \mathcal{Q}}(A \cap B)) \leq \sum_{B \in \mathcal{Q}} u(\mu(A \cap B)),$$

d'après la sous-additivité de la fonction u, ce qui montre que $H(\mathcal{P}|\mathcal{B}) \leq H(\mathcal{Q}|\mathcal{B})$.

2) En appliquant l'inégalité (10.13) aux fonctions 1_A et à la fonction u, on obtient :

$$H(\mathcal{P}|\mathcal{B}) = \sum_{A \in \mathcal{P}} \int u(\mathbb{E}[1_A|\mathcal{B}])\ d\mu \leq \sum_{A \in \mathcal{P}} \int u(\mathbb{E}[1_A|\mathcal{C}])\ d\mu = H(\mathcal{P}|\mathcal{C}).$$

Si $\mathcal{C}$ est la tribu triviale, $H(\mathcal{P}|\mathcal{C})$ se réduit à $H(\mathcal{P})$.

Si la partition $\mathcal{P}$ est indépendante d'une sous-tribu $\mathcal{B}$, on a $H(\mathcal{P}|\mathcal{B}) = H(\mathcal{P})$.

Soit, inversement, $\mathcal{B}$ une sous-tribu telle que $H(\mathcal{P}|\mathcal{B}) = H(\mathcal{P})$. Les inégalités précédentes impliquent que l'on a, pour chaque $A \in \mathcal{P}$, l'égalité

$$\int u(\mathbb{E}[1_A|\mathcal{B}])\ d\mu = u(\int 1_A\ d\mu).$$

La stricte concavité de u implique alors : $\mathbb{E}[1_A|\mathcal{B}] = \mu(A)$. La partition $\mathcal{P}$ est donc indépendante de la sous-tribu $\mathcal{B}$. $\square$

Proposition 10.4.5 *Soient $(\mathcal{B}_n)$ une suite croissante de sous-tribus et $\mathcal{B} = \bigvee \mathcal{B}_n$ la tribu engendrée par la réunion des $\mathcal{B}_n$, $n \geq 1$. Pour toute partition finie ou dénombrable d'entropie finie $\mathcal{P}$, on a :*

$$\lim_n \downarrow H(\mathcal{P}|\mathcal{B}_n) = H(\mathcal{P}|\mathcal{B}).$$

Preuve : Pour chaque élément A de $\mathcal{P}$, les fonctions $u(\mathbb{E}[1_A|\mathcal{B}_n])$ forment une suite majorée par $1/e$, d'intégrales décroissantes (inégalités de Jensen) et convergeant presque partout (théorème de convergence des martingales) vers $u(\mathbb{E}[1_A|\mathcal{B}]) = -\mathbb{E}[1_A|\mathcal{B}] \log \mathbb{E}[1_A|\mathcal{B}]$.

Si $\mathcal{P}$ est une partition finie, on peut alors appliquer le théorème de convergence dominée de Lebesgue.

Si $\mathcal{P} = (A_i, i \in I)$ est une partition dénombrable d'entropie finie, étant donné $\varepsilon > 0$, il existe un ensemble fini d'indices $J \subset I$ tel que

$$\sum_{i \notin J} \int u(\mathbb{E}[1_{A_i}|\mathcal{B}_n])\ d\mu \leq \sum_{i \notin J} u(1_{A_i}) \leq \varepsilon.$$

On se ramène ainsi au cas fini. $\square$

Conditionnement par une partition

Reprenons la définition de l'entropie conditionnelle $H(\mathcal{P}|\mathcal{B})$, quand le conditionnement est effectué par une tribu finie que nous identifions à la partition finie $\mathcal{Q}$ en ses atomes. L'entropie conditionnelle s'écrit dans ce cas :

$$\begin{aligned} H(\mathcal{P}|\mathcal{Q}) &= -\sum_{A\in\mathcal{P},B\in\mathcal{Q}} \mu(A\cap B)\ \log(\frac{\mu(A\cap B)}{\mu(B)}) \\ &= -\sum_{A\in\mathcal{P},B\in\mathcal{Q}} \mu(A\cap B)\log\mu(A\cap B) + \sum_{B\in\mathcal{Q}} \mu(B)\ \log\mu(B); \end{aligned}$$

d'où :

$$H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}\vee\mathcal{Q}) - H(\mathcal{Q}). \tag{10.14}$$

On obtient ainsi les relations :

$$H(\mathcal{P}\vee\mathcal{Q}) = H(\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q}|\mathcal{P}). \tag{10.15}$$

Ces relations impliquent notamment :

$$H(\mathcal{P}\vee\mathcal{Q}) \leq H(\mathcal{P}) + H(\mathcal{Q}), \tag{10.16}$$

avec égalité si seulement si $\mathcal{P}$ et $\mathcal{Q}$ sont indépendantes, et l'équivalence :

$$H(\mathcal{P}|\mathcal{Q}) = 0 \Longleftrightarrow \mathcal{P} \leq \mathcal{Q}.$$

Les relations (10.15) sont valables sous une forme conditionnelle. Si $\mathcal{P}$, $\mathcal{Q}$ et $\mathcal{S}$ sont trois partitions finies ou dénombrables, on a la relation :

$$H(\mathcal{P}\vee\mathcal{Q}|\mathcal{S}) = H(\mathcal{Q}|\mathcal{S}) + H(\mathcal{P}|\mathcal{Q}\vee\mathcal{S}) = H(\mathcal{P}|\mathcal{S}) + H(\mathcal{Q}|\mathcal{P}\vee\mathcal{S}). \tag{10.17}$$

Une distance sur l'espace des partitions

Notons $Z(E,\mathcal{E},\mu)$ l'ensemble des partitions dénombrables $\mathcal{P}$ de $(E,\mathcal{E})$ telle que $H(\mathcal{P}) < +\infty$. Pour $\mathcal{P}$ et $\mathcal{Q}$ dans $Z(E,\mathcal{E},\mu)$, posons

$$\delta(\mathcal{P},\mathcal{Q}) = H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{P}).$$

Nous avons l'inégalité

$$|H(\mathcal{P}) - H(\mathcal{Q})| \leq \delta(\mathcal{P},\mathcal{Q}), \tag{10.18}$$

et l'inégalité triangulaire d'après l'inégalité :

$$H(\mathcal{P}|\mathcal{Q}) \leq H(\mathcal{P}\vee\mathcal{R}|\mathcal{Q}) = H(\mathcal{R}|\mathcal{Q}) + H(\mathcal{P}|\mathcal{R}\vee\mathcal{Q}) \leq H(\mathcal{R}|\mathcal{Q}) + H(\mathcal{P}|\mathcal{R}).$$

On montre (exercice) la proposition :

Proposition 10.4.6 *δ définit une distance sur $Z(E,\mathcal{E},\mu)$. L'espace métrique $(Z(E,\mathcal{E},\mu),\delta)$ est complet et l'ensemble des partitions finies est dense dans cet espace.*

10.5 Entropie d'un système dynamique

Considérons maintenant un système dynamique $(E, \mathcal{E}, \mu, \tau)$.

Nous pouvons faire opérer τ sur l'espace des partitions dénombrables : si $\mathcal{P}$ est une partition dénombrable de E en sous-ensembles $A \in \mathcal{E}$, deux à deux disjoints, la partition $\tau^{-1}\mathcal{P}$ est formée des sous-ensembles $\tau^{-1}A, A \in \mathcal{P}$. Si p et q sont deux entiers, $0 \leq p < q$, on note $\mathcal{P}_p^q$ la partition

$$\mathcal{P}_p^q = \tau^{-p}\mathcal{P} \vee \cdots \vee \tau^{-q}\mathcal{P}. \tag{10.19}$$

En particulier, pour $n \geq 0$, on a $\mathcal{P}_0^n = \mathcal{P} \vee \tau^{-1}\mathcal{P} \cdots \vee \tau^{-n}\mathcal{P}$.

Si τ est inversible, la notation (10.19) s'étend aux indices p et q, $p < q$, quelconques.

On peut d'autre part donner un sens à la notation (10.19) avec des bornes infinies. Ainsi $\mathcal{P}_0^\infty$ est la tribu engendrée par la suite croissante de partitions $\mathcal{P} \vee \tau^{-1}\mathcal{P} \cdots \vee \tau^{-n}\mathcal{P}$, quand n tend vers $+\infty$.

Exemples : a) Si $\mathcal{P}$ est la partition de $[0, 1[$ en les intervalles $[0, \frac{1}{2}[$ et $]\frac{1}{2}, 1[$, $\mathcal{P}_0^n$ est la tribu atomique engendrée par les intervalles dyadiques d'ordre n et $\mathcal{P}_0^\infty$ est la tribu borélienne.

b) Dans le cas du décalage sur un espace produit unilatère $\Omega = I^{\mathbb{N}}$, si $\mathcal{P}$ est la partition de Ω correspondant à la première coordonnée, $\mathcal{P}_0^n$ est la tribu engendrée par les n premières coordonnées et $\mathcal{P}_0^\infty$ est la tribu borélienne produit.

Notons que l'invariance de la mesure μ par la transformation τ entraîne :

$$H(\tau^{-1}\mathcal{P}) = H(\mathcal{P}).$$

Cette propriété d'invariance et l'inégalité (10.16) assurent que la suite $(H(\mathcal{P}_0^{n-1}))$ est sous-additive, ce qui entraîne la convergence de la suite $(\frac{1}{n}H(\mathcal{P}_0^{n-1}))_{n\geq 1}$ vers $\inf_n \frac{1}{n}H(\mathcal{P}_0^{n-1})$.

Définition 10.5.1 On appelle **entropie moyenne** de la partition $\mathcal{P}$ sous l'action de τ la quantité

$$h(\mathcal{P}, \tau) := \lim_n \frac{1}{n} H(\mathcal{P}_0^{n-1}).$$

On vérifiera en exercice l'inégalité

$$|h(\mathcal{P}, \tau) - h(\mathcal{Q}, \tau)| \leq \delta(\mathcal{P}, \mathcal{Q}).$$

Proposition 10.5.2 $h(\mathcal{P}, \tau) = H(\mathcal{P}|\mathcal{P}_1^\infty)$.

Preuve :

$$\begin{aligned} H(\mathcal{P}_0^n) &= H(\tau^{-1}(\mathcal{P} \vee \cdots \vee \tau^{-n+1}\mathcal{P})) + H(\mathcal{P}|\tau^{-1}\mathcal{P} \vee \cdots \vee \tau^n\mathcal{P}) \\ &= H(\mathcal{P}_0^{n-1}) + H(\mathcal{P}|\tau^{-1}\mathcal{P}_0^{n-1}) \\ &= H(\mathcal{P}_0^{n-1}) + H(\mathcal{P}|\mathcal{P}_1^n). \end{aligned}$$

Quand n tend vers $+\infty$, l'entropie conditionnelle $H(\mathcal{P}|\mathcal{P}_1^n)$ converge en décroissant vers $H(\mathcal{P}|\mathcal{P}_1^\infty)$ (proposition 10.2.5).

On a donc :

$$\lim_n (H(\mathcal{P}_0^n) - H(\mathcal{P}_0^{n-1}) = H(\mathcal{P}|\mathcal{P}_1^\infty).$$

et a fortiori

$$\lim_n \frac{1}{n} H(\mathcal{P}_0^{n-1}) = H(\mathcal{P}|\mathcal{P}_1^\infty).$$

□

Dans la proposition qui suit, on suppose la transformation τ inversible.

Proposition 10.5.3 *Pour toutes partitions $\mathcal{P}$ et $\mathcal{Q}$ finies de E, on a :*

$$h(\mathcal{P} \vee \mathcal{Q}, \tau) \leq h(\mathcal{P}, \tau) + H(\mathcal{Q}|\mathcal{P}_{-\infty}^{\infty}).$$

Preuve : D'après les relations sur l'entropie, on a :

$$H(\mathcal{P}_0^{n-1} \vee \mathcal{Q}_0^{n-1}) = H(\mathcal{P}_0^{n-1}) + H(\mathcal{Q}_0^{n-1}|\mathcal{P}_0^{n-1}).$$

Fixons un entier p tel que $0 \leq 2p \leq n-1$. En développant, on obtient :

$$\begin{aligned} H(\mathcal{Q}_0^{n-1}|\mathcal{P}_0^{n-1}) &= H(\mathcal{Q} \vee \cdots \vee \tau^{-n+1}\mathcal{Q}|\ \mathcal{P} \vee \cdots \vee \tau^{-n+1}\mathcal{P}) \\ &\leq \sum_{k=0}^{n-1} H(\tau^{-k}\mathcal{Q}|\mathcal{P} \vee \cdots \vee \tau^{-n+1}\mathcal{P}) \\ &\leq (n-1-2p)H(\mathcal{Q}|\tau^p\mathcal{P} \vee \cdots \vee \tau^{-p}\mathcal{P}) + 2pH(\mathcal{Q}); \end{aligned}$$

d'où en divisant par n et prenant la limite quand n tend vers l'infini :

$$h(\mathcal{P} \vee \mathcal{Q}, \tau) \leq h(\mathcal{P}, \tau) + H(\mathcal{Q}|\tau^p\mathcal{P} \vee \cdots \vee \tau^{-p}\mathcal{P}),\ \forall p \geq 0.$$

On en déduit :

$$h(\mathcal{P} \vee \mathcal{Q}, \tau) \leq h(\mathcal{P}, \tau) + \lim_{p \to \infty} H(\mathcal{Q}|\tau^p \vee \cdots \vee \tau^{-p}\mathcal{P}) = h(\mathcal{P}, \tau) + H(\mathcal{Q}|\mathcal{P}_{-\infty}^{\infty}).$$

□

Si τ est non inversible, nous avons la majoration :

$$h(\mathcal{P} \vee \mathcal{Q}, \tau) \leq h(\mathcal{P}, \tau) + H(\mathcal{Q}|\mathcal{P}_0^{\infty}).$$

Le théorème de Shannon-McMillan

Etant donnée une partition finie ou dénombrable $\mathcal{P}$, notons pour $x \in E$, $\mathcal{P}(x)$ l'élément de $\mathcal{P}$ contenant un point x.

La *fonction d'incertitude* de $\mathcal{P}$ est définie sur E par : $I(\mathcal{P})(x) = -\log \mu(\mathcal{P}(x))$.

Si $\mathcal{B}$ est une sous-tribu, la fonction d'incertitude conditionnelle de $\mathcal{P}$ est définie par

$$I(\mathcal{P}|\mathcal{B})(x) = -\sum_{A \in \mathcal{P}} 1_A(x) \log(\mathbb{E}[1_A|\mathcal{B}](x)).$$

Sa valeur, en un point x de E, est donc égale à l'opposé du logarithme de la mesure conditionnelle de l'atome de $\mathcal{P}$ qui contient x. On note la relation analogue à (10.16) permettant de développer I pour l'opération $\vee$:

$$I(\mathcal{P} \vee \mathcal{Q}) = I(\mathcal{P}) + I(\mathcal{Q}|\mathcal{P}). \tag{10.20}$$

L'entropie $H(\mathcal{P})$ et l'entropie conditionnelle $H(\mathcal{P}|\mathcal{B})$ sont respectivement les intégrales des fonctions $I(\mathcal{P})$ et $I(\mathcal{P}|\mathcal{B})$.

Nous avons vu que la suite des moyennes $(\frac{1}{n}H(\mathcal{P}_0^{n-1}))$ des intégrales converge vers $h(\mathcal{P}, \tau)$, l'entropie moyenne de la partition $\mathcal{P}$ sous l'action de τ. Nous allons voir qu'il y a convergence ponctuelle de la suite $\frac{1}{n}I(\mathcal{P}_0^{n-1})$, la limite étant égale dans le cas ergodique à $h(\mathcal{P}, \tau)$.

Théorème 10.5.4 *(Shannon-McMillan) La suite* $(-\frac{1}{n}\log\mu(\mathcal{P}_0^{n-1}(.)))_{n\geq 1}$ *converge presque partout et en norme* $\| \ \|_1$ *vers* $\mathbb{E}[I(\mathcal{P}|\mathcal{P}_1^\infty)|\mathcal{J}]$.

Si le système est ergodique, la limite est égale à la constante $\int I(\mathcal{P}|\mathcal{P}_1^\infty)\ d\mu = h(\mathcal{P},\tau)$.

Montrons d'abord deux lemmes. Le deuxième est une extension du théorème de Birkhoff.

Lemme 10.5.5 *Si* $(\mathcal{B}_n)_{n\geq 0}$ *est une suite croissante de tribus, on a, pour tout ensemble mesurable* A *:*

$$\sup_n(-1_A\log\mathbb{E}[1_A|\mathcal{B}_n]) \in L^1(\mu).$$

<u>Preuve</u> : Soit a un réel > 0. Posons $C_a = \{\sup_n(-\log\mathbb{E}[1_A|\mathcal{B}_n]) > a\}$. On a :

$$C_a = \{\inf_n \mathbb{E}[1_A|\mathcal{B}_n] < e^{-a}\} = \bigcup_n C_{a,n},$$

où $C_{a,n} = \{\mathbb{E}[1_A|\mathcal{B}_n] < e^{-a}\} \cap \{\mathbb{E}[1_A|\mathcal{B}_k] \geq e^{-a},\ \forall 0 \leq k < n\}$. Ces ensembles sont $\mathcal{B}_n$-mesurables et deux à deux disjoints. On a donc :

$$\begin{aligned}\mu(A\cap C_a) = \sum_n \mu(A\cap C_{a,n}) &= \sum_n \int_{C_{a,n}} 1_A\ d\mu \\ &= \sum_n \int_{C_{a,n}} \mathbb{E}[1_A|\mathcal{B}_n]\ d\mu \\ &\leq e^{-a}\sum_n \mu(C_{a,n}) \leq e^{-a}.\end{aligned}$$

Il en résulte que, pour tout $a > 0$, $\mu\{\sup_n(1_A\log\mathbb{E}[1_A|\mathcal{B}_n]) > a\} \leq e^{-a}$. Cette condition entraîne l'intégrabilité annoncée. □

Lemme 10.5.6 *Si* (f_n) *est une suite dans* $L^1(\mu)$ *convergeant, presque partout et dans* $L^1(\mu)$*, vers une limite* f *et telle que* $\sup_n |f_n| \in L^1(\mu)$*, on a, presque partout et dans* $L^1(\mu)$ *:*

$$\lim_n \frac{1}{n}\sum_{k=0}^{n-1} f_{n-k}\circ\tau^k = \mathbb{E}[f|\mathcal{J}].$$

<u>Preuve</u> : Dans la majoration

$$|\frac{1}{n}\sum_{k=0}^{n-1} f_{n-k}\circ\tau^k - \mathbb{E}[f|\mathcal{J}]| \ \leq \ |\frac{1}{n}\sum_{k=0}^{n-1} f\circ\tau^k - \mathbb{E}[f|\mathcal{J}]| + \frac{1}{n}\sum_{k=0}^{n-1}|f_{n-k}-f|\circ\tau^k,$$

le premier terme à droite tend vers 0, d'après le théorème de Birkhoff.

Pour majorer le deuxième terme, posons $g_L = \sup_{n\geq L}|f_n - f|$. Les fonctions g_L sont intégrables et la suite (g_L) tend vers 0, presque partout et dans L^1, d'après les hypothèses sur la suite (f_n).

En appliquant à nouveau le théorème de Birkhoff, on obtient :

$$\limsup_n \frac{1}{n}\sum_0^{n-1}|f_{n-k}-f|\circ\tau^k \leq \lim_n \frac{1}{n}\sum_{k=0}^{n-1} g_L\circ\tau^k = \mathbb{E}[g_L|\mathcal{J}],\ \forall L. \tag{10.21}$$

La fonction limite à droite dans (10.21) a pour intégrale $\int g_L\ d\mu$, quantité qui tend vers 0, d'après le théorème de Lebesgue. Donc la limite supérieure à gauche, dans (10.21), majorée par des fonctions d'intégrale arbitrairement petite, est nulle. □

Preuve du théorème 10.5.4 : En développant "l'incertitude conditionnelle", nous obtenons :

$$\begin{aligned} I(\mathcal{P}_0^n) &= I(\mathcal{P}|\mathcal{P}_1^n) + I(\tau^{-1}\mathcal{P}|\mathcal{P}_2^n) + I(\tau^{-2}\mathcal{P}|\mathcal{P}_3^n) + \cdots + I(\tau^{-n}\mathcal{P}) \\ &= I(\mathcal{P}|\mathcal{P}_1^n) + I(\mathcal{P}|\mathcal{P}_1^{n-1}) \circ \tau + I(\mathcal{P}|\mathcal{P}_1^{n-2}) \circ \tau^2 + \cdots + I(\mathcal{P}). \end{aligned}$$

Nous avons la convergence presque partout et dans L^1 de la suite $I(\mathcal{P}|\mathcal{P}_1^n)$ vers $I(\mathcal{P}|\mathcal{P}_1^\infty)$, d'après le théorème des martingales.

Compte-tenu du lemme 10.5.5, on obtient le résultat en appliquant le lemme 10.5.6 à cette suite. □

Définition 10.5.7 On appelle **entropie du système dynamique** $(E, \mathcal{E}, \mu, \tau)$ la quantité

$$h(\tau) = \sup_{\mathcal{P}} h(\mathcal{P}, \tau),$$

la borne supérieure étant prise sur l'ensemble des partitions finies de E.

L'entropie est une fonction de la mesure μ et de la transformation τ préservant μ. Pour indiquer la dépendance en μ, on notera aussi $h(\mu, \tau)$ ou $h_\mu(\tau)$ l'entropie.

Montrons que l'entropie d'un système décroit par passage à un facteur du système (cf. 8.4.1 pour la notion de facteur).

Théorème 10.5.8 *Si un système dynamique $(E', \mathcal{E}', \mu', \tau')$ est un facteur d'un système dynamique $(E, \mathcal{E}, \mu, \tau)$, on a : $h(\tau') \leq h(\tau)$.*

Preuve : Soit ϕ l'application de E dans E' qui applique l'espace mesuré $(E, \mathcal{E}, \mu)$ sur $(E', \mathcal{E}', \mu')$. On a le diagramme commutatif :

$$\begin{array}{ccc} (E, \mathcal{E}, \mu) & \xrightarrow{\tau} & (E, \mathcal{E}, \mu) \\ \phi \big\downarrow & & \phi \big\downarrow \\ (E', \mathcal{E}', \mu') & \xrightarrow{\tau'} & (E', \mathcal{E}', \mu'). \end{array}$$

Quand $\mathcal{P}'$ décrit les partitions finies de E', $\phi^{-1}\mathcal{P}'$ décrit un sous-ensembles de l'ensemble des partitions finies de E. Il en résulte que

$$h(\tau') = \sup_{\mathcal{P}'} h(\mathcal{P}', \tau') = \sup_{\mathcal{P}'} h(\phi^{-1}\mathcal{P}', \tau) \leq \sup_{\mathcal{P}} h(\mathcal{P}, \tau).$$

□

Corollaire 10.5.9 *Deux systèmes dynamiques isomorphes ont la même entropie.*

L'entropie des systèmes dynamiques est ainsi un invariant de conjugaison. Son introduction en théorie ergodique en 1958 par A. Kolmogorov et Ya. Sinaï a permis de résoudre le problème longtemps ouvert de la conjugaison des schémas de Bernoulli.

La notion de partition génératrice réduit le calcul de l'entropie, effectué a priori sur l'ensembles des partitions finies, à une seule partition.

Définition 10.5.10 Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique. Une partition finie ou dénombrable $\mathcal{P}$ est dite *fortement génératrice*, si la suite des partitions $\mathcal{P} \vee \tau^{-1}\mathcal{P} \cdots \vee \tau^{-n}\mathcal{P}$, $n \geq 1$, engendre la tribu $\mathcal{E}$ (autrement dit $\mathcal{P}_0^\infty = \mathcal{E}$).

Si la transformation τ est inversible, la partition $\mathcal{P}$ est dite *génératrice* (au sens bilatère), si la suite des partitions $\tau^{-n}\mathcal{P} \cdots \vee \tau^{-n}\mathcal{P}$, $n \geq 1$, engendre la tribu $\mathcal{E}$ (autrement dit $\mathcal{P}_{-\infty}^\infty = \mathcal{E}$).

De la proposition 10.5.3, on déduit immédiatement le corollaire suivant :

Théorème 10.5.11 *Si $\mathcal{P}$ est une partition fortement génératrice pour un système dynamique $(E, \mathcal{E}, \mu, \tau)$, ou génératrice (dans le cas où τ est inversible), on a $h(\tau) = h(\mathcal{P}, \tau)$.*

Exemple Soit $p = (p_1, \ldots, p_{|I|})$ un vecteur de probabilité, I étant un alphabet fini ou dénombrable. Considérons le schéma de Bernoulli unilatère (resp. bilatère) correspondant sur $I^{\mathbb{N}}$ (resp. $I^{\mathbb{Z}}$) muni de la mesure de probabilité produit $\mathbb{P}_p$ et du décalage θ. La partition $\mathcal{P} = \{X_0^{-1}(a), a \in I\}$ est fortement génératrice (resp. génératrice) et par indépendance, $h(\mathcal{P}, \theta) = H(p)$. On a donc $h(\mathbb{P}_p, \theta) = H(p)$, d'après le théorème 10.5.12.

Corollaire 10.5.12 *Si p et q sont deux vecteurs de probabilité tels que $H(p) \neq H(q)$, les schémas de Bernoulli construits sur les vecteurs de probabilité p et q ne sont pas isomorphes.*

Le théorème d'Ornstein (? ?) établit la réciproque : *si $H(p) = H(q)$, les schémas de Bernoulli correspondants sont isomorphes.*

• **Partition génératrice et codage**

Un codage d'un système inversible $(E, \mathcal{A}, \mu, \tau)$ est un isomorphisme entre $(E, \mathcal{A}, \mu, \tau)$ et le système formé par $I^{\mathbb{Z}}$ muni d'une mesure stationnaire sur lequel opère le décalage θ. Si le système est non inversible, on cherche à réaliser cet isomorphisme entre le système et le décalage opérant sur l'espace des suites unilatères $I^{\mathbb{N}}$.

On construit des codages en se donnant une partition $\mathcal{P}$ finie ou dénombrable en ensembles mesurables A_i, $i \in I$, de E. Soit $\Omega = I^{\mathbb{N}}$. La suite $\omega(x) = (\omega_n(x))$ telle que

$$T^n x \in A_{\omega_n(x)}, n \in \mathbb{N},$$

définit une application de E dans Ω. Si l'application est injective, on dit que la partition $\mathcal{P}$ est **génératrice**. On peut montrer que cette définition équivaut à (10.5.10).

La construction de représentations symboliques des systèmes dynamiques est importante pour des raisons théoriques et pratiques, comme le montre l'exemple le plus ancien qui est celui de la représentation des nombres réels dans une base entière. Par exemple, le système $x \to 2x \bmod 1$, correspondant au développement en base 2, est codé à l'aide de la partition en les deux intervalles $A_0 = [0, \frac{1}{2}[$, $A_1 = [\frac{1}{2}, 1[$. Le système symbolique associé est le décalage unilatère sur l'espace $\{0, 1\}^{\mathbb{N}}$ muni de la mesure produit $\bigotimes_{\mathbb{N}}(\frac{1}{2}, \frac{1}{2})$.

Chapitre 11

TCL pour des accroissements de martingale, applications

Dans ce chapitre nous montrons le théorème central limite pour les accroissements de martingales, sous la forme donnée par B.M. Brown, et nous l'appliquons aux processus stationnaires et markoviens et aux systèmes dynamiques.

11.1 Le théorème de la limite centrale pour les martingales

Notations 11.1.1 Pour $\sigma > 0$, nous notons ϕ_σ la fonction de répartition de la loi normale $\mathcal{N}(0, \sigma^2)$, en convenant que ϕ_0 est la fonction de répartition de la mesure de Dirac en 0.

La convergence en loi (en distribution) sous une mesure de probabilité $\mathbb{P}$ d'une suite de v.a. (Z_n) vers la loi normale centrée de variance σ^2 sera notée : $Z_n \overset{d/\mathbb{P}/}{\to} \mathcal{N}(0, \sigma^2)$.

Rappelons que, pour $\sigma \neq 0$, on a alors $\lim_{n\to+\infty} \ \sup_{\alpha\in\mathbb{R}} |\mathbb{P}(\sigma^{-1} Z_n \leq \alpha) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-\frac{u^2}{2}} \, du| = 0$.

Définitions 11.1.2 Rappelons qu'étant donnés un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, un processus $(S_n)_{n\geq 0}$ à valeurs réelles et une filtration $(\mathcal{F}_n)_{n\geq 0}$ définis sur $(\Omega, \mathcal{F}, \mathbb{P})$, on dit que $(S_n, \mathcal{F}_n)_{n\geq 0}$ est une *martingale de carré intégrable* si les conditions suivantes sont satisfaites :

i) $\mathbb{E}[|S_n|^2] < +\infty, \ \forall n \geq 0$,

ii) X_n est $\mathcal{F}_n$-mesurable, $\forall n \geq 0$,

ii) $\mathbb{E}[S_{n+1}|\mathcal{F}_n] = S_n$, $\forall n \geq 0$.

A une martingale $(S_n, \mathcal{F}_n)_{n\geq 0}$, on associe le processus $(X_n)_{n\geq 1}$ formé des *accroissements* : $X_0 = S_0$, $X_n = S_n - S_{n-1}$, pour $n \geq 1$. On a donc $S_n = S_0 + X_1 + \ldots + X_n$, $\forall n \geq 0$.

La condition *iii)* de martingale se traduit par

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] = 0, \forall n \geq 1.$$

En particulier, on a $\mathbb{E}[S_n] = \mathbb{E}[S_0]$, $\forall n \geq 0$.

A une martingale $(S_n, \mathcal{F}_n)_{n\geq 0}$ de carré intégrable, on associe le *processus croissant* $(V_n, n \geq 0)$ défini par $V_0 = 0$ et

$$V_n = \sum_{k=1}^{n} \mathbb{E}[X_k^2|\mathcal{F}_{k-1}],\ \forall n \geq 1.$$

On note que, par construction, $(S_n^2 - V_n, \mathcal{F}_n)_{n\geq 0}$ est une martingale. Le processus $(S_n)_{n\geq 0}$ est à accroissements orthogonaux et on a, en posant $\sigma_n = \sqrt{\mathbb{E}[S_n^2]}$,

$$\sigma_n^2 = \mathbb{E}[S_n^2] = \mathbb{E}[S_0^2] + \sum_{k=1}^{n} \mathbb{E}[X_k^2] = \mathbb{E}[S_0^2] + \mathbb{E}[V_n].$$

Nous commençons par un lemme de majoration.

Lemme 11.1.3 *Soient n un entier ≥ 1, $(\mathcal{F}_0, ..., \mathcal{F}_n)$ une famille croissante de sous-tribus de $\mathcal{F}$ et $Y_1, ..., Y_n$ des variables aléatoires de carré intégrable telles que Y_k soit $\mathcal{F}_k$-mesurable et $\mathbb{E}[Y_k|\mathcal{F}_{k-1}] = 0$, pour $1 \leq k \leq n$. Posons*

$$T_k = Y_1 + \ldots + Y_k,\ W_k = \sum_{j=1}^{k} \mathbb{E}[Y_j^2|\mathcal{F}_{j-1}],\ 1 \leq k \leq n.$$

Sous l'hypothèse $\mathbb{P}(W_n \leq C) = 1$, où C est un réel > 0, nous avons, pour tout $\varepsilon > 0$,

$$|\mathbb{E}[e^{iT_n + \frac{1}{2}W_n}] - 1| \leq e^{\frac{1}{2}C}((\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})C + (1 + \frac{C}{8}) \sum_{1\leq k\leq n} \mathbb{E}[Y_k^2 1_{\{|Y_k|>\varepsilon\}}]).$$

<u>Preuve</u> : Les développements de Taylor avec reste intégral à l'ordre 2 et 3 permettent d'écrire

$$\forall x \in \mathbb{R},\ \exp(ix) = 1 + ix - \frac{1}{2}x^2 + u(x),\ \text{avec } |u(x)| \leq \inf(x^2, \frac{|x|^3}{6}). \tag{11.1}$$

On a également :

$$e^{-x} = 1 - x + v(x),\ \text{avec } \forall x \geq 0,\ 0 \leq v(x) \leq \frac{x^2}{2}.$$

Posons, pour $1 \leq k \leq n$, $R_k = e^{iT_k + \frac{1}{2}W_k}$ et $Z_k = R_k - R_{k-1}$, en convenant que $R_0 = 1$.

Nous avons :

$$\begin{aligned} Z_k &= e^{iT_{k-1}+\frac{1}{2}W_k}(e^{iY_k} - e^{-\frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}]}) \\ &= e^{iT_{k-1}+\frac{1}{2}W_k}(iY_k - \frac{1}{2}Y_k^2 + u(Y_k) + \frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}] - v(\frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}])); \end{aligned}$$

d'où, en notant que les variables aléatoires T_{k-1} et W_k sont $\mathcal{F}_{k-1}$-mesurables,

$$|\mathbb{E}[Z_k|\mathcal{F}_{k-1}]| \leq e^{\frac{1}{2}W_k}(|\mathbb{E}[u(Y_k)|\mathcal{F}_{k-1}]| + v(\frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}])).$$

On a alors :

$$\begin{aligned} |\mathbb{E}[Z_k]| &\leq \mathbb{E}[|\mathbb{E}[Z_k|\mathcal{F}_{k-1}]|] \\ &\leq e^{\frac{1}{2}C}(\mathbb{E}[|u(Y_k)|] + \mathbb{E}[v(\frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}])]). \end{aligned}$$

D'après (11.1), le premier terme $\mathbb{E}[|u(Y_k)|]$ est majoré, pour tout $\varepsilon > 0$, par

$$\mathbb{E}(Y_k^2 1_{\{|Y_k|>\varepsilon\}}) + \frac{\varepsilon}{6}\mathbb{E}(Y_k^2)$$

et le second terme $\mathbb{E}[v(\frac{1}{2}\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}])]$ par

$$\begin{aligned}\frac{1}{8}\mathbb{E}[(\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}])^2] &\leq \frac{1}{8}\mathbb{E}[\mathbb{E}[Y_k^2|\mathcal{F}_{k-1}](\varepsilon^2 + \mathbb{E}[Y_k^2 1_{\{|Y_k|>\varepsilon\}}|\mathcal{F}_{k-1}])] \\ &\leq \frac{1}{8}\varepsilon^2\mathbb{E}[Y_k^2] + \frac{C}{8}\mathbb{E}[Y_k^2 1_{\{|Y_k|>\varepsilon\}}].\end{aligned}$$

On en déduit que :

$$\begin{aligned}|\mathbb{E}[R_n] - 1| &\leq \sum_{k=1}^{n} |\mathbb{E}[Z_k]| \\ &\leq e^{\frac{1}{2}C}((\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})\mathbb{E}[W_n] + (1 + \frac{1}{8}C)\sum_{1\leq k\leq n} \mathbb{E}[Y_k^2 1_{\{|Y_k|>\varepsilon\}}]).\end{aligned}$$

□

Théorème 11.1.4 *(B.M. Brown [Brow71]) Soit $(S_n, \mathcal{F}_n)_{n\geq 0}$ une martingale non identiquement nulle définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ de carré intégrable, d'accroissements $X_n = S_n - S_{n-1}$, $n \geq 1$. Soit $\sigma_n^2 = \mathbb{E}[S_n^2]$. Supposons vérifiées les conditions suivantes :*

i) $\sigma_n^{-2}V_n \to 1$, en probabilité,

ii) $\forall\varepsilon > 0$, $\lim_{n\to\infty} \sigma_n^{-2}\sum_{k=1}^{n}\mathbb{E}[X_k^2 1_{\{|X_k|>\varepsilon\sigma_n\}}] = 0$ (condition de Lindeberg).

Alors on a : $\sigma_n^{-1}S_n \overset{d/\mathbb{P}}{\to} \mathcal{N}(0,1)$.

<u>Preuve</u> : La condition *ii)* impliquant que $\lim_{n\to\infty} \sigma_n = +\infty$, on peut supposer que $S_0 = 0$.

1) *Cas borné* : Nous supposons qu'il existe un réel $C > 0$ tel que

$$\sigma_n^{-2}V_n \leq C, \mathbb{P} - \text{p.s.}, \forall n \geq 0.$$

Pour $n \geq 1$, appliquons le lemme précédent aux v.a. $Y_{k,n} = \frac{tX_k}{\sigma_n}$, $1 \leq k \leq n$, avec $t \in \mathbb{R}$. Les $Y_{k,n}$ vérifient les hypothèses du lemme et nous avons

$$\sum_{k=1}^{n} \mathbb{E}[Y_{k,n}^2|\mathcal{F}_{k-1}] = t^2\sigma_n^{-2}V_n \leq Ct^2, \mathbb{P} - \text{p.s.}$$

On obtient donc, pour tout $\varepsilon > 0$, la majoration

$$\begin{aligned}&|\mathbb{E}[e^{it\sigma_n^{-1}S_n}\, e^{\frac{t^2}{2}\sigma_n^{-2}V_n}] - 1| \\ &\leq e^{C\frac{t^2}{2}}(C(\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})t^2 \ + (1 + \frac{C}{8}t^2)\sum_{1\leq k\leq n} \frac{t^2}{\sigma_n^2}\mathbb{E}[X_k^2 1_{\{|t||X_k|>\varepsilon\sigma_n\}}]).\end{aligned}$$

De cette majoration et de l'hypothèse *ii)*, on déduit que

$$\lim_{n\to\infty} \mathbb{E}[e^{it\sigma_n^{-1}S_n}\, e^{\frac{t^2}{2}\sigma_n^{-2}V_n}] = 1.$$

L'hypothèse $i)$ et le fait que la suite $(\sigma_n^{-2}V_n)$ soit bornée permettent d'obtenir la convergence en norme $\| \|_1$ de la suite $(e^{\frac{t^2}{2\sigma_n^2}V_n})$ vers la constante $e^{\frac{t^2}{2}}$.

(En effet, pour t fixé, pour $\varepsilon > 0$, soit $\delta(\varepsilon) > 0$ tel que : $|x-1| < \delta(\varepsilon) \Rightarrow |e^{\frac{t^2}{2}x} - e^{\frac{t^2}{2}}| < \varepsilon$. On a :

$$\begin{aligned}
\mathbb{E}[|e^{\frac{t^2 V_n}{2\sigma_n^2}} - e^{\frac{t^2}{2}}|] &= \mathbb{E}[|e^{\frac{t^2 V_n}{2\sigma_n^2}} - e^{\frac{t^2}{2}}|1_{\{|\frac{V_n}{\sigma_n}-1|<\delta\}}] + \mathbb{E}[|e^{\frac{t^2 V_n}{2\sigma_n^2}} - e^{\frac{t^2}{2}}|1_{\{|\frac{V_n}{\sigma_n}-1|\geq\delta\}}] \\
&\leq \varepsilon + (e^{\frac{t^2}{2}} + e^{C\frac{t^2}{2}})\mathbb{P}[\{|\frac{V_n}{\sigma_n} - 1| \geq \delta(\varepsilon)\}].
\end{aligned}$$

On a donc :

$$|\mathbb{E}[e^{i\frac{tS_n}{\sigma_n}}(e^{\frac{t^2 V_n}{2\sigma_n^2}} - e^{\frac{t^2}{2}})]| \leq \mathbb{E}[|e^{\frac{t^2 V_n}{2\sigma_n^2}} - e^{\frac{t^2}{2}}|] \to 0.)$$

D'où la convergence de la fonction caractéristique $\mathbb{E}[e^{i\frac{tS_n}{\sigma_n}}]$ vers $e^{-\frac{t^2}{2}}$, ce qui termine la démonstration du théorème dans le cas borné.

2) *Cas général*

Montrons maintenant comment on peut se ramener par troncature au cas où la suite $(\sigma_n^{-2}V_n)_{n\geq 0}$ est bornée, $\mathbb{P}$-p.s..

Soit C un réel > 1. Pour k tel que $1 \leq k \leq n$, posons $Y_{k,n} := \frac{t}{\sigma_n}X_k 1_{\{V_k \leq C\sigma_n^2\}}$.

Pour tout $k \geq 1$, V_k étant $\mathcal{F}_{k-1}$ mesurable, on a encore $\mathbb{E}[Y_{k,n}|\mathcal{F}_{k-1}] = 0$, $\forall 1 \leq k \leq n$. En reprenant les notations du lemme, définissons les v.a.

$$\begin{aligned}
T_n &:= \sum_{k=1}^n Y_{k,n} = \frac{t}{\sigma_n}\sum_{k=1}^n X_k 1_{\{V_k \leq C\sigma_n^2\}} \\
W_n &:= \sum_{k=1}^n \mathbb{E}[Y_{k,n}^2|\mathcal{F}_{k-1}].
\end{aligned}$$

La v.a. W_n vérifie

$$W_n = \frac{t^2}{\sigma_n^2}\sum_{k=1}^n \mathbb{E}[X_k^2 1_{\{V_k \leq C\sigma_n^2\}}|\mathcal{F}_{k-1}] = \frac{t^2}{\sigma_n^2}\sum_{k=1}^n \mathbb{E}[X_k^2|\mathcal{F}_{k-1}]1_{\{V_k \leq C\sigma_n^2\}} \leq Ct^2.$$

On déduit du lemme que :

$$\begin{aligned}
&|\mathbb{E}[e^{iT_n}\, e^{\frac{W_n}{2}}] - 1| \\
&\leq e^{C\frac{t^2}{2}}((\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})Ct^2 + (1 + \frac{C}{8}t^2)\frac{t^2}{\sigma_n^2}\sum_{1\leq k\leq n}\mathbb{E}[X_k^2 1_{\{|t||X_k| > \varepsilon\sigma_n\}}].
\end{aligned}$$

car

$$\mathbb{E}[Y_{k,n}^2 1_{\{|Y_{k,n}|>\varepsilon\}}] \leq \frac{t^2}{\sigma_n^2}\mathbb{E}[X_k^2 1_{\{|t||X_k|>\varepsilon\sigma_n\}}.$$

Soient $\widetilde{S}_n = \sum_{k=1}^n X_k 1_{\{V_k \leq C\sigma_n^2\}}$, et $\widetilde{V}_n = \sum_{k=1}^n \mathbb{E}[X_k^2|\mathcal{F}_{k-1}]1_{\{V_k \leq C\sigma_n^2\}}$.

Les inégalités $V_n 1_{\{V_n \leq C\sigma_n^2\}} \leq \widetilde{V}_n \leq V_n$ montrent, compte-tenu de l'hypothèse $i)$, que $(\widetilde{V}_n/\sigma_n^2)$ converge en probabilité vers 1.

Comme dans le premier cas, on en déduit que la suite de v.a. $(\widetilde{S}_n/\sigma_n)$ converge en loi vers la loi $\mathcal{N}(0,1)$.

Il reste à comparer $\widetilde{S}_n$ et S_n. L'inclusion $\{S_n \neq \widetilde{S}_n\} \subset \{V_n > C\sigma_n^2\}$ montre que $\frac{S_n}{\sigma_n} - \frac{\widetilde{S}_n}{\sigma_n} \to 0$ en probabilité.

D'où le résultat dans le cas général. □

Donnons maintenant une forme un peu plus générale du résultat précédent, adaptée au cas d'un théorème de la limite centrale avec mélange de lois gaussiennes.

Théorème 11.1.5 *Soit $(S_n, \mathcal{F}_n)_{n\geq 0}$ une martingale de carré intégrable, non identiquement nulle, définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, d'accroissements $X_n = S_n - S_{n-1}$, $n \geq 1$. Sous les conditions*

i) La suite $\sigma_n^{-2} V_n$ converge en probabilité vers une v.a. U,

ii) $\forall \varepsilon > 0$, $\lim_{n\to\infty} \sigma_n^{-2} \sum_{k=1}^{n} \mathbb{E}(X_k^2 1_{\{|X_k|>\varepsilon\sigma_n\}}) = 0$, (condition de Lindeberg),

nous avons, pour tout $C > 0$:

$$\limsup_n |1 - \mathbb{E}[e^{it\frac{S_n}{\sigma_n} + \frac{1}{2}t^2 U} 1_{\{U<C\}}]| \leq e^{\frac{1}{2}Ct^2} \mathbb{P}(U \geq C), \ \forall t \in \mathbb{R}.$$

En particulier, si U est telle que, pour tout $t \in \mathbb{R}$, $\mathbb{E}[e^{\frac{1}{2}t^2 U}] < \infty$, alors on a :

$$\lim_n \mathbb{E}[e^{it\frac{S_n}{\sigma_n} + \frac{1}{2}t^2 U}] = 1, \ \forall t \in \mathbb{R}.$$

Si $U \equiv 1$, on a la convergence vers la loi normale.

Preuve : La preuve est analogue à celle du théorème 11.1.4, mais nous procédons ici en traitant directement le processus tronqué. La condition $ii)$ impliquant que $\lim_{n\to\infty} \sigma_n = +\infty$, on peut supposer que $S_0 = 0$.

Soit C un réel > 0. Pour $1 \leq k \leq n$, posons $Y_{k,n} := \frac{t}{\sigma_n} X_k 1_{\{V_k \leq C\sigma_n^2\}}$.

Pour tout $k \geq 1$, V_k étant $\mathcal{F}_{k-1}$ mesurable, on a encore $\mathbb{E}[Y_{k,n}|\mathcal{F}_{k-1}] = 0$, $\forall 1 \leq k \leq n$. Comme précédemment, avec les notations du lemme, nous définissons, pour $1 \leq k \leq n$,

$$\begin{aligned} T_n &:= \sum_{k=1}^{n} Y_{k,n} = \frac{t}{\sigma_n} \sum_{k=1}^{n} X_k 1_{\{V_k \leq C\sigma_n^2\}} \\ W_n &:= \sum_{k=1}^{n} \mathbb{E}[Y_{k,n}^2 | \mathcal{F}_{k-1}]. \end{aligned}$$

La suite (V_k) étant croissante, on a :

$$W_n = \frac{t^2}{\sigma_n^2} \sum_{k=1}^{n} \mathbb{E}[X_k^2 1_{\{V_k \leq C\sigma_n^2\}} | \mathcal{F}_{k-1}] = \frac{t^2}{\sigma_n^2} \sum_{k=1}^{n} \mathbb{E}[X_k^2 | \mathcal{F}_{k-1}] 1_{\{V_k \leq C\sigma_n^2\}} = Ct^2.$$

On déduit du lemme 11.1.3 la majoration :

$$\begin{aligned} &|\mathbb{E}[e^{iT_n + \frac{1}{2}W_n}] - 1| \\ &\leq e^{C\frac{t^2}{2}}\Big(\Big(\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8}\Big)Ct^2 + \Big(1 + \frac{C}{8}t^2\Big)\frac{t^2}{\sigma_n^2} \sum_{1\leq k\leq n} \mathbb{E}[X_k^2 1_{\{|t||X_k|>\varepsilon\sigma_n\}}]\Big), \end{aligned}$$

car

$$\mathbb{E}[Y_{k,n}^2 1_{\{|Y_{k,n}|>\varepsilon\}}] \leq \frac{t^2}{\sigma_n^2} \mathbb{E}[X_k^2 1_{\{|t||X_k|>\varepsilon\sigma_n\}}];$$

d'où :

$$\lim_n \mathbb{E}[e^{iT_n+\frac{1}{2}W_n}] = 1.$$

On a la majoration :

$$\begin{aligned}&|\mathbb{E}[e^{iT_n+\frac{1}{2}W_n}] - \mathbb{E}[e^{iT_n+\frac{t^2}{2}U}1_{\{U<C\}}]| \\ &\quad \leq \mathbb{E}[e^{\frac{1}{2}W_n}1_{\{U\geq C\}}] + \mathbb{E}[|e^{\frac{1}{2}W_n} - e^{\frac{t^2}{2}U}|1_{\{U<C\}}].\end{aligned} \tag{11.2}$$

Le premier terme de (11.2) est majoré par $e^{\frac{1}{2}Ct^2}\mathbb{P}(U \geq C)$. Etudions le deuxième terme.

On a l'encadrement

$$\frac{t^2}{\sigma_n^2}V_n 1_{\{V_n\leq C\sigma_n^2\}} \leq W_n \leq \frac{t^2}{\sigma_n^2}V_n$$

et la convergence en probabilité de (V_n/σ_n^2) vers U impliquent la convergence en probabilité de $(W_n 1_{\{U<C\}})$ vers $t^2U1_{\{U<C\}}$. Pour $\varepsilon > 0$, soit $\delta(\varepsilon) > 0$ tel que, les conditions $|x-y| < \delta$, $|x|, |y| \leq \frac{1}{2}Ct^2$ impliquent $|e^x - e^y| < \varepsilon$. On a alors, pour n assez grand :

$$\mathbb{E}[|e^{\frac{1}{2}W_n} - e^{\frac{t^2}{2}U}|1_{\{U<C\}}] \leq \varepsilon + 2e^{\frac{1}{2}Ct^2}\mathbb{P}(\{|W_n - t^2U| \geq \delta(\varepsilon)\}\bigcap\{U<C\}) < 2\varepsilon.$$

Ceci montre que le deuxième terme de (11.2) tend vers zéro, quand n tend vers l'infini. En remplaçant $\mathbb{E}[e^{iT_n+\frac{1}{2}W_n}]$ par sa limite 1, on obtient ainsi :

$$\limsup_n |1 - \mathbb{E}[e^{iT_n+\frac{t^2}{2}U}1_{\{U<C\}}]| \leq e^{\frac{1}{2}Ct^2}\mathbb{P}(U\geq C).$$

D'autre part, compte-tenu de l'inclusion $\{T_n \neq \frac{t}{\sigma_n}S_n\} \subset \{V_n > C\sigma_n^2\}$, on a la majoration :

$$\begin{aligned}|\mathbb{E}[(e^{iT_n+\frac{1}{2}t^2U} - e^{it\frac{S_n}{\sigma_n}+\frac{1}{2}t^2U})1_{\{U<C\}}]| &\leq 2\mathbb{E}[e^{\frac{1}{2}t^2U}1_{\{V_n>C\sigma_n^2\}\bigcap\{U<C\}}] \\ &\leq 2e^{\frac{1}{2}Ct^2}\mathbb{P}(\{V_n > C\sigma_n^2\}\bigcap\{U<C\}) \to 0;\end{aligned}$$

d'où :

$$\limsup_n |1 - \mathbb{E}[e^{it\frac{S_n}{\sigma_n}+\frac{1}{2}t^2U}1_{\{U<C\}}]| \leq e^{\frac{1}{2}Ct^2}\mathbb{P}(U\geq C),\ \forall t \in \mathbb{R}.$$

Sous la condition $\mathbb{E}[e^{\frac{1}{2}t^2U}] < \infty, \forall t \in \mathbb{R}$, on peut majorer l'expression précédente par $\mathbb{E}[e^{\frac{1}{2}t^2U}1_{\{U\geq C\}}]$ et on obtient, pour toute constante $C > 0$, la majoration :

$$\limsup_n |1 - \mathbb{E}[e^{it\frac{S_n}{\sigma_n}+\frac{1}{2}t^2U}]| \leq 2\mathbb{E}[e^{\frac{1}{2}t^2U}1_{\{U\geq C\}}],$$

qui implique

$$\lim_n \mathbb{E}[e^{it\frac{S_n}{\sigma_n}+\frac{1}{2}t^2U}] = 1.$$

□

Définition 11.1.6 Soit $(Y_n)_{n\geq 0}$ un processus à valeurs réelles et $(\mathcal{G}_n)_{n\geq 0}$ une suite décroissante de sous-tribus de $\mathcal{F}$. On dit que $((Y_n)_{n\geq 0}, (\mathcal{G}_n)_{n\geq 0})$ est une *suite d'accroissements de martingale renversée de carré intégrable* si, pour tout entier $n \geq 0$:

i) $\mathbb{E}[|Y_n|^2] < +\infty$;
ii) la v.a. Y_n est $\mathcal{G}_n$-mesurable ;
iii) $\mathbb{E}[Y_n|\mathcal{G}_{n+1}] = 0$.

Théorème 11.1.7 *Soit $((Y_n)_{n\geq 0}, (\mathcal{G}_n)_{n\geq 0})$ une suite d'accroissements de martingale renversée de carré intégrable. Pour tout entier $n \geq 0$, nous posons :*

$$S_n := Y_0 + \ldots + Y_{n-1}, \quad \sigma_n^2 = \sum_{k=0}^{n-1} \mathbb{E}[Y_k^2] \text{ et } V_n = \sum_{k=0}^{n-1} \mathbb{E}[Y_k^2|\mathcal{G}_{k+1}].$$

Supposons que les deux conditions suivantes soient satisfaites :

i) la suite de v.a.r. $(\sigma_n^{-2}\ V_n)_{n\geq 1}$ converge en probabilité vers 1,

ii) pour tout $\varepsilon > 0$, $\lim_{n\to+\infty} \sigma_n^{-2} \sum_{k=0}^{n-1} \mathbb{E}[Y_k^2 1_{\{|Y_k|>\varepsilon\sigma_n\}}] = 0$,

Alors nous avons $\sigma_n^{-1} S_n \overset{d/\mathbb{P}}{\to} \mathcal{N}(0,1)$.

Preuve : Soit C un réel > 1. Pour tout entier $0 \leq k \leq n$, posons

$$Z_k = \sum_{j=k}^{n-1} \mathbb{E}[Y_j^2|\mathcal{G}_{j+1}], U_k = \frac{t}{\sigma_n} Y_{n-k} 1_{\{Z_{n-k}\leq C\sigma_n^2\}} \text{ et } \mathcal{F}_k = \mathcal{G}_{n-k}.$$

Nous avons, pour $1 \leq k \leq n$,

$$W_k = \sum_{j=1}^{k} \mathbb{E}[U_j^2|\mathcal{F}_{j-1}] \ = \frac{t^2}{\sigma_n^2} \sum_{j=n-k}^{n-1} \mathbb{E}[Y_j^2|\mathcal{G}_{j+1}]\ 1_{\{Z_j<C\sigma_n^2\}} \leq Ct^2 \ ;$$

$$\text{et } \mathbb{E}[U_k|\mathcal{F}_{k-1}] = \frac{t}{\sigma_n}\ 1_{\{Z_{n-k}<C\sigma_n^2\}}\ \mathbb{E}[Y_k|\mathcal{G}_{n-k}] = 0.$$

Les familles $(\mathcal{F}_k)_{0\leq k\leq n}$ et $(U_k)_{1\leq k\leq n}$ vérifient donc les hypothèses du lemme 11.1.3.

En posant

$$\tilde{S_n} := U_1 + \ldots + U_n = Y_0 1_{\{Z_0<C\sigma_n^2\}} + \ldots + Y_{n-1} 1_{\{Z_{n-1}<C\sigma_n^2\}},$$

on en déduit :

$$\begin{aligned} |\mathbb{E}[e^{i\frac{t}{\sigma_n}\ \tilde{S_n}} e^{\frac{1}{2}W_n}] - 1]| &\leq e^{\frac{Ct^2}{2}} \left((\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})\ Ct^2 + (1 + \frac{Ct^2}{8}) \sum_{k=1}^{n} \mathbb{E}[U_k^2 1_{\{|U_k|>\varepsilon\}}] \right) \\ &\leq e^{\frac{Ct^2}{2}} \left((\frac{\varepsilon}{6} + \frac{\varepsilon^2}{8})\ Ct^2 \right. \\ &\qquad \left. + (1 + \frac{Ct^2}{8})\ \frac{t^2}{\sigma_n^2} \sum_{k=0}^{n-1} \mathbb{E}[Y_k^2 1_{\{|tY_k|>\sigma_n\varepsilon\}}] \right). \end{aligned}$$

L'hypothèse *ii)* nous assure alors que la suite $|\mathbb{E}[e^{i\frac{t}{\sigma_n}\ \tilde{S_n}} e^{\frac{1}{2}W_n}] - 1]|$ converge vers zéro.

Les inégalités

$$\sigma_n^{-2} t^2 V_n 1_{\{V_n<C\sigma_n^2\}} \leq W_n \leq \sigma_n^{-2} t^2 V_n$$

et l'hypothèse *i)* montrent que la suite (W_n) converge en probabilité vers t^2.

Comme dans la démonstration du théorème 11.1.4, on montre alors que la suite des fonctions caractéristiques $\mathbb{E}[e^{i\frac{t}{\sigma_n}\tilde{S_n}}]$ converge vers $e^{-\frac{t^2}{2}}$.

Il reste à comparer S_n et $\tilde{S_n}$. Nous avons : $\{\tilde{S_n} \neq S_n\} \subset \{V_n > C\sigma_n^2\}$; d'où le résultat. $\square$

• **TCL pour des sous-suites**

Pour un accroissement de martingale, le TCL est encore vérifié quand on "observe" la suite $(S(n,.))$ à des temps aléatoires de densité > 0. Ce résultat est classique dans le cas indépendant (cf. Renyi [Ren ? ?], p. 144).

Théorème 11.1.8 *Sous les hypothèses du théorème 11.1.4, si $(k_n(x))_{n\geq 1}$ est une suite strictement croissante de fonctions mesurables à valeurs dans $\mathbb{N}$ telle que la suite $(\frac{k_n(x)}{n})$ converge en probabilité vers une constante $a \in [1,\infty[$, on a :*

$$\frac{1}{\sqrt{n}} S(k_n(.),.) \xrightarrow{d} \mathcal{N}(0, \sigma^2/a).$$

Preuve : Dans le cas indépendant, le résultat repose sur l'inégalité de Kolmogorov. On utilise de même ici une inégalité résultant de la propriété de martingale.

Rappelons que, si $(Y_\ell, \mathcal{F}_\ell)_{\ell=1,\dots,n}$ est une sous-martingale sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ (i.e. une suite de v. a. réelles (Y_ℓ) intégrables et une famille croissante de sous-tribus $\mathcal{F}_\ell$, telles que Y_ℓ soit $\mathcal{F}_\ell$-mesurable et $\mathbb{E}(Y_\ell|\mathcal{F}_{\ell-1}) \; = Y_{\ell-1})$, on a, pour tout $\alpha > 0$:

$$\mathbb{P}(\sup_{1\leq \ell\leq n} |Y_\ell| \; = \alpha) \leq \frac{1}{\alpha}\mathbb{E}(|Y_n|).$$

Si (X_j) est une suite d'accroissements stationnaires de martingale de carré intégrable, de variance σ^2, on obtient, en appliquant l'inégalité précédente à $Y_\ell = |\sum_{j=1}^{\ell} X_j|^2$:

$$\mathbb{P}(\sup_{n_0\leq \ell< n_1} |\sum_{j=n_0}^{\ell} X_j| \; = \alpha) \leq \frac{(n_1-n_0)\sigma^2}{\alpha^2}.$$

Soient $\delta > 0, \epsilon > 0$. L'inégalité précédente implique, en prenant $\alpha = \epsilon\sqrt{n}$, $n_0 = [an]$ et $n_1 = [an] + [\epsilon^2\delta n]$,

$$\mathbb{P}(\sup_{0\leq \ell-[an]\leq \epsilon^2\delta n} |\sum_{j=[an]}^{\ell} X_j| \; = \epsilon\sqrt{n}) \leq \delta\sigma^2.$$

Prenons comme espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ l'espace $(E, \mathcal{E}, \mu)$ et pour X_j les variables $f \circ T^j$, où f vérifie les conditions du théorème (1.1). On obtient :

$$\mu(\sup_{0\leq \ell-[an]<\epsilon^2\delta n} |\sum_{j=[an]}^{\ell} f\circ T^j| \; = \epsilon\sqrt{n}) < \delta\sigma^2.$$

On applique cette inégalité à $\sum_{j=[an]}^{k_n(.)} f\circ T^j$, pour $k_n(.) \; = [an]$ et on traite de façon analogue la somme $\sum_{j=[an]}^{k_n(.)} f\circ T^j$, pour le cas $k_n(.) < [an]$. On obtient ainsi :

$$\begin{aligned}
&\mu(|\frac{Sf([an],.)}{\sqrt{an}} - \frac{Sf(k_n(.),.)}{\sqrt{an}}| \; = \epsilon) \\
&\leq \mu(\sup_{|\ell-[an]|<\epsilon^2\delta n} \frac{1}{\sqrt{an}} |\sum_{j=[an]}^{\ell} f\circ T^j| \; = \epsilon) + \mu(|k_n(.) - [an]| \; = \epsilon^2\delta n) \\
&\leq 2\delta\sigma^2/a + \mu(|k_n(.) - [an]| \geq \epsilon^2\delta n), \qquad (11.3)
\end{aligned}$$

ce qui montre la convergence en loi annoncée. □

Notons que la majoration utilisée dans la démonstration est essentiellement celle qui fournit la condition de "tension" permettant classiquement d'obtenir un théorème fonctionnel pour les accroissements de martingales.

11.2 Cas des différences de martingales stationnaires

Soit $(X_n)_{n\geq 0}$ un processus à valeurs réelles, défini sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ *stationnaire* au sens strict (pour tous entiers $p \geq 1, n \geq 0$ la distribution du p-uple $(X_n, \ldots, X_{n+p-1})$ est indépendante de n).

Rappelons (cf. 9.1) qu'un tel processus stationnaire peut toujours être réalisé comme le processus des applications coordonnées sur l'espace produit $\mathbb{R}^{\mathbb{N}}$ muni d'une probabilité $\mathbb{P}$ invariante par le décalage θ opèrant sur l'espace $\mathbb{R}^{\mathbb{N}}$ par décalage des coordonnées.

Il existe en effet une unique mesure (nécessairement invariante par le décalage) sur $\mathbb{R}^{\mathbb{N}}$ telle que, pour tout $n \geq 0$, la distribution du n-uplet des n premières applications coordonnées soit égale à celle du n-uplet $(X_0, \cdots, X_{n-1})$. On peut ensuite étendre la probabilité $\mathbb{P}$ en une probabilité définie sur l'espace $\mathbb{R}^{\mathbb{Z}}$ invariante par le décalage θ opérant sur l'espace des suites bilatères.

Inversement, étant donnés un système dynamique $(E, \mathcal{E}, \mu, \tau)$ inversible et une fonction f dans $L^2(\mu)$, la suite $(f \circ \tau^n)_{n\geq 0}$ forme un processus stationnaire, qui est ergodique si le système $(E, \mathcal{E}, \mu, \tau)$ l'est.

(Voir la section 9.1 du chapitre 9 (Processus stationnaires) dont nous reprenons les notations.)

Le théorème suivant obtenu indépendemment par P. Billingsley et I. A. Ibragimov en 1961 et 1963 est un cas particulier du théorème de B.M. Brown de 1971.

Théorème 11.2.1 *(Ibragimov, Billingsley) Soit $((X_n)_{n\in\mathbb{Z}}, (\mathcal{F}_n)_{n\in\mathbb{Z}})$ une suite d'accroissements de martingale de carré intégrable définie sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, telle que le processus (X_n) soit stationnaire et ergodique avec $\sigma^2 := \mathbb{E}[X_1^2] > 0$. Alors, nous avons $\dfrac{X_1 + \ldots + X_n}{\sigma\sqrt{n}} \overset{d/\mathbb{P}}{\to} \mathcal{N}(0,1)$.*

Preuve : Pour tout entier $n \geq 1$, posons

$$V_n = \sum_{k=1}^{n} \mathbb{E}_\mu[X_k^2 | \mathcal{F}_{k-1}^X].$$

D'une part, nous avons :

$$\frac{1}{n}\, V_n = \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[Y_0^2 | \mathcal{F}_{-1}] \circ \theta^k.$$

Ceci montre, d'après le théorème ergodique, que la suite de v.a. $(\dfrac{1}{n\sigma^2}\, V_n)_{n\geq 1}$ converge dans $L^1(\mu)$, et donc en probabilité, vers 1.

D'autre part, pour tout $\varepsilon > 0$:

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[X_k^2 1_{\{|X_k| > \varepsilon\sigma\sqrt{n}\}}] \;=\; \mathbb{E}[X_1^2 1_{\{|X_1| > \varepsilon\sigma\sqrt{n}\}}].$$

D'où l'on déduit

$$\lim_{n\to+\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[X_k^2 1_{\{|X_k| > \varepsilon\sigma\sqrt{n}\}}] \;=\; 0.$$

Le résultat est alors une conséquence du théorème 11.1.4. □

De même, nous avons, en appliquant le théorème 11.1.7 :

Théorème 11.2.2 *Soit $((X_n)_{n\geq 0}, (\mathcal{G}_n^X)_{n\geq 0})$ une suite d'accroissements de martingale renversée de carré intégrable défini sur l'espace probabilisé $(B, \mathcal{B}, \lambda)$, tel que le processus $(X_n)_{n\geq 0}$ soit stationnaire et ergodique, avec $\sigma^2 := \mathbb{E}[X_1^2] > 0$. Alors, nous avons $\dfrac{X_1 + \ldots + X_n}{\sigma\sqrt{n}} \overset{d/\mathbb{P}}{\to} \mathcal{N}(0,1)$.*

La formulation du théorème 11.1.5 permet de traiter le cas non ergodique.

Théorème 11.2.3 *Soit $(X_n)_{n\in\mathbb{Z}}$ un processus stationnaire de carré intégrable, formant une suite d'accroissements de martingale pour sa filtration naturelle. Alors il vérifie le théorème de la limite centrale généralisé :*

$$\lim_{n\to\infty} \mathbb{P}(\frac{X_0 + X_1 + ... + X_{n-1}}{\sqrt{n}} \leq \alpha) = \int \phi_y(\alpha)\ d\mathbb{P}_U(y),$$

où U est la v.a. $\mathbb{E}[X_0^2|\mathcal{J}]$, espérance de X_0^2 par rapport à la tribu $\mathcal{J}$ des invariants du système associé.

Preuve : En reprenant les notations précédentes, il résulte du théorème ergodique que $\mathbb{P}$-p.s. et dans $L^1(\mathbb{P})$, on a :

$$\lim_n \frac{1}{n} V_n = \lim_n \frac{1}{n} \sum_{k=1}^{n} \mathbb{E}[X_0^2|\mathcal{F}_{-1}] \circ \theta^k = \mathbb{E}[\mathbb{E}[X_0^2|\mathcal{F}_{-1}]|\mathcal{J}] = \mathbb{E}[X_0^2|\mathcal{J}].$$

Pour une fonction borélienne strictement positive et bornée g, soit ${}^g\mathbb{P}$ la probabilité définie par

$${}^g\mathbb{P} = \frac{g(U)}{\mathbb{E}[g(U)]}\mathbb{P}.$$

En appliquant le théorème ergodique aux moyennes (pour les "temps négatifs")

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbb{E}[X_0^2|\mathcal{F}_{-1}] \circ \theta^{-k},$$

on obtient une limite presque-sûrement égale à U. Ceci implique que U, et donc $g(U)$ est $\mathcal{F}_n$-mesurable, pour tout n, ce qui implique :

$${}^g\mathbb{E}[X_k^2|\mathcal{F}_{k-1}] = \mathbb{E}[X_0^2|\mathcal{F}_{-1}] \circ \theta^k,\ {}^g\mathbb{P} - p.s.,$$

et donc le processus croissant sous ${}^g\mathbb{P}$ s'écrit

$$V_n = \sum_{k=1}^{n} {}^g\mathbb{E}[X_k^2|\mathcal{F}_{k-1}] = \sum_{k=1}^{n} \mathbb{E}[X_0^2|\mathcal{F}_{-1}] \circ \theta^k,\ {}^g\mathbb{P} - p.s..$$

Les moyennes convergent ${}^g\mathbb{P}$-p.s. et dans $L^1({}^g\mathbb{P})$ vers U, et on a donc, ${}^g\mathbb{P}$-p.s. et dans $L^1({}^g\mathbb{P})$:

$$\frac{V_n}{\sigma_n^2} \to \frac{U\mathbb{E}[g(U)]}{\mathbb{E}[Ug(U)]}.$$

Notons $\widetilde{U}$ la limite. Il résulte du théorème 11.1.5 que, pour tout $t \in \mathbb{R}$ et tout $C > 0$,

$$\limsup_n |1 - \ {}^g\mathbb{E}[e^{it\frac{S_n}{\sigma_n} + \frac{1}{2}t^2\widetilde{U}} 1_{\{\widetilde{U}<C\}}]| \leq e^{\frac{1}{2}Ct^2} \ {}^g\mathbb{E}[1_{\{\widetilde{U}\geq C\}}],$$

soit

$$\limsup_n |\mathbb{E}[g(U)] - \mathbb{E}[e^{it\frac{S_n}{\sigma_n} + \frac{1}{2}t^2\widetilde{U}} \ g(U) 1_{\{\widetilde{U}<C\}}]| \leq e^{\frac{1}{2}Ct^2}\mathbb{E}[g(U) 1_{\{\widetilde{U}\geq C\}}].$$

En changeant t en $t(\frac{\mathbb{E}[Ug(U)]}{\mathbb{E}[g(U)]})^{\frac{1}{2}}$ et C en $C\frac{\mathbb{E}[Ug(U)]}{\mathbb{E}[g(U)]}$, on obtient, pour tout $t \in \mathbb{R}$ et tout $C > 0$:

$$\limsup_n |\mathbb{E}[g(U)] - \mathbb{E}[e^{it(\frac{\mathbb{E}[Ug(U)]}{\mathbb{E}[g(U)]})^{\frac{1}{2}}\frac{S_n}{\sigma_n} + \frac{1}{2}t^2 U} \ g(U) 1_{\{U<C\}}]| \leq e^{\frac{1}{2}Ct^2}\mathbb{E}[g(U) 1_{U\geq C\}}].$$

En appliquant cette majoration à $g(.) = e^{-\frac{1}{2}t^2 .}$, on obtient :

$$\begin{aligned}
&\limsup_n |\mathbb{E}[e^{-\frac{1}{2}t^2 U}] - \mathbb{E}[e^{it(\frac{\mathbb{E}[Ug(U)]}{\mathbb{E}[g(U)]})^{\frac{1}{2}}\frac{S_n}{\sigma_n}} 1_{\{U<C\}}| \\
&\leq e^{\frac{1}{2}Ct^2}\mathbb{E}[e^{-\frac{1}{2}t^2 U} 1_{U\geq C\}}] \\
&\leq \mathbb{P}(U \geq C).
\end{aligned}$$

Il en résulte que :

$$\lim_n \mathbb{E}[e^{it(\frac{\mathbb{E}[Ug(U)]}{\mathbb{E}[g(U)]})^{\frac{1}{2}}\frac{S_n}{\sigma_n}}] = \mathbb{E}[e^{-\frac{1}{2}t^2 U}],$$

et par suite

$$\lim_n \mathbb{E}[e^{it\frac{S_n}{\sqrt{n}}}] = \mathbb{E}[e^{-\frac{1}{2}t^2 U}].$$

□

11.2.1 Variance, cobords

L'étude de la variance et la réduction d'une fonction présentée maintenant s'applique à un système dynamique $(E, \mathcal{E}, \mu, \tau)$, mais il s'agit essentiellement d'une étude hilbertienne qui peut être appliquée de façon générale à un "processus du second ordre" (cf. chapitre 9.1).

Cette étude est une étape importante dans l'établissement d'un TCL, en particulier pour établir que la loi limite n'est pas dégénérée.

Soit f une fonction à valeurs réelles dans $L^2(\mu)$ que, dans tout ce pararagraphe, nous supposerons d'intégrale nulle, $\mu(f) = 0$. Dans la recherche de conditions assurant que la suite $(S_n f)_{n\geq 1}$ des sommes ergodiques ($S_n f = f + Tf + ... + T^{n-1}f$) vérifie le TCL, un premier pas consiste à étudier la **variance** $\sigma^2(f)$ définie par :

$$\sigma^2(f) = \limsup_n \frac{1}{n}\|f + Tf + ... + T^{n-1}f\|_2^2 \in [0, \infty]. \tag{11.4}$$

Un problème est de montrer que la limite, $\lim_n \dfrac{1}{n}\|S_n f\|_2^2$, d'une part existe et est *finie*, d'autre part est *non nulle*. Ce deuxième point est lié à l'équation de cobord $f = h - Th$.

Notons que, si une fonction f dans $L^2(\mu)$ vérifie les conditions du théorème 11.4.1, les fonctions $T^n f$, $n \in \mathbb{Z}$, sont deux à deux orthogonales. On a donc dans ce cas : $\frac{1}{n}\|S_n f\|^2 = \|f\|^2 = \sigma^2(f)$.

Définitions 11.2.4 Une fonction f est un *cobord* (pour l'action de τ) s'il existe une fonction h mesurable telle que $f = h - Th$. Si la fonction h est dans $L^2(\mu)$, nous dirons que f est un cobord dans $L^2(\mu)$.

L'étude de la variance est une question spectrale, i.e. ne mettant en jeu que l'opérateur T opérant dans l'espace de Hilbert $L^2(\mu)$. La condition $\sigma^2(f) < \infty$ résulte d'une hypothèse de vitesse de **décorrélation** de la suite $(T^n f)$, i.e. de vitesse de convergence vers 0 de la suite $(\langle T^n f, f\rangle)$ (mais non l'inverse), et elle est préservée si l'on ajoute à f un cobord dans L^2.

Rappelons d'abord quelques résultats obtenus en 9.3. Nous avons, d'après la stationnarité :

$$\frac{1}{n}\|S_n f\|^2 = \|f\|^2 + 2\sum_{k=1}^{n}(1-\frac{k}{n})\langle T^k f, f\rangle. \tag{11.5}$$

Il en résulte que, si la série $\sum_{k\in\mathbb{Z}} |\langle T^k f, f\rangle|$ converge, alors la variance $\sigma^2(f)$ existe et est donnée par l'expression

$$\sigma^2(f) = \sum_{k\in\mathbb{Z}}\langle T^k f, f\rangle = \|f\|^2 + 2\sum_{k=1}^{\infty}\langle T^k f, f\rangle. \tag{11.6}$$

Soit ν_f la mesure spectrale du processus stationnaire $(T^n f)_{n\in\mathbb{Z}}$, i.e. la mesure positive sur le cercle (théorème de Bochner 9.2.2) dont les coefficients de Fourier sont donnés par la fonction de corrélation : $\hat{\nu}(n) = \langle T^n f, f\rangle,\ n \in \mathbb{Z}$.

Si $\sum_{k\in\mathbb{Z}} |\langle T^k f, f\rangle| < +\infty$, la mesure spectrale ν_f a une densité ϕ_f, qui est une fonction paire (car f est réelle) et continue, somme de sa série de Fourier :

$$\phi_f(t) = \sum_{k\in\mathbb{Z}}\langle T^k f, f\rangle e^{ikt}. \tag{11.7}$$

On a alors :

$$\sigma^2(f) = \phi_f(0). \tag{11.8}$$

Cette relation est vérifiée si la mesure spectrale associée à f a une densité ϕ_f continue à l'origine.

Rappelons également :

Lemme 11.2.5 *La suite $(\|S_n f\|_2)_{n\in\mathbb{N}}$ est bornée si, et seulement si, f est un cobord dans $L^2(\mu)$.*

Sous la condition de décorrélation $\sum_{n\geq 0} n|\langle T^n f, f\rangle| < +\infty$, la suite $(\|S_n f\|^2 - n\sigma^2(f))_{n\geq 0}$ est bornée. La variance $\sigma^2(f)$ est nulle alors si, et seulement si, f est un cobord dans L^2.

Remarque 11.2.6 L'application $f \to \sigma(f) := \limsup_n \frac{1}{n}\|f + Tf + ... + T^{n-1}f\|_2^2 \in [0,\infty]$ vérifie l'inégalité triangulaire. L'ensemble des fonctions f, telles que $\sigma^2(f) < \infty$ vérifiant le TCL (y compris celles pour lesquelles la limite est dégénérée), forment un sous-ensemble de $L^2(\mu)$, qui est fermé pour la "semi-norme" définie par $\sigma()$ d'après le lemme suivant :

Lemme 11.2.7 *Si $f = f_k + h_k$ avec les conditions $\sigma^2(h_k) \to 0$, $0 < \sigma^2(f) < \infty$ et si, pour tout $k \in \mathbb{N}$,*

$$\frac{1}{\sigma(f_k)\sqrt{n}}\sum_{j=0}^{n-1} T^j f_k \xrightarrow{d} \mathcal{N}(0,1),$$

on a alors

$$\frac{1}{\sigma(f)\sqrt{n}}\sum_{j=0}^{n-1} T^j f \xrightarrow{d} \mathcal{N}(0,1).$$

Preuve : Soit J un intervalle $[a,b]$ de $\mathbb{R}$ et, pour $\delta > 0$, J_δ l'intervalle $[a-\delta, b+\delta]$. Si X et Y sont deux v.a. de carré intégrable, nous avons :

$$\mathbb{P}(X \in J) \leq \mathbb{P}(Y \in J_\delta) + \frac{1}{\delta^2}\|X-Y\|_2^2.$$

Il en résulte, pour tout $k \in \mathbb{N}$,

$$\mathbb{P}(\frac{1}{\sigma(f_k)\sqrt{n}}\sum_0^{n-1} T^j f \in J) \leq \mathbb{P}(\frac{1}{\sigma(f_k)\sqrt{n}}\sum_{j=0}^{n-1} T^j f_k \in J_\delta) + \frac{1}{\delta^2\sigma^2(f_k)}\sigma_n^2(h_k).$$

D'où, pour tout $\delta > 0$,

$$\limsup_n \mathbb{P}(\frac{1}{\sigma(f)\sqrt{n}}\sum_{j=0}^{n-1} T^j f \in J) \leq \mathcal{N}(J_\delta) + \frac{1}{\delta^2\sigma^2(f_k)}\sigma^2(h_k),$$

et donc :

$$\limsup_n \mathbb{P}(\frac{1}{\sigma(f)\sqrt{n}}\sum_{j=0}^{n-1} T^j f \in J) \leq \mathcal{N}(J_\delta).$$

On a de même

$$\liminf_n \mathbb{P}(\frac{1}{\sigma(f)\sqrt{n}}\sum_{j=0}^{n-1} T^j f \in J) \geq \mathcal{N}(J_{-\delta}),$$

et donc, δ étant arbitraire,

$$\lim_n \mathbb{P}(\frac{1}{\sigma(f)\sqrt{n}}\sum_0^{n-1} T^j f \in J) = \mathcal{N}(J).$$

□

11.3 Application aux chaînes de Markov

Considérons un espace probabilisé $(E, \mathcal{E}, \mu)$ et un opérateur de transition P sur cet espace laissant la mesure μ invariante. L'opérateur P définit une contraction de $L^2(\mu)$ que nous noterons encore P. Il lui est associée (cf. 3.4, sur l'espace de trajectoires, la chaîne de Markov canonique

$$(\Omega = E^{\mathbb{N}}, \mathcal{F} = \bigoplus_{\mathbb{N}} \mathcal{E}, (X_n)_{n\in\mathbb{N}}, (\mathbb{P}_x)_{x\in E}).$$

• **Processus de Markov stationnaires**

Soit f une fonction dans $L^2(\mu)$. Pour la loi initiale μ, i.e. en munissant Ω de la probabilité invariante par décalage $\mathbb{P}_\mu$, le processus $(f(X_n))_{n\geq 0}$ est stationnaire. Donnons un premier théorème de la limite centrale pour la suite stationnaire $(f(X_n)_{n\geq 0})$ sous $\mathbb{P}_\mu$, dû à M. I. Gordin et B. A. Lifsic.

Théorème 11.3.1 *Supposons que les fonctions P-harmoniques bornées soient μ-presque partout constantes. Si $f \in L^2(\mu)$ est de la forme $f = g - Pg$ avec $g \in L^2(\mu)$, alors f a une variance $\sigma^2(f) = \int_E [g^2 - (Pg)^2] \, d\mu$ et la suite $(f(X_n)_{n\geq 0})$ vérifie :*

$$\frac{f(X_0) + \ldots + f(X_{n-1})}{\sqrt{n}} \stackrel{d/\mathbb{P}_\mu}{\longrightarrow} \mathcal{N}(0, \sigma(f)^2).$$

Preuve : Nous avons :

$$f(X_0) + \ldots + f(X_{n-1}) = \sum_{k=1}^{n} (g(X_k) - Pg(X_{k-1})) + g(X_0) - Pg(X_n).$$

Le lemme 8.6.5, appliqué à la suite $(\frac{1}{n}(Pg(X_0))^2 \circ \theta^n)$, implique la convergence : $\frac{1}{\sqrt{n}} Pg(X_n) \to 0$, $\mathbb{P}_\mu$-p.s..

D'autre part, la somme du second membre est une martingale stationnaire. D'après l'hypothèse sur les fonctions P-harmoniques, le système dynamique $(\Omega, \mathcal{F}, \mathbb{P}_\mu, \theta)$ est ergodique (cf. le corollaire 3.5.2). On conclut à l'aide du théorème 11.2.1. $\square$

• **Convergence sous $\mathbb{P}_x$**

Le théorème 11.1.4 permet de préciser le résultat du théorème 11.3.1.

Théorème 11.3.2 *Sous les hypothèses du théorème (11.3.1), le TCL est vérifié, sous la probabilité $\mathbb{P}_x$, pour μ-presque tout x.*

Preuve : Il suffit, dans la démonstration précédente, de vérifier les conditions $i)$ et $ii)$ du théorème 11.1.4 pour la martingale $\sum_{k=1}^n (g(X_k) - Pg(X_{k-1}))$ considérée sous la probabilité $\mathbb{P}_x$.

Pour tout $k \geq 1$, posons $Y_k = g(X_k) - Pg(X_{k-1})$. Le processus croissant sous $\mathbb{P}_x$, $V_n = \sum_{k=1}^n \mathbb{E}_x[Y_k^2 | \mathcal{F}_{k-1}]$, $n \geq 1$, associé à la suite d'accroissements de martingale $(Y_n)_{n\geq 1}$ s'écrit :

$$V_n = \sum_{k=1}^{n} (Pg^2 - (Pg)^2)(X_{k-1}),$$

et la variance

$$\mathbb{E}_x[V_n] = \sum_{k=1}^{n} \mathbb{E}_x[Y_k^2] = \sum_{k=1}^{n} P^{k-1} \, (Pg^2 - (Pg)^2)(x).$$

Le théorème ergodique implique que nous avons, $\mathbb{P}_\mu$-p.s. :

$$\lim_n \frac{1}{n} V_n = \mu(Pg^2 - (Pg)^2) = \sigma^2(f),$$
$$\lim_n \mathbb{E}_\mu[|\frac{1}{n} V_n - \sigma^2(f)|] = 0,$$

d'où il résulte, pour μ-presque tout x :

$$\lim_n \frac{1}{n} V_n = \sigma^2(f), \ \mathbb{P}_x - \text{p.s.},$$
$$\lim_n \mathbb{E}_x[\frac{1}{n} V_n] = \sigma^2(f).$$

La condition $i)$ du théorème 11.1.4 est donc vérifiée pour μ-presque tout x.

D'autre part, pour tout nombre réel $C > 0$, nous avons

$$\frac{1}{n}\sum_{k\geq 1}\mathbb{E}_x[Y_k^2 1_{|Y_k|>C}] = \frac{1}{n}\sum_{k=1}^{n} P^{k-1}H_C(x);$$

où

$$H_C(x) = \int (g(u) - Pg(x))^2 1_{\{|g(u)-Pg(x)|>C\}}\ P(x, du).$$

D'après le théorème ergodique, cette moyenne converge vers $\mu(H_C)$ pour μ-presque tout x. On en déduit que, pour tout $\varepsilon > 0$ et tout $C > 0$,

$$\limsup_n \frac{1}{n}\sum_{k\geq 1}\mathbb{E}_x[Y_k^2 1_{|Y_k|>\varepsilon\sigma_n}] \leq \mu(H_C),\ \text{pour } \mu\text{-presque tout } x.$$

Le théorème de convergence dominée assure que $\lim_{C\to+\infty}\mu(H_C) = 0$; d'où la condition de Lindeberg *ii)* du théorème (11.1.4), pour μ-presque tout x.

□

• Cas de l'équicontinuité

Dans la suite de cette section, nous considérons un espace métrique compact E muni de sa tribu borélienne. Notons $\mathcal{C}(E)$ l'espaced es fonctions continues sur E. Soit P un opérateur de transition sur E tel que $Pf \in \mathcal{C}(E)$, pour $f \in \mathcal{C}(E)$. Sous des conditions sur f et sur P, nous allons obtenir des convergences au sens de la probabilité $\mathbb{P}_x$, pour tout x : d'abord dans le cas où il y a unicité de la mesure invariante, puis sous une hypothèse plus générale d'"équicontinuité en moyenne".

Rappelons que P vérifie la propriété d'équicontinuité en moyenne si, *pour toute fonction continue f, la famille $(\frac{1}{n}\sum_{k=1}^{n} P^k f)_{n\geq 1}$ est équicontinue.* Nous utilisons les notations et les résultats du chapitre 6 (notamment de 6.3.1). D'après la proposition 6.3.2, nous avons :

Théorème 11.3.3 *Supposons que P possède une unique mesure invariante μ. Si $f \in \mathcal{C}(E)$ est de la forme $f = g - Pg$ avec $g \in \mathcal{C}(E)$, alors f a une variance $\sigma^2(f) = \int_E [g^2 - (Pg)^2]\ d\mu$. La suite $(f(X_n))_{n\geq 0}$ vérifie le TCL pour toutes les probabilités $\mathbb{P}_x, x \in E$.*

Preuve : La fonction f étant bornée, il suffit d'étudier la martingale $(\sum_{k=1}^{n}(g(X_k) - Pg(X_{k-1})))_{n\geq 1}$. Appliquons le théorème 11.1.4.

La condition *ii)* du théorème 11.1.4 est automatiquement vérifiée, car g est bornée.

Pour la condition *i)*, on reprend la démonstration du théorème précédent en utilisant le théorème 6.3.1, qui permet d'avoir des convergences sous $\mathbb{P}_x$, *pour tout $x \in E$.* □

Théorème 11.3.4 *Supposons que P vérifie la propriété d'équicontinuité en moyenne. Si $f \in \mathcal{C}(E)$ est de la forme $f = g - Pg$ avec $g \in \mathcal{C}(E)$. Alors, pour tout $x \in E$, nous avons :*

$$\lim_{n\to\infty} \mathbb{P}_x(\frac{f(X_0) + \ldots + f(X_n)}{\sqrt{n}} \leq \alpha) = \int_{\mathcal{B}} \phi_{\sigma_\gamma(f)}(\alpha)\ \rho(x, d\gamma),\ \forall \alpha \in \mathbb{R},$$

où $\sigma_\gamma^2(f) = \int_E [g^2 - (Pg)^2]\ d\nu_\gamma$.

Preuve : Soient $\widehat{\xi}$ une fonction strictement positive dans $\mathcal{C}(\mathcal{B})$ et ξ la fonction dans $\mathcal{H}_c$ associée, définie par :

$$\xi(u) = \mathbb{E}_u[\widehat{\xi}\circ\eta] = \int \widehat{\xi}(\gamma)\rho(u, d\gamma),\ \forall u \in E.$$

Considérons la chaîne relativisée définie par l'opérateur de transition

$$^{\xi}Pf(u) = \frac{1}{\xi(u)} P(\xi f)(u), \ u \in E,$$

et la probabilité associée sur $E^{\mathbb{N}}$

$$^{\xi}\mathbb{P}_x = \frac{\widehat{\xi}(\eta)}{\xi(x)} \mathbb{P}_x.$$

Notons $^{\xi}\mathbb{E}_x$ l'espérance correspondante. Les applications coordonnées $(X_n)_{n\geq 0}$ sur l'espace produit $(E^{\mathbb{N}}, \mathcal{F})$ définissent sous $^{\xi}\mathbb{P}_x$ une chaîne de Markov de probabilité de transition $^{\xi}P$, partant de x.

Posons

$$\begin{aligned} S_n &= \sum_{k=1}^{n} (g(X_k) - \ Pg(X_{k-1})), \\ V_n &= \sum_{k=1}^{n} (Pg^2 - (Pg)^2)(X_{k-1}), \end{aligned}$$

et soit $\sigma_n^2 = \mathbb{E}[V_n]$ la variance associée.

Notons également, pour la chaîne relativisée :

$$\begin{aligned} \widetilde{S}_n &= \sum_{k=1}^{n} (g(X_k) - {}^{\xi}Pg(X_{k-1})), \\ \widetilde{V}_n &= \sum_{k=1}^{n} ({}^{\xi}Pg^2 - ({}^{\xi}Pg)^2)(X_{k-1}). \end{aligned}$$

Sous $^{\xi}\mathbb{P}_x$, la suite $(\widetilde{S}_n)$ est une suite d'accroissements de martingale et $(\widetilde{V}_n)$ est le processus croissant associé. Soit $\widetilde{\sigma}_n^2 = {}^{\xi}\mathbb{E}[\widetilde{V}_n]$ la variance correspondante. Vérifions les conditions du théorème 11.1.4.

On a :

$$\frac{\widetilde{V}_n}{n} \to \nu_\eta({}^{\xi}Pg^2 - ({}^{\xi}Pg)^2),$$

la convergence ayant lieu $\mathbb{P}_x$-p.s. et donc aussi $^{\xi}\mathbb{P}_x$-p.s.. Pour tout u appartenant à une classe ergodique $\gamma \in \mathcal{B}$, on a $^{\xi}Pf(u) = Pf(u)$, tout élément de $\mathcal{H}_c$ étant constant sur γ. La limite ci-dessus est donc aussi égale à $\nu_\eta(Pg^2 - (Pg)^2)$.

En appliquant le théorème de convergence dominée, on obtient :

$$\begin{aligned} n^{-1}\sigma_n^2 \quad &\to \ \mathbb{E}_x[\nu_{\eta()}(Pg^2 - (Pg)^2)] \\ &= \int \nu_\gamma(Pg^2 - (Pg)^2) \ \rho(x, d\gamma), \mathbb{P}_x - \text{p.s.} \\ n^{-1}\widetilde{\sigma}_n^2 \quad &\to {}^{\xi}\mathbb{E}_x[\nu_{\eta()}(Pg^2 - (Pg)^2)] \\ &= \frac{\int \widehat{\xi}(\gamma)\nu_\gamma(Pg^2 - (Pg)^2) \ \rho(x, d\gamma)}{\xi(x)}, \ {}^{\xi}\mathbb{P}_x - \text{p.s..} \end{aligned}$$

Posons $\sigma_x^2 = \int \nu_\gamma(Pg^2 - (Pg)^2) \ \rho(x, d\gamma)$.

Dans le cas où $\sigma_x^2 = 0$, on a $\frac{S_n}{\sqrt{n}} \to 0$, dans $L^2(\mathbb{P}_x)$, et le résultat est vérifié.

Supposons $\sigma_x > 0$. On a alors :

$$\widetilde{\sigma}_n^{-2}\widetilde{V}_n \to \xi(x)(\int \widehat{\xi}(\gamma)\nu_\gamma(Pg^2 - (Pg)^2) \ \rho(x, d\gamma))^{-1}\nu_\eta(Pg^2 - (Pg)^2), \ {}^{\xi}\mathbb{P}_x - \text{p.s..}$$

Par le théorème 11.1.5, il en résulte que :

$$\lim_n {}^\xi\mathbb{E}_x[e^{it\frac{\widetilde{S}_n}{\sqrt{n}}+\frac{t^2}{2}(\nu_\eta(Pg^2-(Pg)^2)}] = 1, \ \forall t \in \mathbb{R}. \tag{11.9}$$

Montrons que l'on peut remplacer $\widetilde{S}_n$ par S_n dans la relation (11.9). D'après la majoration

$$\frac{|(S_n - \widetilde{S}_n)|}{\sqrt{n}} \leq \frac{1}{\sqrt{n}} \sum_1^n |({}^\xi Pg - Pg)(X_{k-1})|,$$

il suffit de montrer que

$$\frac{1}{\sqrt{n}} \sum_0^n {}^\xi\mathbb{E}_x[|{}^\xi Pg - Pg|(X_{k-1})] \to 0. \tag{11.10}$$

On peut écrire

$$\begin{aligned} ({}^\xi Pg - Pg)(X_{k-1}) &= \frac{1}{\xi(X_{k-1})}(P(\xi g)(X_{k-1}) - \xi(X_{k-1})Pg(X_{k-1})) \\ &= \frac{1}{\xi(X_{k-1})}\mathbb{E}_x[(\xi(X_k) - \xi(X_{k-1}))g(X_k)|\mathcal{F}_{k-1}]. \end{aligned}$$

La convergence (11.10) s'obtient en utilisant la majoration

$$\begin{aligned} &\frac{1}{\sqrt{n}} \sum_{k=1}^n {}^\xi\mathbb{E}_x[|{}^\xi Pg - Pg|(X_{k-1})] \\ \leq\ &\frac{1}{\sqrt{n}} \sum_{k=1}^n {}^\xi\mathbb{E}_x[\frac{1}{\xi(X_{k-1})}\mathbb{E}_x[|\xi(X_k) - \xi(X_{k-1})||g(X_k)||\mathcal{F}_{k-1}]] \\ \leq\ &\|g\|_\infty \frac{1}{\xi(x)} \frac{1}{\sqrt{n}} \sum_{k=1}^n \mathbb{E}_x[|\xi(X_k) - \xi(X_{k-1})|], \\ \leq\ &\|g\|_\infty \frac{1}{\xi(x)}[\frac{1}{\sqrt{n}} \sum_{k=1}^L \mathbb{E}_x[|\xi(X_k) - \xi(X_{k-1})| + (\sum_{k=L}^\infty \mathbb{E}_x[|\xi(X_k) - \xi(X_{k-1})|^2)^{\frac{1}{2}}], \end{aligned}$$

pour tout entier L, $1 \leq L \leq n$, et le fait que

$$\sum_{k \geq 1} \mathbb{E}_x[(\xi(X_k) - \xi(X_{k-1}))^2] \leq \|\widehat{\xi}\|_\infty^2 < +\infty.$$

De la convergence dans (11.9), il résulte alors que, $\forall t \in \mathbb{R}$ et $\forall \widehat{\xi} \in \mathcal{C}_+(\mathcal{B})$ strictement positive,

$$\mathbb{E}_x[e^{it\frac{S_n}{\sqrt{n}}+\frac{t^2}{2}\nu_\eta(Pg^2-(Pg)^2)}\widehat{\xi} \circ \eta] \to \mathbb{E}_x[\widehat{\xi} \circ \eta] = \xi(x), \tag{11.11}$$

Cette relation, qui est vraie aussi pour $\widehat{\xi}$ constant, est donc vérifiée par toute fonction continue $\widehat{\xi}$ (remplacer $\widehat{\xi}$ par $\widehat{\xi} - a$, où $a < \inf_{\gamma \in \mathcal{B}} \widehat{\xi}(\gamma)$).

Le premier facteur étant uniformément borné pour t fixé, la convergence (11.11) s'étend à toutes les fonctions ξ dans $L^1(\rho(x, d\gamma))$.

Prenons la fonction sur $\mathcal{B}$ définie par $\widehat{\xi}(\gamma) = e^{-\frac{t^2}{2\sigma_x^2}\nu_\gamma(Pg^2-(Pg)^2)}$, qui est une fonction borélienne bornée. On obtient, en appliquant (11.11),

$$\mathbb{E}_x[e^{it\frac{S_n}{\sqrt{n}}}] \to \int e^{-\frac{t^2}{2}\nu_\gamma(Pg^2-(Pg)^2)} \ \rho(x, d\gamma), \forall t \in \mathbb{R},$$

ce qui prouve le résultat. □

En particulier, dans le cas d'un nombre fini de mesures invariantes extrémales (autrement dit de classes ergodiques), nous avons obtenu :

Corollaire 11.3.5 *Supposons que P vérifie la propriété d'équicontinuité en moyenne et possède un nombre fini de mesures invariantes extrémales $\nu_1, \dots, \nu_r$. Si $f \in \mathcal{C}(E)$ est de la forme $f = g - Pg$ avec $g \in \mathcal{C}(E)$, alors :*

$$\lim_{n\to\infty} \mathbb{P}(\frac{f(X_0) + \ldots + f(X_n)}{\sqrt{n}} \leq \alpha) = \sum_{i=1}^{r} h_i(x)\phi_{\sigma_i}(\alpha),$$

où $\sigma_i^2(f) = \int_E [g^2 - (Pg)^2]\, d\nu_i$; les $h_i(x)$ sont les fonctions P-harmoniques continues à valeurs ≥ 0, avec $h_i(x) = 1$ sur le support de ν_i, $h_i(x) = 0$, sur le support de ν_j, $j \neq i$, et $\sum_{i=1}^r h_i(x) = 1, \forall x \in E$.

Théorème 11.3.6 *Soit $f \in L^2(\mu)$, de la forme $f = g_n - Pg_n + h_n$, avec $g_n, h_n \in L^2(\mu)$ et $\sigma(h_n) \to 0$. Alors f vérifie le TCL.*

On notera que la condition du théorème peut être réalisée, même si $\lim_n \|g_n\| = \infty$. En particulier, on peut avoir $\sigma^2(g_n - Pg_n) \to \sigma^2(f) < \infty$, sans que la suite des normes $(\|g_n\|_{n\geq 0})$ soit bornée.

Extension dans le cas symétrique

Dans le cas des noyaux symétriques, Kipnis et Varadhan [] ont montré que la condition plus faible $f \in Im(I - P)^{\frac{1}{2}}$ permet d'établir le TCL. On est en effet, dans ce cas, dans les conditions du théorème (4.1).

(6.1) Théorème : (Kipnis, Varadhan) *Si l'opérateur de transition P est auto-adjoint, et si f non nulle est de la forme*

$$f = (Id - P)^{1/2} g, (4.3.1)$$

avec $f, g \in L^2(\mu)$, alors f a une variance $\sigma^2(f) = \|g\|_2^2 + \|P^{\frac{1}{2}} g\|_2^2$ et la suite $f(X_n)$ vérifie le TCL. (loi de Gauss centrée de variance

$$\sigma^2 = \lim_{n\to+\infty} \frac{1}{n} E(\sum_{k=0}^{n-1} f \circ X_k)^2,$$

si ce nombre est strictement positif, sinon la loi est la mesure de Dirac en 0.

Preuve : Supposons qu'on ait $f = (Id - P)^{1/2} g$. De l'expression

$$\frac{1}{n}\|S_n f\|_2^2 = -\|f\|_2^2 + 2\sum_{k=0}^{n-1}(1 - \frac{k}{n})\langle P^k f, f\rangle,$$

on déduit :

$$\frac{1}{n}\|S_n f\|_2^2 = -\|f\|_2^2 + 2\ [\|g\|_2^2 - \frac{1}{n}\langle Pg + ... + P^n g, g\rangle].$$

On a donc :

$$\sigma^2(f) = \lim_n 2\|g\|_2^2 - \|f\|_2^2 = 2\ \|g\|_2^2 - \|f\|_2^2 = \|g\|_2^2 + \|P^{\frac{1}{2}} g\|_2^2.$$

Cette quantité n'est nulle que si f est nulle presque partout. Posons $g_n = f + ... + P^{n-1} f$, on a

$$f = g_n - Pg_n + P^n f,$$

et $\sigma^2(P^n f) = 2\|P^n g\|_2^2 - \|P^n f\|_2^2 \to 0$, quand n tend vers l'infini (en effet $(P^n g)$ converge en norme vers une fonction invariante que l'on peut supposer nulle).

On peut donc appliquer le théorème (4.1) ? ?. $\square$

11.4 Application aux systèmes dynamiques

Avec les notations des systèmes dynamiques, on obtient un énoncé équivalent au théorème de Billingsley-Ibragimov, dans lequel on exprime que les sommes ergodiques forment une martingale relativement à une suite croissante de tribus $(\tau^{-n}\mathcal{F})_{n\geq 0}$, où $\mathcal{F}$ est une tribu telle que $\mathcal{F} \subset \tau^{-1}\mathcal{F}$.

Théorème 11.4.1 *Soient $(E, \mathcal{E}, \mu, \tau)$ un système dynamique ergodique inversible, f une fonction dans $L^2(\mu)$ (avec $\sigma^2 = \int f^2 \, d\mu > 0$), $\mathcal{F}$ une sous-tribu de $\mathcal{E}$ telles que*
i) $\mathcal{F} \subset \tau^{-1}\mathcal{F}$ ($\mathcal{F}$ est croissante pour τ^{-1}),
ii) f est $\mathcal{F}$-mesurable,
iii) $\mathbb{E}[f|\tau\mathcal{F}] = 0$.

Alors la suite $(T^n f)_{n\geq 0}$ vérifie le théorème de la limite centrale : $\frac{1}{\sqrt{n}} S_n f \overset{d/\mu}{\to} \mathcal{N}(0, \sigma^2)$.

Remarquons que si les conditions du théorème sont vérifiées par rapport à une filtration $(\mathcal{F}_n)$, elles le sont par rapport à la tribu $\sigma(f \circ \tau^k, k \geq 0)$ engendrée par f et ses images par les itérées de τ. Cependant, dans les applications, la tribu $\mathcal{F}$ (et donc la filtration $(\mathcal{F}_n)$), est généralement construite indépendamment de f. Par exemple, pour un système symbolique sur lequel la transformation qui opère est le décalage τ, il est naturel de prendre pour filtration $(\tau^{-n}\mathcal{F})_{n\geq 0}$ la filtration engendrée par les applications coordonnées, alors que f peut être une fonction dépendant de toutes les coordonnées. Il faut alors, comme nous le verrons, modifier f par "homologie".

Système dynamique non inversible et TCL

Dans le théorème 11.4.1, nous avons supposé le système inversible. La notion de martingale est relative à une filtration croissante correspondant au "passé" de la transformation. Dans le cas où τ est non inversible, par exemple pour le décalage unilatère, on est amené à raisonner par rapport à une filtration décroissante, correspondant à la tribu du "futur". On est alors dans la situation du théorème 11.1.7 (martingales renversées) dont l'énoncé prend la forme suivante :

Théorème 11.4.2 *Soient $(E, \mathcal{E}, \mu, \tau)$ un système dynamique ergodique, f une fonction dans $L^2(\mu)$ (avec $\sigma^2 = \int f^2 \, d\mu > 0$), $\mathcal{F}$ une sous-tribu de $\mathcal{E}$ telles que*
i) $\tau^{-1}\mathcal{F} \subset \mathcal{F}$,
ii) f est $\mathcal{F}$-mesurable,
iii) $\mathbb{E}[f|\tau^{-1}\mathcal{F}] = 0$.

Alors la suite $(T^n f)_{n\geq 0}$ vérifie le théorème de la limite centrale : $\frac{1}{\sqrt{n}} S_n f \overset{d/\mu}{\to} \mathcal{N}(0, \sigma^2)$.

Notons également une version "multidimensionnelle" qui se déduit facilement du cas scalaire.

Théorème 11.4.3 *Soient $(E, \mathcal{E}, \mu, \tau)$ un système dynamique ergodique, $F = (f_1, ..., f_d)$ une fonction à valeurs dans $\mathbb{R}^d$ de carré intégrable, $\mathcal{F}$ une sous-tribu de $\mathcal{E}$ telles que*
i) $\tau^{-1}\mathcal{F} \subset \mathcal{F}$,
ii) F est $\mathcal{F}$-mesurable,
iii) $\mathbb{E}[F|\tau^{-1}\mathcal{F}] = 0$.

Si F est non dégénérée, dans le sens qu'il n'existe pas de vecteur $(a_1, \cdots, a_d)$ de $\mathbb{R}^d$ non nul tel que $\sum_i a_i f_i = 0$, p.p., alors le vecteur $\frac{1}{\sqrt{n}}(F + TF + ... + T^{n-1}F)$ est asymptotiquement de loi gaussienne, de matrice de covariance Γ non dégénérée.

11.4.1 Réduction par homologie à un accroissement de martingale

Le champ d'application de la méthode des martingales, dans l'étude du comportement en loi des sommes ergodiques, peut être étendu grâce au procédé introduit par I. M. Gordin à une classe plus large de fonctions. Ce procédé consiste à exprimer, sous certaines conditions, une fonction comme la somme d'une fonction engendrant un accroissement de martingale et d'un "cobord".

L'addition d'un cobord intégrable ne modifie pas le comportement en loi des sommes ergodiques normalisées, si la normalisation tend vers l'infini. En effet, soient J un intervalle $[a,b]$ de $\mathbb{R}$ et J_δ l'intervalle $[a-\delta, b+\delta]$. En effet, nous avons pour $\delta > 0$:

$$\mathbb{P}(a_n^{-1} S_n f \in J) \leq \mathbb{P}(a_n^{-1} S_n f + T^n h - h \in J_\delta) + 2a_n^{-1}\delta^{-1}\|h\|_1,$$
$$\mathbb{P}(a_n^{-1} S_n f + T^n h - h \in J) \leq \mathbb{P}(a_n^{-1} S_n f \in J_\delta) + 2a_n^{-1}\delta^{-1}\|h\|_1.$$

Définitions 11.4.4 Nous dirons qu'une fonction f est *homologue* dans $L^2(m)$ à une fonction g engendrant une suite d'*accroissements de martingale* (renversée), s'il existe une sous-tribu $\mathcal{A}_0$, vérifiant $\tau^{-1}\mathcal{A}_0 \subset \mathcal{A}_0$, et des fonctions h et g dans $L^2(m)$ avec g $\mathcal{A}_0$-mesurable, vérifiant $\mathbb{E}(g|\tau^{-1}\mathcal{A}_0) = 0$ et

$$f = g + Th - h. \tag{11.12}$$

La suite des sommes partielles $(\sum_0^{n-1} T^k g)_{n\geq 1}$ forme alors, pour la filtration $(\tau^{-n}\mathcal{A}_0)_{n\in\mathbb{Z}}$, une martingale renversée et la variance $\sigma^2(f)$ associée au processus $(T^n f)$ est égale à $\|g\|_2^2$.

On définit de façon analogue la propriété d'être "homologue" à une suite d'*accroissements de martingale directe.*

Dans ce cas, le processus $(f \circ \tau^n)$ engendré par f vérifie encore certains résultats obtenus pour les différences de martingales, en particulier le TCL. Ces propriétés restent stables par "homologie".

Le problème est donc maintenant de donner des conditions qui assurent la validité de la représentation (11.12). Nous allons donner, dans un cadre hilbertien, une méthode de réduction par homologie, que nous appliquerons ensuite pour caractériser, dans un système dynamique, des classes de fonctions homologues à une fonction engendrant une différence de martingale.

Considérons une isométrie T dans un espace de Hilbert $\mathcal{H}$. Comme pour les fonctions sous l'action d'une transformation, nous dirons que deux vecteurs f, g sont **homologues** s'il existe $h \in \mathcal{H}$ tel que $f = g + Th - h$.

Cas unilatère

Nous appellerons **filtration vectorielle unilatère** pour T dans l'espace de Hilbert $\mathcal{H}$ une famille décroissante de sous-espaces fermés

$$\{0\} \subset ... \subset \mathcal{H}_n \subset ... \subset \mathcal{H}_1 \subset \mathcal{H}_0 = \mathcal{H}, \tag{11.13}$$

telle que $T^n\mathcal{H} = \mathcal{H}_n$, $\forall n \geq 0$.

En appliquant à l'opérateur adjoint P le lemme (9.3.1), on obtient :

Théorème 11.4.5 *Etant donnée une filtration unilatère (11.13), pour tout $f \in \mathcal{H}$ les conditions suivantes sont équivalentes :*
i)

$$\sup_n \|\sum_{k=0}^{n} P^k f\| < \infty, \tag{11.14}$$

ii) il existe $g \in L^2(\mu)$ telle que $f = (I - P)g$,
iii) il existe h et $g_1 \in L^2(\mu)$ telles que $f = g_1 + Th - h$ et $Pg_1 = 0$.

Notons que la condition (11.14) résulte de la condition :

$$\sum_{n\geq 0} \|P^n f\| = \sum_{n\geq 0} \|\Pi_{T^n\mathcal{H}} f\| < \infty. \tag{11.15}$$

Cette condition est plus directement accessible dans les exemples. Dans ce cas la fonction g est égale à $\sum_{k=0}^{\infty} P^k f$.

Cas bilatère

Traitons maintenant le cas "bilatère". Considèrons un opérateur T isométrique **surjectif** sur un espace de Hilbert $\mathcal{H}$ (T est donc un opérateur unitaire). Nous appelerons **filtration vectorielle bilatère** pour T une suite croissante de sous-espaces :

$$\{0\} \subset ... \subset \mathcal{H}_n \subset ... \subset \mathcal{H}_1 \subset \mathcal{H}_0 \subset \mathcal{H}_{-1} \subset ... \subset \mathcal{H}_{-n} \subset ... \subset \mathcal{H}, \tag{11.16}$$

telle que $T^n\mathcal{H}_0 = \mathcal{H}_n$, $\forall n \in \mathbb{Z}$.

Par exemple, on obtient une filtration vectorielle, dans le cadre vectoriel des processus "du second ordre", en formant, pour un vecteur f dans $\mathcal{H}$, la suite des sous-espaces $\mathcal{H}_n$ engendrés (linéairement) par les vecteurs $T^j f, j \leq n$.

Notons $\mathcal{W}_0 = \mathcal{H}_0^\perp$, $\mathcal{W}_n = T^{-n}\mathcal{W}_0$. La condition $T\mathcal{H}_0 \subset \mathcal{H}_0$ équivaut à $T^{-1}\mathcal{W}_0 \subset \mathcal{W}_0$. On a les relations $T^{-n}\mathcal{W}_0 = \mathcal{H}_{-n}^\perp$ et donc les inclusions :

$$\{0\} \subset ... \subset \mathcal{W}_n \subset ... \subset \mathcal{W}_1 \subset \mathcal{W}_0.$$

Remarquons l'égalité $\mathcal{W}_0 \ominus T^{-1}\mathcal{W}_0 = T^{-1}\mathcal{H}_0 \ominus \mathcal{H}_0$.

Rappelons maintenant (cf. 9.4.5) une condition suffisante pour qu'un vecteur de $\mathcal{H}$ soit égal, à un cobord près, à un vecteur de $\mathcal{H}_0$ orthogonal à $T\mathcal{H}_0$.

Théorème 11.4.6 *Etant donnée une filtration bilatère (11.16), si $f \in \mathcal{H}$ vérifie les conditions*

$$\sum_{n\geq 0} \|\Pi_{T^n\mathcal{H}_0} f\| < \infty \ \text{et} \ \sum_{n\geq 0} \|f - \Pi_{T^{-n}\mathcal{H}_0} f\| < \infty, \tag{11.17}$$

alors il existe g et h dans $\mathcal{H}$, avec $g \in \mathcal{H}_0 \ominus T\mathcal{H}_0$, tels que

$$f = g + Th - h.$$

Sous ces conditions le vecteur f est égal, à un cobord $Th - h$ près, à un vecteur g orthogonal à ses images par T^n. En particulier, il est de variance finie et vérifie :

$$| \| \sum_0^{n-1} T^k f\| - \sqrt{n}\|g\| \, | \leq 2\|h\|. \tag{11.18}$$

Nous allons maintenant montrer comment obtenir une réduction à un accroissement de martingale en utilisant des filtrations.

11.4.2 Application aux systèmes dynamiques

Cas non inversible

Considérons d'abord un système dynamique $(E, \mathcal{E}, \mu, \tau)$ non inversible. La suite des tribus $(\tau^{-n}\mathcal{E})$ forme une filtration décroissante. Pour toute fonction f de $L^2(\mu)$ telle que $\mathbb{E}[f|\tau^{-1}\mathcal{E}] = 0$, $((f \circ \tau^n)_{n\geq 0}, (\tau^{-n}\mathcal{E})_{n\geq 0})$ est une suite d'accroissements de martingale renversée de carré intégrable. D'après le théorème 11.1.7, nous avons donc :

Théorème 11.4.7 *Si le système dynamique $(E, \mathcal{E}, \mu, \tau)$ est ergodique, pour toute fonction f de $L^2(\mu)$ non nulle presque partout telle que $\mathbb{E}[f|\tau^{-1}\mathcal{E}] = 0$, nous avons (avec $\sigma = \|f\|_2$) :*

$$\frac{1}{\sqrt{n}} S_n f(.) \stackrel{d/\mu}{\to} \mathcal{N}(0, \sigma^2).$$

La condition $\mathbb{E}[f|\tau^{-1}\mathcal{E}] = 0$ s'écrit $Pf = 0$, P étant l'opérateur adjoint de l'opérateur T de composition par τ. Nous allons affaiblir l'hypothèse $Pf = 0$ en appliquant le théorème 11.4.5 à la filtration vectorielle $(L^2(E, \tau^{-n}\mathcal{E}, \mu))$.

Théorème 11.4.8 *Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique ergodique (non inversible), soit P l'opérateur μ-adjoint de l'opérateur T de composition par τ. Soit f une fonction de $L^2(\mu)$ telle que $\int_E f \, d\mu = 0$ et*

$$\sup_{n\geq 0} \|\sum_{k=0}^{n} P^k f\|_2 < +\infty.$$

Alors la suite $(n^{-\frac{1}{2}}\|S_n f\|_2)_{n\geq 1}$ converge vers un réel σ, qui est strictement positif si, et seulement si, f n'est pas un cobord dans $L^2(\mu)$ et f vérifie le TCL :

$$\frac{1}{\sqrt{n}} S_n f(.) \stackrel{d/\mu}{\to} \mathcal{N}(0, \sigma^2).$$

Preuve : D'après le théorème 11.4.5, il existe $g, h \in L^2$ telle que $f = g + h - Th$, et $Pg = 0$. On applique alors le théorème 11.4.7.

□

Remarquons à nouveau que la condition $\sup_{n\geq 0} \|\sum_{k=0}^{n} P^k f\|_2 < +\infty$ est satisfaite lorsque la série $\sum_{n\geq 0} \|P^n f\|_2 = \sum_{n\geq 0} \|\mathbb{E}[f|\tau^{-n}\mathcal{E}]\|_2$ est convergente.

Exemple :

Pour la transformation τ définie par $\tau x \to 2x \bmod 1$ sur $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$ muni de la mesure de Lebesgue, l'opérateur adjoint P est donné par

$$P : g(x) \to Pg(x) = \frac{1}{2}[g(\frac{x}{2}) + g(\frac{x}{2} + \frac{1}{2})].$$

Si $f \in L^2(\mathbb{T}^1)$ est une fonction "impaire" $(f(x + 1/2) = -f(x))$, elle vérifie $Pf = 0$ et donc la conclusion du théorème 11.4.7.

Soit f une fonction d'intégrale nulle, höldérienne d'ordre $0 < \alpha \leq 1$ (c'est-à-dire telle que, pour une constante $C > 0$, on ait : $\forall x, y \in [0, 1[,\ |f(x) - f(y)| \leq C|x - y|^\alpha$).

Pour tout $x \in [0,1[$, nous avons :

$$\begin{aligned} P^n f(x) &= \frac{1}{2^n} \sum_{k=0}^{2^n-1} f(\frac{x+k}{2^n}) = \sum_{k=0}^{2^n-1} \int_{\frac{k}{2^n}}^{\frac{k+1}{2^n}} (f(\frac{x+k}{2^n}) - f(u))\, du \\ &= \frac{1}{2^n} \sum_{k=0}^{2^n-1} \int_0^1 (f(\frac{x+k}{2^n}) - f(\frac{u+k}{2^n}))\, du \leq C\, \frac{1}{2^{n\alpha}}. \end{aligned}$$

Il s'ensuit que la série $\sum \|P^n f\|_\infty$, et a fortiori la série $\sum \|P^n f\|_2$, convergent. D'après 11.4.8 le théorème de la limite centrale est donc vérifié par f.

• Systèmes exacts

Considérons comme en 10.3 un système $(E, \mathcal{E}, \mu, \tau)$ *exact*, i.e. tel que l'intersection $\bigcap_n \tau^{-n}\mathcal{E}$ est triviale (aux ensembles négligeables près). Cette propriété est équivalente à

$$\lim_n \mathbb{E}(f|\tau^{-n}\mathcal{E}) = \mu(f), \forall f \in L^2(\mu).$$

L'opérateur P, adjoint de l'opérateur T de composition par τ dans $L^2(E, \mathcal{E}, \mu)$, est un inverse à gauche de T : $PT = Id$ et, plus généralement, nous avons, comme au paragraphe 4, les relations :

$$P^n T^n f = f,\ \forall f \in L^2(\mu). \tag{11.19}$$

Notons qu'ici nous avons une propriété plus forte :

$$P^n(uT^n f) = P^n(u) f,\ \forall n \in \mathbb{N},\ \forall f \in L^2(\mu), \forall u \in L^\infty(\mu). \tag{11.20}$$

Les fonctions de la forme $T^n g$, $g \in L^2$, appartiennent à $L^2(E, \tau^{-n}\mathcal{E}, \mu)$. Pour tout $n \geq 1$, l'opérateur $T^n P^n$, projection orthogonale sur le sous-espace $\mathcal{H}_n = L^2(E, \tau^{-n}\mathcal{E}, \mu)$ de $L^2(E, \mathcal{E}, \mu)$, est l'espérance conditionnelle par rapport à cette sous-tribu :

$$T^n P^n f = \mathbb{E}(f|\tau^{-n}\mathcal{E}). \tag{11.21}$$

La propriété de système exact peut être exprimée à l'aide de l'opérateur adjoint P. Elle est équivalente à la convergence en norme vers 0, des itérées $P^n f$, pour chaque fonction f dans $L^2(\mu)$ d'intégrale nulle (convergence au sens fort) :

$$\lim_n \|T^n P^n f\|_2 = \lim_n \|P^n f\|_2 = 0, \forall f \in L^2(\mu),\ \text{d'intégrale nulle.}$$

C'est une propriété plus forte que celle de mélange dit "fort" ($\lim_n \langle T^n f, g\rangle = 0$), qui n'assure que la convergence au sens faible vers 0 de la suite $(T^n f)$.

Pour obtenir un TCL, on est conduit à considérer des classes de fonctions pour lesquelles une condition de vitesse permet d'assurer la convergence du "potentiel" $\sum P^n f$.

Exemples :

1) Reprenons la transformation τ définie par $\tau x \to 2x \bmod 1$ sur $\mathbb{R}/\mathbb{Z}$ muni de la mesure de Lebesgue, pour tout $n \geq 1$, l'opérateur $T^n P^n = \Pi_n$ est la projection orthogonale sur le sous-espace des fonctions f invariantes par le groupe dyadique d'ordre n, c'est-à-dire telles que $f(x + k2^{-n}) = f(x), \forall k = 0, 1, ..., 2^n - 1$.

Il est obtenu par moyennisation sur l'orbite du groupe dyadique d'ordre n. La convergence $\lim_n P^n f = \mu(f)$, établie en 5.4 pour les fonctions höldériennes, s'étend aux fonctions dans L^2

et montre que le système est exact. Notons que, pour une fonction arbitraire de L^2, on ne dispose pas de vitesse de convergence.

2) Etant donné un alphabet fini I, considérons le décalage unilatère sur $I^{\mathbb{N}}$ muni d'une mesure produit. L'opérateur $\Pi_n = T^n P^n$ est ici la projection orthogonale sur le sous-espace des fonctions ne dépendant pas des n premières coordonnées. Le système dynamique associé est exact.

Décomposition d'une fonction pour un système exact

Pour toute $f \in L^2(\mu)$, la série $\sum_{k=0}^{\infty}[\mathbb{E}(f|\tau^{-k}\mathcal{E}) - \mathbb{E}(f|\tau^{-(k+1)}\mathcal{E})]$ est une série orthogonale qui, dans le cas d'un système exact, converge en norme L^2 vers $f - \mu(f)$.

Choisissons une base orthonormée $(f_j)_{j\in J}$ de $L^2(\mathcal{E},\mu)\ominus L^2(\tau^{-1}\mathcal{E},\mu)$. On peut montrer que l'ensemble d'indices J indexant la base est infini dénombrable. Les fonctions $(T^n f_j)_{n\geq 0, j\in J}$ forment ainsi une base orthonormée de l'orthogonal des constantes dans $L^2(\mathcal{E},\mu)$ (propriété de spectre de Lebesgue).

L'étude effectuée au paragraphe précédent s'applique à la "filtration" $(\mathcal{H}_n = L^2(\tau^{-n}\mathcal{E},\mu), n \geq 0)$. En particulier, le théorème 11.4.5 peut être réénoncé sous la forme suivante :

Lemme 11.4.9 *???Pour toute $f \in L^2(\mu)$, les conditions suivantes sont équivalentes :*
i) la suite $(\sum_0^{n-1} P^k f)_{n\geq 1}$ est bornée dans $L^2(\mu)$,
ii) il existe $g \in L^2(\mu)$ telle que $f = (I-P)g$,
iii) il existe h et $g_1 \in L^2(\mu)$ telles que $f = g_1 + Th - h$ et $Pg_1 = 0$.

Lorsque ces conditions sont satisfaites, on a $g_1 = g - TPg$ et $h = Pg$ et f est homologue dans $L^2(\mu)$ à une fonction g_1 engendrant une différence de martingale.

Remarque 11.4.10 Sous les conditions du lemme, nous avons : $|\|\sum_0^{n-1} T^k f\| - \sqrt{n}\|g_1\|| \leq 2\|h\|$ et la variance $\sigma^2(f)$ est donnée par :

$$\begin{aligned}\sigma^2(f) = \|g_1\| &= \|g - TPg\|^2 = \|g\|^2 - \|Pg\|^2 \\ &= \int [P(g^2) - (Pg)^2]\ \mu(dx).\end{aligned}$$

Elle est nulle si, et seulement si, $g = TPg$, condition équivalente au fait que f est un cobord dans L^2.

Quand on peut exprimer l'adjoint de T comme un opérateur de transition, d'après la dernière égalité, la condition $\sigma(f) = 0$ est aussi équivalente à : *pour μ-presque tout $x \in E$, $g(y) = Pg(x)$, pour $P(x,\cdot)$-presque tout $y \in E$.*

Systèmes inversibles, K-systèmes

Considérons un système dynamique inversible **ergodique** $(E,\mathcal{E},\mu,\tau)$.

Pour toute sous-tribu $\mathcal{F}$ de $\mathcal{E}$ par rapport à laquelle τ est mesurable (i.e. $\tau^{-1}\mathcal{F} \subset \mathcal{F}$), le système dynamique $(E,\mathcal{F},\mu,\tau)$ est non inversible. Le dual de l'opérateur T de composition par τ opérant dans $L^2(E,\mathcal{F},\mu)$ est l'opérateur P défini par

$$\forall g \in L^2(E,\mathcal{F},\mu),\ Pg = \mathbb{E}[g|\tau^{-1}\mathcal{F}] \circ \tau^{-1}.$$

Le TCL est vérifié pour toute fonction f de $L^2(E,\mathcal{F},\mu)$ telle que

$$\int_E f\ d\mu = 0 \quad \text{et} \quad \sup_{n\geq 1} \|\sum_{k=0}^{n-1} \mathbb{E}[f|\tau^{-k}\mathcal{F}] \circ \tau^{-k}\|_2 < +\infty.$$

Le résultat est s'applique aux fonctions $\mathcal{F}$-mesurables. Nous allons étendre la classe de fonctions en appliquant le lemme 9.3.1.

Théorème 11.4.11 *Soit $(E, \mathcal{E}, \mu, \tau)$ un système dynamique inversible ergodique. Soit $\mathcal{F}$ une sous-tribu de $\mathcal{E}$ telle que $\tau^{-1}\mathcal{F} \subset \mathcal{F}$ et $\vee_{n\in\mathbb{Z}} \tau^n \mathcal{F} = \mathcal{E}$. Soit f une fonction de $L^2(E, \mathcal{E}, \mu)$ vérifiant $\int_E f \, d\mu = 0$ et telle que*

$$\sup_{n\geq 1} \|\sum_{k=0}^{n-1} E[f|\tau^{-k}\mathcal{F}] \circ \tau^{-k}\|_2 < +\infty \text{ et } \sup_{n\geq 1} \|\sum_{k=0}^{n-1} (f - E[f|\tau^k\mathcal{F}]) \circ \tau^k\|_2 < +\infty.$$

Alors la suite $(\|\frac{S_n f}{\sqrt{n}}\|_2)_{n\geq 1}$ converge vers un réel σ, qui est strictement positif si et seulement si f n'est pas un cobord et f vérifie le TCL :

$$\frac{1}{\sqrt{n}} S_n f(.) \overset{d/\mu}{\to} \mathcal{N}(0, \sigma^2).$$

La structure de K-système garantit l'existence de "beaucoup" de fonctions pour lesquelles le théorème 11.4.11 s'applique.

Considérons un système (inversible) Si $(E, \mathcal{E}, \mu, \tau)$ a la propriété de *K-système* (cf. 10.3.2), on peut appliquer ce qui précède à une filtration de K-système $(\tau^n \mathcal{F})$.

Toute fonction f dans $L^2(\mu)$ d'intégrale nulle se décompose en série orthogonale :

$$f = \sum_{n\in\mathbb{Z}} [\mathbb{E}(f|\tau^{-n}\mathcal{F}) - \mathbb{E}(f|\tau^{-(n+1)}\mathcal{F})]. \tag{11.22}$$

Si f est F-mesurable, la série se réduit aux termes d'indices < 0. On peut appliquer à la filtration le théorème 11.4.6.

Les fonctions f appartenant à $L^2(\mathcal{F}) \ominus L^2(\tau^{-1}\mathcal{F})$ sont telles que les sommes ergodiques $(S_n f)$ forment une martingale. Nous avons vu (théorème 11.1.7) qu'elles vérifient le TCL, avec la variance $\|f\|^2$. Nous allons maintenant étendre la classe des fonctions vérifiant le TCL.

Soit f appartenant au sous-espace $L^2(\mathcal{F}) \ominus L^2(\tau^{-k}\mathcal{F})$, $k > \ell$. Alors la fonction $f + Tf + ... + T^{k-1}f$ engendre un suite de différences de martingale, sous l'action de la transformation τ^k. Par décalage, cette propriété s'étend à $\mathcal{W}_{k,\ell} = L^2(\tau^{-\ell}\mathcal{F}) \ominus L^2(\tau^{-k}\mathcal{F})$, $k > \ell$.

On obtient des classes de fonctions vérifiant le TCL, en renforçant la condition de convergence dans la série (11.22). Pour expliciter ces classes de fonctions, on peut soit utiliser l'approximation par des fonctions de $\mathcal{W}_{k,\ell}$ du sous-espace vectoriel $\mathcal{R}$ réunion des sous-espaces $\mathcal{W}_{k,l}$ au sens de la "semi-norme" définie par $\sigma()$ et appliquer le lemme 11.2.7, soit essayer d'écrire la fonction étudiée comme la somme d'un cobord et d'une fonction dans $\mathcal{W}$. Ceci conduit aux deux résultats suivants (Gordin [Gor69]) :

Théorème 11.4.12 *Soit $f \in L^2(\mu)$ telle qu'il existe une suite (g_j) dans $\mathcal{R}$ vérifiant :*

$$\lim_j \sigma(f - g_j) = 0.$$

Alors la variance $\sigma^2(f)$ existe. Lorsque cette variance est non nulle, f vérifie le TCL.

Théorème 11.4.13 *Soit $f \in L^2(\mu)$, telle que,*

$$\sum_0^\infty (\|\mathbb{E}(f|\mathcal{F}_n)\|_2 + \|f - \mathbb{E}(f|\mathcal{F}_{-n})\|_2) < \infty,$$

alors f est de la forme $f = g + Th - h$, avec $g \in L^2(\mathcal{F}) \ominus L^2(\tau^{-1}\mathcal{F})$ et f vérifie le TCL.

Le lemme 9.3.1 permet aussi d'obtenir plus généralement :

Théorème 11.4.14 *Soit $f \in L^{2+\delta}(\mu)$ telle que, pour un $\delta \in [0, \infty[$, on ait :*

$$\sum_0^\infty \|\mathbb{E}(f|\mathcal{F}_n)\|_{(2+\delta)/(1+\delta)} + \sum_0^\infty \|f - \mathbb{E}(f|\mathcal{F}_n)\|_{(2+\delta)/(1+\delta)} < \infty, \tag{11.23}$$

alors f est de la forme $f = g + Th - h$, avec $g \in L^2(\mathcal{F}) \ominus L^2(\tau^{-1}\mathcal{F})$ et $h \in L^{\frac{2+\delta}{1+\delta}}(\mu)$. En particulier f vérifie le TCL.

On note que, sous les hypothèses du théorème 11.4.13 ou du théorème 11.4.14, les conditions du théorème (11.4.12) sont vérifiées, mais que la représentation de f sous la forme $g + Th - h$ donne une information plus précise. En particulier, sous les hypothèses du théorème (11.4.13) (resp. du théorème 11.4.14), si $\sigma^2(f) = 0$, ou si f est un cobord mesurable, alors f est un cobord dans $L^2(\mu)$ (resp. dans $L^{\frac{2+\delta}{1+\delta}}(\mu)$).

Dans l'étude de systèmes dynamiques explicites, le problème est de vérifier les conditions d'applications des théorèmes précédents. Même si l'on sait que le système est un K-système et donc possède des fonctions vérifiant le TCL, l'étude doit être reprise pour chaque classe de fonctions spécifique. C'est par exemple le cas pour un système différentiable sur une variété et pour la classe des fonctions différentiables.

Extension aux sous-suites

Donnons un exemple d'extension du TCL qui reste valide pour des fonctions homologues dans $L^2(\mu)$) à une différence de martingale.

Théorème 11.4.15 *Si une fonction f est homologue dans $L^2(\mu)$) à une fonction engendrant une différence de martingale, si $(k_n(x))_{n\geq 1}$ est une suite strictement croissante de fonctions mesurables à valeurs dans $\mathbb{N}$ telle que la suite $(\frac{k_n(x)}{n})$ converge en probabilité vers une constante $a \in [1, \infty[$, on a :*

$$\frac{1}{\sqrt{n}} S(k_n(.), .) \xrightarrow{d} \mathcal{N}(0, \sigma^2/a).$$

<u>Preuve</u> : Soit f dans L^2 telle que l'on ait $f = g + Th - h$, avec h dans $L^2(\mu)$. On a alors :

$$Sf(k_n(x), x) = Sg(k_n(x), x) + h(T^{k_n(x)}x) - h(x).$$

On applique le théorème (11.1.8). Par ailleurs, le terme de cobord peut être contrôlé par la majoration :

$$\begin{aligned} \int_{|k_n(.)-[an]|<n\epsilon} \frac{1}{n} h^2(T^{k_n(x)}x)\, d\mu(x) &\leq \frac{1}{n} \sum_{[an]-\epsilon n \leq \ell \leq [an]+\epsilon n} \int h^2(T^\ell x)\, d\mu(x) \\ &\leq \epsilon \frac{2n}{n} \int h^2\, d\mu = 2\epsilon \|h\|_2^2. \end{aligned}$$

□

11.5 Commentaires

Les travaux sur le théorème de la limite centrale pour les processus stationnaires et les systèmes dynamiques sont nombreux.

Les premiers résultats concernant le TCL pour les martingales, après les premiers théorèmes limites de Bernstein (1927), Lévy (1935, 1937), sont dus à Billingsley (1961) et Ibragimov (1963) qui ont montré indépendamment le TCL pour des différences de martingales stationnaires et ergodiques (théorème 11.1.6 ci-dessous).

En 71, B.M. Brown a donné un énoncé général, dans un cadre non nécessairement stationnaire, sous l'hypothèse de la convergence de la "variance conditionnelle" et une "condition de Lindeberg". Ces conditions sont satisfaites dans le cas stationnaire ergodique traité par Billingsley et Ibragimov.

On trouvera dans M. Denker ([Denk89]) une revue de résultats et de méthodes. Mentionnons également P. Crépel [??] pour une revue et une discussion des théorèmes limites dans le cas stationnaire et non stationnaire.

Les méthodes

Il est possible de montrer que, dans tout système dynamique ergodique (cf. []), il existe des fonctions f telle que le comportement en loi asymptotique des sommes $(S_n f)$ soit régi par le TCL. Cependant, il paraît raisonnable de chercher à établir le TCL pour des systèmes définis sur des espaces métriques et pour des fonctions régulières sur ces espaces. La propriété de TCL apparaît alors, à travers le comportement des processus ainsi engendrés, comme un indicateur de la stochasticité du système dynamique.

Plusieurs types de méthodes permettent de prouver un TCL pour certains processus stationnaires :

- Une méthode basée sur les propriétés spectrales de l'opérateur de transfert. Elle s'applique à l'étude des processus stationnaires assez "réguliers" définis à l'aide d'une mesure de Gibbs et permet de traiter de nombreux exemples de nature géométrique et algébrique, quand la construction de "partitions markoviennes" rend possible l'utilisation d'un codage ramenant au cas des mesures de Gibbs. Dans ce cas, des résultats beaucoup plus précis que le TCL, tels que le théorème de la limite locale, peuvent être obtenus. Citons par exemple :

Théorème 11.5.1 *(théorème local) (Guivarc'h-Hardy [GuiHar88]) Soit f une fonction höldérienne sur le sous-shift muni d'une mesure de Gibbs μ associée à une fonction höldérienne. Si f est d'intégrale nulle et apériodique, pour tout intervalle I de longueur $\ell(I)$, on a :*

$$\sigma\sqrt{2\pi n}\ \mu(\sum_0^{n-1} T^k f \in I) \to \ell(I).$$

- Les méthodes de mélange basées sur l'utilisation de coefficients de mélange, qui expriment une propriété d'indépendance asymptotique du processus (voir par exemple Hall et Heyde [HaHe80], Doukhan ([Douk94]). Cette approche s'applique à des processus stationnaires ou non, mais, dans le cas des processus définis par les systèmes dynamiques, paraît surtout utilisables dans les situations où les méthodes d'opérateurs s'appliquent également.

- Une troisième méthode, moins exigeante que les deux précédentes, est la méthode des différences de martingale. Pour certains processus stationnaires ergodiques, elle permet d'obtenir le TCL (théorème de Billingsley-Ibragimov) ainsi que plusieurs résultats renforçant ce théorème. Le champ d'application de cette méthode est élargi par le procédé introduit par I.M. Gordin, qui consiste à exprimer, sous certaines conditions, un processus comme une somme d'accroissements de martingales et d'un terme "cobord", dont on contrôle les fluctuations. Cette méthode, bien adaptée au traitement du cas stationnaire, s'applique à des situations où les méthodes de mélange ne peuvent être employées.

C'est cette dernière méthode que nous avons présenté ici : si l'on peut représenter la fonction f sous la forme $f = g + Th - h$, avec h dans $L^2(\mu)$, et g engendrant un processus $(g \circ \tau^n)$ qui est une "suite de différences de martingales", certaines propriétés asymptotiques vérifiées par les différences de martingales le sont encore par le processus $(f \circ \tau^n)$ engendré par f. On obtient ainsi une classe étendue de fonctions pour lesquelles les sommes $(\sum_0^{n-1} T^k f)$ vérifient de bonnes propriétés stochastiques.

Décrivons enfin quelques cadres naturels, dans lesquels ces différentes méthodes peuvent être mises en œuvre pour obtenir un TCL et des propriétés dérivées :

- Un cadre abstrait qui est celui des systèmes exacts et des K-systèmes. Ces systèmes ont des propriétés stochastiques fortes qui se traduisent notamment en terme d'entropie. Par la méthode de différences de martingale, on montre que "beaucoup" de fonctions dans ces systèmes vérifient le TCL.

- Le cadre plus explicite des processus définis sur un espace produit muni d'une mesure stationnaire (cette mesure contient toutes les informations statistiques sur le processus aléatoire sous-jacent et sa stochasticité). Un cas important, généralisant l'exemple des mesures produits, est celui des mesures de Gibbs sur les "sous-shift de type fini". On peut alors, considérer le processus des applications coordonnées, ou plus généralement le processus défini à l'aide d'une fonction de toutes les coordonnées et non pas seulement de la première. Cette situation peut être traitée par des méthodes de coefficients de mélange ; mais aussi par une méthode d'opérateurs, ou par réduction à une différence de martingale.

- Le cadre des systèmes différentiables, tels que les systèmes hyperboliques, pour lesquels il est naturel de chercher à établir des théorèmes limites pour des fonctions assez régulières. L'un des procédés pour traiter ce problème est celui des partitions markoviennes. Il est également possible de traiter le cas non-hyperbolique, quand les "feuilles instables" du système possèdent de bonnes propriétés d'équirépartition.

Renforcements du TCL

Le TCL est le premier résultat dans une famille de théorèmes qui peuvent être obtenus sous des hypothèes convenables dans le cadre des martingales. Ceci montre l'intérêt des méthodes de martingales, quand elles peuvent être mises en oeuvre.

Mentionnons au passage qu'une partie des résultats énoncés dans ce paragraphe peut être obtenue par la méthode du plongement dans un processus brownien de Skorokod étendu par Strassen aux martingales. Cette méthode est employée en particulier dans Scott, Hall et Heide [HaHe80]. On notera que, dans le cas de processus ou de fonctions à valeurs vectorielles, on ne dispose pas d'un plongement de Skorokhod "multidimensionnel".

Parmi les renforcements classiques du TCL, mentionnons :
- le théorème central limite fonctionnel (ou principe d'invariance).
- la loi du logarithme itéré,
- le principe d'invariance presque-sûr,
- une vitesse d'approximation par un processus brownien.

Ces différentes propriétés peuvent être établies en particulier pour les suites stationnaires ergodiques d'accroissement de martingales. Nous les présentons brièvement. Une référence est le livre de Hall et Heyde [HaHe80] de 1980, qui fait le point sur l'application de la méthode des martingales au TCL et aux résultats annexes. Notons cependant quelques résultats plus récents, en particulier sur la vitesse de convergence. ? ?

Deux autres extensions classiques dans le cas de v.a. indépendantes (vitesse de convergence, théorème local) nécessitent des hypothèses plus fortes.

• **Théorème central limite fonctionnel (ou principe d'invariance)**

Soit (X_n) une suite de v.a.. Notons $(S(n,t,.))_{n\geq 1}$ le processus interpolé aux points $(\frac{k}{n}, \frac{1}{\sqrt{n}}S(k,x))$ construit à partir des sommes $S_n := \sum_1^n X_k$:

$$S(n,t,x) = S([nt],x) + (nt-[nt])f(T^{[nt]+1}x), \ t \in [0,1].$$

L'application $\omega \to (t \to \frac{1}{\sigma\sqrt{n}}S(n,t,.))_{t\in[0,1]}$ est une v.a. à valeurs dans l'espace $S = C([0,1])$ muni de la distance uniforme. Soit P_n la mesure image (la loi de cette v.a.). La convergence faible de la suite (P_n) vers la loi P du mouvement brownien standard W sur $[0,1]$ (mesure de Wiener) signifie que, pour toute fonction continue h sur l'espace S, $\lim_n \int_S h\ dP_n \to \int_S h\ dP$. De façon équivalente, on a $\lim_n P_n(A) = P(A)$, pour tout A borélien de S de frontière P-négligeable.

Si cette convergence faible est réalisée, on dit que la suite converge en loi ou en distribution vers W et on note :

$$(\frac{1}{\sigma\sqrt{n}}S(n,t,.))_{t\in[0,1]} \xrightarrow{d} (W_t)_{t\in[0,1]}. \tag{11.24}$$

Elle implique par exemple le corollaire suivant : prenons comme fonction "test" continue dans la convergence faible la fonction $h(x) = \sup_{t\in[0,1]} x(t)$, $x \in S$, on obtient alors :

$$\frac{1}{\sigma\sqrt{n}} \sup_{k\leq n} S(n,t.) \xrightarrow{d} \sup_{t\in[0,1]} W_t.$$

La convergence dans (11.24) est liée à une propriété de **tension**. Plus précisément, pour avoir (11.24), il faut et il suffit que les deux conditions suivantes soient vérifiées :

i) Pour tout entier k et tout k-uple $(t_1 < ... < t_k)$, $0 \leq t_1 < ... < t_k \leq 1$,

$$(\frac{1}{\sigma\sqrt{n}}S(n,t_1,.), ..., \frac{1}{\sigma\sqrt{n}}S(n,t_k,.)) \xrightarrow{d} (W_{t_1}, ..., W_{t_k}), \tag{11.25}$$

(convergence des distributions fini-dimensionnelles),

ii) Pour tout $\varepsilon > 0$, on a :

$$\mu(\{x : \sup_{|s-t|<\delta} |S(n,s,x) - S(n,t,x)| > \varepsilon\sqrt{n}\}) \to 0, \tag{11.26}$$

uniformément en n, quand δ tend vers 0 (condition de tension).

La convergence donnée par (11.24) (théorème de Donsker pour le cas indépendant (1951)) a été étendue par Billingsley (1962) aux différences de martingales stationnaires ergodiques. Une version plus générale a été donnée en 71 par Brown [Brow71]. Dans le cas stationnaire le résultat est le suivant :

Théorème 11.5.2 *Les hypothèses du théorème 11.1.4, impliquent la convergence (11.24).*

- **Loi du logarithme itéré**

On note K l'ensemble des fonctions $x \in C[0,1]$ absolument continues telles que $x(0) = 0$ et $\int_0^1 x'(t)^2 \, dt \leq 1$. Cet ensemble est compact.

Soit $(S(n,t,.))$ le processus interpolé. Notons $\phi(t) := (2t \ \ln \ln t)^{\frac{1}{2}}$.

Théorème 11.5.3 *Sous les hypothèses du théorème 11.1.4, la suite* $([\phi(n\sigma^2)]^{-1} S(n,t,.))$ *est presque sûrement relativement compacte et admet* K *pour ensemble limite.*

Chapitre 12

Opérateurs et mesures invariantes

Nous examinons l'existence de mesures de probabilités invariantes pour une transformation opérant sur un espace et donnons des méthodes de construction de ces mesures.

12.1 Mesures invariantes sur un espace compact

Commençons par la question des mesures invariantes dans le cas des opérateurs markoviens sur les espaces de fonctions continues (la composition par une transformation continue est un cas particulier de cette situation).

Soit (E, d) un espace métrique muni de sa tribu borélienne $\mathcal{E}$. Notons $\mathcal{C}(E)$ l'espace des fonctions continues sur E, à valeurs réelles ou complexes, $\mathcal{C}_b(E)$ l'espace de Banach des fonctions continues bornées sur E, muni de la norme uniforme, $\|f\|_\infty = \sup_{x \in E} |f(x)|$, et par $\mathcal{P}(E)$ l'espace des mesures de probabilité sur la tribu borélienne de E.

Soit P un *opérateur markovien* opérant sur $\mathcal{C}_b(E)$: $Pf \in \mathcal{C}_b(E)$, $\forall f \in \mathcal{C}_b(E)$, $P1 = 1$ et $Pf \geq 0$, pour $f \geq 0$.

L'opérateur P opère par dualité sur l'espace $\mathcal{P}(E)$ des probabilités sur les boréliens de E. Si ν est dans $\mathcal{P}(E)$, nous noterons, νP l'action de P sur ν définie par :

$$\int f(x)\ \nu P(dx) = \int Pf(x)\ \nu(dx).$$

Losque l'espace (E, d) est compact, il existe toujours au moins une mesure de probabilité P-invariante sur E. En effet, pour tout $x \in E$, toute valeur d'adhérence de la suite $(\frac{1}{n} \sum_0^{n-1} \delta_x P^k)$ est une mesure de probabilité sur E qui est invariante par P. Plus précisément, on peut faire la construction suivante :

Proposition 12.1.1 *Pour tout $x \in E$, il existe une mesure P-invariante μ et une sous-suite strictement croissante (n_j) telle que*

$$\lim_j \frac{1}{n_j} \sum_{k=0}^{n_j - 1} P^k f(x) = \mu(f),\ \forall f \in \mathcal{C}(E).$$

Preuve : Soit D une famille dénombrable de fonctions f, dense pour la norme uniforme dans $\mathcal{C}(E)$. Pour tout point x, par le procédé diagonal, on peut construire une suite (n_j) telle que la sous-suite des moyennes

$$M_{n_j} f(x) = \frac{1}{n_j} \sum_{k=0}^{n_j-1} P^k f(x)$$

converge, pour toute fonction f dans D. Comme P est une contraction pour la norme uniforme et par densité de D, on en déduit que, pour toute fonction continue f, la suite des moyennes $(M_{n_j} f)$ est de Cauchy pour la norme uniforme, donc converge uniformément. $\square$

Un cas important est celui où il y a unicité de la mesure invariante par P.

Proposition 12.1.2 *Soient (E, d) un espace métrique compact. et P un opérateur markovien opérant sur les fonctions continues sur E. Les conditions suivantes sont équivalentes :*
a) il existe une unique probabilité μ invariante par P,
b) pour toute fonction $f \in \mathcal{C}(E)$, pour tout $x \in E$, la suite des moyennes définies par

$$M_n f(x) = \frac{1}{N} \sum_{k=0}^{n-1} P^k f(x)$$

converge vers une constante,
c) la suite des moyennes converge vers une constante uniformément en x.

Preuve : On peut soit reprendre la méthode de la proposition précédente, soit utiliser le théorème d'Hahn-Banach de la façon suivante :

Soit $V_0 = \overline{\{h - Th, h \in \mathcal{C}(E)\}}$, le sous-espace fermé de $\mathcal{C}(E)$ formé des fonctions qui sont limite uniforme de "cobords". Une mesure μ est invariante si et seulement si $\mu(g) = 0$, $\forall g \in V_0$.

Le sous-espace $\mathcal{M}(\tau)$ des mesures τ-invariantes s'identifie donc à l'orthogonal (au sens de la dualité) de V_0. Par le théorème de Hahn-Banach, il est de dimension 1 si, et seulement si, $\mathbb{C} + V_0$ est égal à $\mathcal{C}(E)$.

En effet, soit μ_0 une mesure de probabilité P-invariante fixée. Considérons le sous-espace fermé V_1 formé des fonctions continues g telles que $\mu_0(g) = 0$. Supposons que V_0 ne soit pas égal à V_1. Soit $f_0 \notin V_0$ telle que $\mu_0(f) = 0$. Il existe (Hahn-Banach) une forme linéaire continue ν nulle sur V_0 et telle que $\nu(f_0) \neq 0$.

La forme linéaire ν s'écrit $\nu = \nu^+ - \nu^-$, où ν^+ et ν^- sont des mesures positives. Pour une fonction continue positive f, on a $\nu^+(f) = \sup\{\nu(g), 0 \leq g \leq f\}$. Ceci implique $\nu^+(Pf) \geq \nu^+(f)$, et donc aussi, $\nu^+(P(\|f\|_\infty - f)) \geq \nu^+(\|f\|_\infty - f)$. Comme $P1 = 1$, ces inégalités sont des égalités : $\nu^+(Pf) = \nu^+(f)$.

Les mesures positives ν^+ et ν^- sont donc P-invariantes et l'une des deux au moins n'est pas un multiple de μ_0. $\square$

Nous mettrons en évidence des situations où, pour toute fonction continue f, la suite $(P^n(f))_{n \in \mathbb{N}}$ des itérés de l'opérateur est elle-même convergente.

• Cas d'un système dynamique topologique

Nous faisons maintenant le lien avec l'étude des transformations. Soient (E, d) un espace métrique compact muni de sa tribu borélienne $\mathcal{E}$ et τ une application continue surjective de E sur lui-même.

On dit que le couple (E, τ) forme un *système dynamique topologique*.

Nous recherchons une ou des mesures de probabilité sur la tribu borélienne $\mathcal{E}$ de E invariantes par τ ; c'est-à-dire, de façon équivalente, des mesures de Radon μ, positives de masse 1 sur E, vérifiant pour toute fonction f continue sur E :

$$\int_E f(\tau x)\ \mu(dx) = \int_E f(x)\ \mu(dx).$$

Pour une telle mesure invariante μ, nous obtenons un système dynamique mesuré $(E, \mathcal{E}, \tau, \mu)$.

Notons $\mathcal{P}(E, \tau)$ l'ensemble des mesures de $\mathcal{P}(E)$ qui sont invariantes par τ.

L'opérateur T de composition par τ est un opérateur markovien opérant sur les fonctions continues. On peut lui appliquer les résultats du paragraphe précédent. D'autres opérateurs markoviens apparaîssent dans ce contexte, comme duaux vis à vis d'une mesure invariante de l'opérateur T. Nous en verrons des exemples, ainsi qu'une discussion générale plus loin.

Le convexe fermé $\mathcal{P}(E, \tau)$ est toujours non vide, d'après le théorème de Markov-Kakutani ou la proposition 12.1.1 appliquée à l'opérateur T. L'extrémalité d'une mesure dans ce convexe est caractérisée par la propriété d'ergodicité.

Lemme 12.1.3 *Les points extrémaux du convexe $\mathcal{P}(E, \tau)$ sont les mesures de probabilité μ pour lesquelles le système mesuré $(E, \mathcal{E}, \mu, \tau)$ est ergodique.*

<u>Preuve</u> : En effet, si le système n'est pas ergodique, il existe un ensemble A tel que $\tau^{-1}A \subset A$ et $0 < \mu(A) < 1$. La mesure μ est alors le barycentre des mesures μ_1 et μ_2 définies par $\mu_1 = \mu(A)^{-1}1_A\mu$, $\mu_2 = \mu(A^c)^{-1}1_{A^c}\mu$, ce qui contredit l'extrémalité.

Inversement, si μ n'est pas extrémale, μ est le barycentre non trivial de deux mesures de probabilité invariantes par τ. Les densités de ces mesures par rapport à la mesure μ sont des fonctions mesurables bornées, invariantes par l'opérateur adjoint de T, qui est une contraction de $L^2(\mu)$, donc par T, non constantes μ-presque partout. $\square$

L'existence de points extrémaux dans le convexe $\mathcal{P}(E, \tau)$ assure ainsi l'existence d'au moins une mesure de probabilité μ invariante par τ pour laquelle le système $(E, \mathcal{E}, \mu, \tau)$ est ergodique.

• Points génériques

La preuve de la proposition 12.1.1 pose la question de l'existence de points x tels que la limite de la suite des moyennes $(\frac{1}{n}\sum_{k=0}^{n} f(\tau^k x))$ existe, pour toute fonction f continue, non pas seulement sur une sous-suite, mais pour la suite des entiers naturels.

Ces points seront appelés **points génériques**. Si x est générique, il "engendre" une mesure de probabilité invariante μ, la mesure définie par

$$f \in \mathcal{C}(E) \to \mu(f) = \lim_n \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k x).$$

On dit alors que x est **μ-générique**. L'existence de points génériques est une conséquence de l'existence d'au moins une mesure τ-invariante ergodique et du théorème de Birkhoff qui implique le résultat suivant :

Proposition 12.1.4 *Si E est un espace métrique compact, pour toute probabilité μ τ-invariante ergodique, l'ensemble des points génériques est de μ-mesure égale à 1.*

Exemples 12.1.5

Rappelons deux exemples de transformations déjà étudiés en 8.5.

Rotation

Soit ρ_α la rotation d'angle α sur le cercle unité $E = \mathbb{T}^1$ identifié à $\mathbb{R}/\mathbb{Z}$, muni de sa tribu borélienne. La transformation laisse la mesure de Lebesgue λ invariante et elle est inversible. Pour α irrationnel, le système $(\mathbb{R}/\mathbb{Z}, \rho_\alpha, \lambda)$ est donc ergodique.

Nous avons vu également que λ est la seule mesure de probabilité invariante par ρ_α. (Si μ est une mesure de probabilité invariante sur $\mathbb{T}$, ses coefficients de Fourier vérifient $\widehat{\mu}(n) = e^{2\pi i n\alpha}\widehat{\mu}(n)$, d'où $\widehat{\mu}(n) = 0, \forall n \neq 0$.)

Un exercice classique utilisant la densité (pour la norme uniforme) des polynômes trigonométriques dans $\mathcal{C}(\mathbb{T})$ montre que, si α est irrationnel, pour tout $x \in \mathbb{T}$ et toute fonction continue 1-périodique, on a $\lambda(f) = \lim_n \frac{1}{n}\sum_{k=0}^{n-1} f(x + k\alpha)$. En d'autre terme, pour une rotation irrationnelle, tout point est λ-générique, i.e. engendre la mesure uniforme sur le cercle, ce qui revient à dire, ici, que, pour tout x, la suite $(x + k\alpha)_{k\geq 0}$ est **équirépartie** (ou équidistribuée) modulo 1.

Cette propriété peut être vue comme une conséquence de l'unique ergodicité des rotations irrationnelles.

Exercice : Etudier les rotations sur un groupe compact abélien G muni de sa mesure λ normalisée invariante par les translations (la mesure de Haar). En utilisant les caractères sur le groupe G, montrer qu'un élément $\alpha \in G$ définit par translation une transformation ρ_α ergodique si, et seulement si, l'ensemble $\{n\alpha,\ n \in \mathbb{Z}\}$ est dense dans G.

Multiplication par 2 mod 1

On considère le cercle unité $\mathbb{T}$ identifié à $\mathbb{R}/\mathbb{Z} = [0, 1[$, muni de sa tribu des boréliens $\mathcal{E}$ et de la mesure de Lebesgue notée λ. On désigne par τ la transformation $x \to 2x \bmod 1$, opérant sur $\mathbb{T}$. Toute fonction sur $\mathbb{T}$ s'identifie à une fonction sur $\mathbb{R}$ périodique et de période 1.

Nous avons montré en 8.5 que le système dynamique $(\mathbb{T}, \mathcal{E}, \lambda, \tau)$ est ergodique. Dans cet exemple, l'opérateur adjoint (pour la mesure λ) est défini par

$$Pf(x) = \frac{1}{2}\ (f(\frac{x}{2}) + f(\frac{x+1}{2})).$$

Les points λ-génériques pour la transformation $x \to 2x \bmod 1$ sont les nombres **normaux** en base 2. L'ergodicité du système et le théorème ergodique permettent d'affirmer que *presque tout nombre est normal en base 2 (et, par le même raisonnement, normal dans toute base b entière).* On retrouve ainsi le théorème d'E. Borel, qui peut être vu aussi comme un cas particulier de la loi forte des grands nombres.

Contrairement au cas d'une rotation irrationnelle sur le cercle, la transformation $\tau : x \to 2x \bmod 1$ possède "beaucoup" de mesures invariantes, et en particulier des mesures discrètes périodiques (cf. 12.3.2).

12.2 Unique ergodicité

Dans le cas où le convexe $\mathcal{P}(E, \tau)$ est réduit à un unique élément μ (le système est dit alors **uniquement ergodique**), pour toute fonction f de $\mathcal{C}(E)$, la suite des moyennes ergodiques

$(\frac{1}{n}S_n f(x))$ converge uniformément en $x \in E$ vers la constante $\mu(f)$. Ceci est un cas particulier de la proposition 12.1.2 :

Proposition 12.2.1 *Soit (E, d) un espace métrique compact. Pour toute transformation continue τ opérant sur E, les conditions suivantes sont équivalentes :*
a) il existe une unique probabilité μ invariante par τ,
b) pour toute fonction $f \in \mathcal{C}(E)$, pour tout $x \in E$, les moyennes $M_n f(x) = \frac{1}{n}\sum_{k=0}^{n-1} f(\tau^k x)$ convergent vers une constante,
c) la convergence des moyennes ergodiques vers une constante dans b) est uniforme en x.

Introduisons maintenant quelques notions purement topologiques :

Définitions 12.2.2

La transformation τ est dite *minimale*, si, pour tout $x \in E$, l'orbite de x sous l'action de τ, $\{\tau^n x, n \geq 0\}$, est dense dans E.

On dit qu'un système (X, T) est *strictement ergodique* s'il est uniquement ergodique et minimal.

Proposition 12.2.3 *Pour un système dynamique topologique (E, τ) sur un espace compact E (τ étant surjective), les conditions suivantes sont équivalentes :*
i) (E, τ) est minimal,
ii) pour tout fermé F non vide de E tel que $F \subset \tau^{-1}F$, on a $F = E$,
iii) pour tout ouvert U non vide de E tel que $\tau^{-1}U \subset U$, on a $U = E$,
iv) pour tout ouvert non vide U dans E, il existe un entier $N \geq 1$ tel que $E = \bigcup_{k=0}^{N} \tau^{-k}U$,
v) toute mesure τ-invariante charge les ouverts non vides.

Preuve : On vérifie facilement l'équivalence entre *i*), *ii*), *iii*) et *iv*).

La condition *iv*) implique que, pour tout ouvert U non vide et pour toute probabilité μ invariante par τ, on a, pour un entier $N \geq 1$, $1 = \mu(E) \leq N\mu(U)$, d'où $\mu(U) > 0$.

Supposons que *iv*) soit vérifiée. Si F est un fermé τ-invariant non vide ($\tau F \subset F$), il existe une probabilité μ invariante par τ à support dans F. L'ouvert F^c est de μ-mesure nulle, donc vide.

□

Par le lemme de Zorn, on peut montrer que, pour toute application continue τ d'un espace compact E dans lui-même, il existe un fermé F dans E invariant par τ, tel que la restriction de τ à F soit minimale.

Dans le cas de l'unique ergodicité, la minimalité est assurée si l'unique mesure τ-invariante charge les ouverts. Les rotations irrationnelles sur le cercle sont des exemples de systèmes strictement ergodiques. Par contre, une application contractante, sur un espace qui n'est pas réduit à un point, est évidemment uniquement ergodique, mais non minimale.

• Extension ou produit gauche d'un système dynamique

Soient un système dynamique $(E, \mathcal{E}, \mu, \tau)$ et un groupe G localement compact muni d'une mesure de Haar à gauche λ. Formons l'espace produit $E \times G$ muni de la tribu produit et du produit des mesures μ et λ.

Soit ϕ une application mesurable de E dans G. Considérons l'application τ_ϕ de $E \times G$ dans lui-même définie par :

$$(x, g) \to \tau_\phi(x, g) = (\tau x, \phi(x)g).$$

La transformation τ_ϕ laisse invariante la mesure produit $\mu \times \lambda$ et définit donc un système dynamique $(E \times G, \mathcal{E} \times \mathcal{B}, \mu \times \lambda, \tau_\phi)$, qui est une extension de $(E, \mathcal{E}, \mu, \tau)$. On dit que ce système est l'extension (ou le produit gauche) de τ par ϕ.

Considérons en particulier le cas où G est le groupe $\mathbb{T}$ identifié à $\mathbb{R}/\mathbb{Z}$ et donnons un critère d'ergodicité pour une extension.

Proposition 12.2.4 *Si $(E, \mathcal{E}, \mu, \tau)$ est un système ergodique, la transformation τ_ϕ est ergodique pour la mesure produit $\mu \times \lambda$ si, et seulement si, pour tout entier $p \neq 0$, l'équation fonctionnelle*

$$h(\tau x) = e^{2\pi i p \phi(x)} h(x), \text{ pour } \mu\text{-presque tout } x, \tag{12.1}$$

n'a pas de solution mesurable h non nulle presque partout.

Preuve : Si h est une solution mesurable de (12.1), la fonction mesurable $F(x, y) = e^{-2\pi i p y} h(x)$ est une fonction invariante par τ_ϕ non constante.

Inversement, supposons qu'il existe une fonction $F(x, y)$ mesurable invariante par τ_ϕ non constante. On peut supposer F dans $L^2(\mu \times \lambda)$. Développons la fonction F en série de Fourier $F(x, y) = \sum_p f_p(x) e^{2\pi i p y}$, par rapport à la variable y.

Les coefficients de Fourier (en y) sont des fonctions f_p mesurables qui vérifient (12.1), d'après l'invariance de F. L'une au moins de ces fonctions est non nulle μ presque partout. □

Nous allons appliquer ce résultat à l'étude de l'extension de systèmes uniquement ergodiques.

Proposition 12.2.5 *Si (E, τ) est un système uniquement ergodique, d'unique probabilité τ-invariante μ, la transformation τ_ϕ est uniquement ergodique, si et seulement si la mesure produit $\mu \times \lambda$ est ergodique pour τ_ϕ.*

Preuve : Le principe de la preuve est le suivant. Partant d'une mesure τ-invariante ν, on la régularise pour obtenir une mesure invariante ayant une densité par rapport à la mesure produit. Par ergodicité de la transformation τ vis-à-vis de la mesure produit, on obtient que la mesure régularisée est égale à la mesure produit. On conclut, par un argument d'identité approchée, que la mesure ν est la mesure produit.

Plus précisément, soient ν une mesure invariante sur $E \times G$ et, pour ϕ positive sur G bornée et telle que $\lambda(\phi) = 1$, soit $\phi * \nu$ définie par

$$\int_{E \times G} f(x, y)\ d(\phi * \nu)(x, y) = \int_{E \times G} \int_G f(x, y - z) \phi(z)\ d\nu(x, y)\ dz.$$

La commutation de l'action de τ_ϕ et de la rotation en la deuxième variable y implique que la mesure $\phi * \nu$ est τ-invariante.

Si ϕ est bornée par C_ϕ, on a, pour $f \geq 0$,

$$\int_G \int_{E \times G} f(x, y + z) \phi(z)\ d\nu(x, y)\ dz \leq C_\phi \int_{E \times G} (\int_G f(x, z)\ dz)\ d\nu(x, y).$$

Par ailleurs, la mesure $1 * \nu$ coïncide avec la mesure produit $\mu \times \lambda$. En effet, elle est invariante par les rotations du groupe G (sur la variable y) et coïncide avec μ sur les fonctions ne dépendant que de x, d'après l'unique ergodicité de τ.

On en déduit que $d(\phi * \nu) << d\mu\ d\lambda$. D'après l'ergodicité de la mesure produit, $\phi * \nu$ est égale à la mesure produit. En prenant une suite (ϕ_n) formant une identité approchée, on obtient que ν elle-même est égale à la mesure produit. □

Un exemple d'extension

Un exemple simple d'extension est obtenu en formant la transformation ρ_ϕ qui est l'extension d'une rotation par la fonction $\phi(x) = x$, c'est-à-dire :

$$\rho_\phi(x, y) \to (x + \alpha, y + x) \text{ mod } 1.$$

La n-ième itérée est : $(x + n\alpha, y + nx + \frac{1}{2}n(n-1)\alpha)$. La mesure uniforme $\lambda \times \lambda$ est invariante par ρ_ϕ. C'est l'unique mesure sur le tore $\mathbb{T}^2$ invariante par ρ_ϕ, d'après la proposition 12.2.5. On a en effet :

Proposition 12.2.6 *Soit ρ_ϕ la transformation $(x, y) \to (x + \alpha, y + x)$ mod 1. Si α est irrationnel, le système $(\mathbb{T}^2, \lambda \times \lambda, \rho_\phi)$ est ergodique.*

Preuve : Montrons que, pour p entier non nul, l'équation fonctionnelle en h

$$h(\tau x) = e^{2\pi i p x} h(x) \tag{12.2}$$

n'a pas de solution mesurable.

Soit h une solution mesurable de (12.2). Elle est de module invariant par la rotation d'angle α, donc de module constant. Elle est donc dans $L^2(\mathbb{T})$ et on peut la développer en série de Fourier : $h(x) = \sum_k c_k e^{2\pi i k x}$. En reportant dans (12.2), on obtient par identification des coefficients de Fourier :

$$c_k = e^{2\pi i (k+p)\alpha} c_{k+p}, \ \forall k, p \in \mathbb{Z}.$$

En passant au module, on en déduit, pour tout entier ℓ, $|c_k| = |c_{k+\ell p}|$. Les coefficients de Fourier de h tendent vers 0 à l'infini. On a donc, pour p non nul : $|c_k| = \lim_{\ell \to +\infty} |c_{k+\ell p}| = 0, \forall k \in \mathbb{Z}$, ce qui implique que h est nulle presque partout. □

Corollaire 12.2.7 *Si α est irrationnel, le système $(\mathbb{T}^2, \lambda \times \lambda, \rho_\phi)$ est uniquement ergodique.*

Application : *unique ergodicité et équirépartition des suites*

Soit α irrationnel. L'équirépartition de la suite $(n\alpha \text{ mod } 1, \frac{1}{2}n(n-1)\alpha \text{ mod } 1)$ dans le carré résulte du corollaire et de la proposition 12.2.1 qui impliquent :

Corollaire 12.2.8 *Pour tout point (x, y) et pour toute fonction continue $f(x, y)$ 1-périodique en x et y, on a la convergence uniforme :*

$$\lim \frac{1}{N} \sum_{n=0}^{N-1} f(x + n\alpha, y + nx + \frac{1}{2}n(n-1)\alpha) = \int_0^1 \int_0^1 f(u, v) \, dudv.$$

12.3 Etude d'exemples, transformation $x \to \{1/x\}$

A l'opposé des systèmes uniquement ergodiques, nous étudions maintenant deux exemples de systèmes dynamiques possédant beaucoup de mesures invariantes.

Automorphisme du tore

Soit M une matrice, $d \times d$, à coefficients entiers, de déterminant non nul. Elle définit une application linéaire de $\mathbb{R}^d$ sur lui-même qui préserve le réseau $\mathbb{Z}^d$. Par passage au quotient, on obtient un endomorphisme τ_M du tore $\mathbb{T}^d$, muni de la tribu borélienne $\mathcal{E}$, qui est un automorphisme, dans le cas où M est de déterminant ± 1.

Notons que l'application τ_M est toujours surjective. Elle n'est injective que si le déterminant de M est ± 1 (la matrice inverse M^{-1} est à coefficients rationnels ; comme $det(M^{-1})\ det(M) = 1$), si M^{-1} est à coefficients entiers, nécessairement $det(M) = \pm 1$.)

Pour une fonction f continue sur $\mathbb{T}^d$, autrement dit, une fonction de d variables, continue sur $\mathbb{R}^d$, $\mathbb{Z}^d$-périodique, on définit $\lambda(f)$ par

$$\lambda(f) = \int_{\mathbb{T}^d} f(x)\ dx = \int_0^1 \cdots \int_0^1 f(x_1, \cdots, x_d)\ dx_1 \cdots dx_d.$$

On obtient ainsi la mesure uniforme (appelée aussi mesure de Lebesgue, ou mesure de Haar) λ sur le tore, qui est l'unique mesure de probabilité invariante par les translations. Montrons que cette mesure uniforme λ est invariante par τ_M.

Soit y un élement quelconque de $\mathbb{T}^d$, que l'on peut écrire sous la forme $y = Mu$, pour un $u \in \mathbb{T}^d$. On a :

$$\begin{aligned}\int_{\mathbb{T}^d} f(Mx+y)\ dx &= \int_{\mathbb{T}^d} f(Mx+Mu)\ dx \\ &= \int_{\mathbb{T}^d} f(M(x+u))\ dx = \int_{\mathbb{T}^d} f(Mx)\ dx.\end{aligned}$$

La forme linéaire $f \in \mathcal{C}(\mathbb{T}^d) \to \lambda(f \circ \tau_M)$ est donc invariante par les translations du tore. On en déduit qu'elle est égale à la mesure λ.

La preuve de l'ergodicité du système dynamique $(\mathbb{T}^d, \lambda, \tau_M)$ est plus difficile que dans l'exemple de l'endomorphisme du cercle $x \to 2x \bmod 1$. Si la transformation est inversible, la méthode de l'opérateur adjoint ne donne rien. En fait, pour utiliser cette méthode, il faudrait construire un **codage** du système à l'aide d'une "partition markovienne". De façon générale pour un système inversible, le codage permet d'obtenir un modèle symbolique du système dans lequel on peut "casser" l'inversibilité, en considérant l'espace **unilatère** du codage.

Dans l'immédiat, bornons nous à prouver l'ergodicité (en fait la propriété de mélange fort), en utilisant l'analyse de Fourier sur le tore $\mathbb{T}^d$.

Le groupe dual du tore de dimension d s'identifie à $\mathbb{Z}^d$. Un élément γ du groupe dual, associé à un multi-entier $(\gamma_1, \cdots, \gamma_d) \in \mathbb{Z}^d$, est le caractère défini par

$$x \in \mathbb{T}^d \to \gamma(x) = \exp(2\pi i \langle \gamma, x \rangle) = \exp(2\pi i (\sum_{j=1}^d \gamma_j x_j)).$$

Les fonctions $\gamma(.)$, quand γ décrit $\mathbb{Z}^d$, forment une base orthonormée de $L^2(\mathbb{T}^d, \lambda)$. La composition par τ_M corespond, sur les caractères, à l'action de la matrice transposée ${}^tM$ sur $\mathbb{Z}^d$.

Proposition 12.3.1 *Le système dynamique $(\mathbb{T}^d, \lambda, \tau_M)$ est ergodique si et seulement si la matrice M ne possède pas de valeur propre racine de l'unité. Dans ce cas, il vérifie la propriété de mélange fort :*

$$\lim_{n\to\infty} \langle T_M^n f, g \rangle = \mu(f)\overline{\mu(g)}, \forall f, g \in L^2(\lambda). \tag{12.3}$$

<u>Preuve</u> : Si la matrice M possède une valeur propre racine d'ordre $p \geq 1$ de l'unité, la matrice ${}^tM^p$ laisse invariant un vecteur à coefficients entiers non nul. On peut supposer p minimal parmi les entiers ≥ 1 vérifiant cette propriété. Le caractère γ non trivial correspondant à ce vecteur ${}^tM^p$-invariant a une orbite finie sous l'action de τ_M : les caractères $\gamma, \gamma \circ \tau_M, \cdots, \gamma \circ \tau_M^{p-1}$ sont deux à deux distincts et on a $\gamma = \gamma \circ \tau_M^p$. La fonction $\phi = \gamma + \gamma \circ \tau_M + \cdots + \gamma \circ \tau_M^{p-1}$ est une fonction dans $L^2(\mathbb{T}^d)$ invariante par τ_M. Elle est non nulle, car c'est la somme de fonctions deux à deux orthogonales non nulles.

Supposons inversement que la matrice M ne possède pas de valeur propre racine de l'unité. L'orbite d'un caractère γ non nul sous l'action de τ_M est alors infinie. En d'autres termes, l'ensemble des images $(T_M^n\gamma, n \in \mathbb{Z}^d)$ de γ sont deux à deux distinctes.

La relation (12.3) est alors vérifiée pour chaque couple de caractères (γ_1, γ_2) non triviaux, puisque, pour n assez grand, $T_M^n\gamma_1$ est différent de γ_2, donc orthogonal à γ_2. Par linéarité et par densité des combinaisons linéaires finies de caractères (i.e. des polynômes trigonométriques), la relation (12.3) s'étend à tout $L^2(\lambda)$. Il est clair qu'elle implique l'ergodicité. $\square$

Exemple : la matrice 2×2

$$M = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

définit un automorphisme ergodique du tore $\mathbb{T}^2$.

Remarque 12.3.2

Nous disposons de deux types d'exemples aux propriétés très différentes :
- les rotations irrationnelles sur le cercle $x \to x + \alpha$ qui sont à spectre discret et uniquement ergodiques, et des transformations qui s'en déduisent par extension ;
- la transformation $x \to 2x$ mod 1, ou son modèle "symbolique", le schéma de Bernoulli, qui ont des propriétés stochastiques fortes (mélange, ...), ou des transformations associées aux matrices à coefficients entiers.

Pour ce deuxième type de transformations, on note l'abondance des mesures invariantes. En particulier, ces transformations ont beaucoup de points périodiques. A partir d'un orbite périodique, $\{x_0, \tau x_0, \cdots, \tau^{p-1}x_0, \tau_p x_0 = x_0\}$, on construit la mesure invariante discrète portée par l'orbite de x_0 :

$$\frac{1}{p}\sum_{j=0}^{p-1} \delta_{\tau^j x_0}.$$

Prenons $x \to 2x$ mod 1. L'orbite d'un point rationnel $x = \frac{k}{r}$, k et r premiers entre eux, est finie modulo 1. Il existe donc des entiers p et ℓ tels que $2^{p+\ell}x - 2^\ell x \in \mathbb{Z}$. Si r est impair, ceci implique que $2^p x = x$ mod 1. Ceci montre, en particulier, que toute fraction irréductible à dénominateur impair peut s'écrire sous la forme (en général non irréductible) $\frac{k}{2s+1} = \frac{k'}{2^n-1}$, pour des entiers k', n.

En exercice, on pourra étudier, de façon générale, les orbites des points rationnels du tore $\mathbb{T}^d$, sous l'action d'un matrice à coefficients entiers, opérant modulo 1.

• Transformation de Gauss (fraction continue)

Notations : On note $[y]$ la partie entière d'un réel y, $\{\frac{1}{y}\} = y - [y]$ sa partie fractionnaire.

Les systèmes de numération (développement en base b, développement en fraction continue qui fait l'objet de ce paragraphe), sont basés sur le principe suivant :

Soient τ une transformation mesurable d'un espace mesuré $(E, \mathcal{E})$ dans lui-même et a une fonction mesurable sur E à valeurs dans $\mathbb{N}$.

A tout point x, on peut associer la suite $\phi(x) = (a(\tau^n x))_{n \geq 0}$. On obtient ainsi une application $x \to \phi(x)$ de E dans l'espace $\Omega = \mathbb{N}^{\mathbb{N}}$, telle que $\phi \circ \tau = \theta \circ \phi$. Cette application est injective si la donnée de la suite $(a(\tau^n x))_{n \geq 0}$ détermine x. Elle fournit un **codage** des points de E, et "conjugue" la transformation τ au décalage θ restreint à un sous-ensemble θ-invariant de Ω.

Considérons la partition ζ de E dont les éléments sont les sous-ensembles $C_k = \{x : a(x) = k\}$. L'injectivité de l'application ϕ se traduit par la propriété que "l'itinéraire" de chaque point x de E, sous l'action de τ, à travers l'élement de la partition ζ détermine x.

Quand on prend pour E l'intervalle $[0, 1]$, on obtient les représentations en base b (dans ce cas, la transformation τ est la transformation $x \to bx \bmod 1$ et la fonction a est $a(x) = [bx]$).

Développement en fraction continue

Soit $(a_n)_{n \geq 1}$ une suite d'entiers ≥ 1. Définissons deux suites d'entiers (p_n) et (q_n), $n \geq -1$, avec les valeurs initiales $p_{-1} = 1$, $p_0 = 0$, $q_{-1} = 0$, $q_0 = 1$, par les relations de récurrence suivantes, pour $n \geq 0$:

$$\begin{aligned} p_{n+1} &= a_{n+1} p_n + p_{n-1}, && (12.4) \\ q_{n+1} &= a_{n+1} q_n + q_{n-1}. && (12.5) \end{aligned}$$

On montre par récurrence les relations

$$p_{n-1} q_n - p_n q_{n-1} = (-1)^n, \ \forall n \geq 0, \tag{12.6}$$

qui impliquent que les fractions $\frac{p_n}{q_n}$ sont irréductibles.

Pour tout $n \geq 1$, considérons les fractions rationnelles

$$\frac{1}{a_1 + \frac{1}{a_2 + \ddots + \frac{1}{a_n + y}}}.$$

On montre par récurrence qu'elles s'écrivent sous la forme

$$\frac{1}{a_1 + \frac{1}{a_2 + \ddots + \frac{1}{a_n + y}}} = \frac{p_n + p_{n-1} y}{q_n + q_{n-1} y}. \tag{12.7}$$

Les relations de récurrence vérifiées par les suites (p_n) et (q_n) montrent que ces suites croissent au moins aussi vite que la suite de Fibonacci (cas où tous les a_n sont égaux à 1), donc avec une vitesse au moins exponentielle. D'autre part, d'après (12.6), on a :

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n-1}}{q_n q_{n-1}}, \tag{12.8}$$

ce qui montre que la suite $(\frac{p_n}{q_n})$ vérifie la condition de Cauchy. Notons x sa limite :

$$x = \lim_n \frac{p_n}{q_n} = \lim_n \frac{1}{a_1 + \frac{1}{a_2 + \ddots + \frac{1}{a_n + \ddots}}}.$$

La suite $(0; a_1, a_2, \cdots)$ forme le développement en fraction continue de x.

Nous montrons comment on obtient le développement en fraction continue des points de l'intervalle $]0,1[$.

On prend pour E l'intervalle $[0,1]$ privé des rationnels. Notons τ la transformation (de Gauss) de $]0,1[$ dans $]0,1[$ définie par

$$\tau : x \to \{\frac{1}{x}\},$$

a la fonction $a(x) = [\frac{1}{x}]$.

Pour $x \in]0,1[$, on a donc :

$$x = \frac{1}{a(x) + \tau x}. \tag{12.9}$$

En itérant la relation (12.9), on obtient :

$$x = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \ddots + \cfrac{1}{a_n(x) + \tau^n x}}}, \ \forall n \geq 1, \tag{12.10}$$

où $(a_n(x))_{n\geq 1}$ est la suite d'entiers ≥ 1 définie par

$$a_n(x) = a(\tau^{n-1}x).$$

On peut étendre la construction aux points x rationnels, mais cette construction, qui correspond à l'algorithme du PGCD, s'arrête à la valeur de n pour laquelle $\tau^n(x) = 0$. Par contre, si x est irrationnel, la construction se poursuit indéfiniment et on peut ainsi écrire tout x irrationnel de $[0,1]$ sous la forme de son "développement" en fraction continue :

$$x = \lim_n \frac{p_n(x)}{q_n(x)} = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \ddots + \cfrac{1}{a_n(x) + \ddots}}}.$$

Les suites d'entiers $(p_n(x))$ et $(q_n(x))$ associées à x, vérifient, d'après (12.5), les relations de récurrence :

$$\begin{aligned} p_{n+1}(x) &= a_{n+1}(x)p_n(x) + p_{n-1}(x), \\ q_{n+1}(x) &= a_{n+1}(x)q_n(x) + q_{n-1}(x), \end{aligned}$$

avec les valeurs initiales : $p_{-1} = 1$, $p_0 = 0$, $q_{-1} = 0$, $q_0 = 1$.

En notant simplement $(a_n), (p_n), (q_n)$ les suites d'entiers associées à x, on obtient, par les relations (12.6) :

$$q_n x - p_n = q_n \frac{p_n + p_{n-1}\tau^n x}{q_n + q_{n-1}\tau^n x} - p_n = \frac{(-1)^n \tau^n x}{q_n + q_{n-1}\tau^n x}, \tag{12.11}$$

et

$$q_{n-1}x - p_{n-1} = q_{n-1}\frac{p_n + p_{n-1}\tau^n x}{q_n + q_{n-1}\tau^n x} - p_{n-1} = \frac{(-1)^{n+1}}{q_n + q_{n-1}\tau^n x}, \tag{12.12}$$

d'où :

$$\frac{q_n x - p_n}{q_{n-1}x - p_{n-1}} = -\tau^n x,$$

et donc, puisque $q_0 x - p_0 = x$,

$$q_n x - p_n = (-1)^n \Pi_{k=0}^n \tau^k x. \tag{12.13}$$

Notons, pour tout reél y, $((y)) = \inf_{p\in\mathbb{Z}} |y-p| = \inf(\{y\}, \{1-y\})$.

On a $((q_n x)) = |q_n x - p_n|$, avec alternance de signe de la quantité $q_n x - p_n$.

A l'ordre n, la relation (12.12) s'écrit :

$$q_n x - p_n = \frac{(-1)^n}{q_{n+1} + q_n \tau^{n+1} x}, \tag{12.14}$$

ce qui fournit l'encadrement :

$$\frac{1}{q_{n+1} + q_n} \leq ((q_n x)) \leq \frac{1}{q_{n+1}}. \tag{12.15}$$

En particulier, on a :

$$|x - \frac{p_n}{q_n}| \leq \frac{1}{q_n\ q_{n+1}} < \frac{1}{a_{n+1}\ q_n^2}. \tag{12.16}$$

Cette inégalité permet de relier la vitesse de convergence de la suite $(\frac{p_n}{q_n})_{n\in\mathbb{N}}$ vers x, à la suite $(a_n(x))$, et donc de décrire les propriétés "diophantiennes" de x (vitesse d'approximation par les rationnels).

Pour étudier les propriétés "métriques" du développement en fraction continue, nous utilisons les informations fournies par la théorie ergodique.

Une mesure "naturelle" invariante pour la transformation τ est une mesure ayant une densité par rapport à la mesure uniforme sur l'intervalle. Une telle mesure ν a été explicitée par Gauss. C'est la mesure ν de densité $\frac{1}{\ln 2}\frac{1}{1+x}$ par rapport à la mesure de Lebesgue.

Nous allons établir l'ergodicité de la transformation pour la mesure ν. Pour cela, calculons l'opérateur P adjoint, pour la mesure ν, de l'opérateur T de composition par τ. On a :

$$\begin{aligned}
\int_0^1 g(\{\frac{1}{x}\})\ f(x)\frac{1}{1+x}\ dx &= \sum_{k=1}^{\infty} \int_{\frac{1}{k+1}}^{\frac{1}{k}} g(\{\frac{1}{x}\})\ f(x)\frac{1}{1+x}\ dx \\
&= \sum_{k=1}^{\infty} \int_{\frac{1}{k+1}}^{\frac{1}{k}} g(\frac{1}{x} - k)\ f(x)\frac{1}{1+x}\ dx \\
&= \sum_{k=1}^{\infty} \int_0^1 g(y)\ f(\frac{1}{y+k})\frac{1+y}{(y+k)(1+y+k)}\ \frac{1}{1+y}\ dy,
\end{aligned}$$

d'où

$$Pf(x) = \sum_{k=1}^{\infty} f(\frac{1}{x+k})\frac{1+x}{(x+k)(1+x+k)}.$$

L'invariance de la mesure ν est équivalente à la propriété $P1 = 1$, qui résulte de la relation

$$\sum_{k=1}^{\infty} \frac{1}{(x+k)(1+x+k)} = \sum_{k=1}^{\infty}[\frac{1}{x+k} - \frac{1}{1+x+k}] = \frac{1}{x+1}.$$

Montrons que l'opérateur markovien P a de bonnes propriétés spectrales sur l'espace des fonctions lipschitziennes. Posons

$$m(f) = \sup_{x,y,x\neq y} |\frac{f(x)-f(y)}{x-y}|.$$

Lemme 12.3.3 *Il existe des constantes $\rho \in [0,1[$ et C, telles que*

$$m(Pf) \leq \rho m(f) + C\|f\|_\infty, \ \forall f \text{ lipschitzienne sur } [0,1]. \tag{12.17}$$

Preuve : On a :

$$\begin{aligned}
&|Pf(x) - Pf(y)| \\
&\leq \sum_k |f(\frac{1}{x+k}) - f(\frac{1}{y+k})| \frac{1+x}{(x+k)(x+k+1)} \\
&+ \sum_{k\geq 1} |f(\frac{1}{y+k})| |\frac{1+x}{(x+k)(x+k+1)} - \frac{1+y}{(y+k)(y+k+1)}|.
\end{aligned}$$

On obtient donc (12.17) par un calcul élémentaire utilisant le théorème des accroissements finis, avec

$$\begin{aligned}
C &\leq \sum_{k\geq 1} \frac{1}{k(k+1)} < +\infty \\
\rho &= \sup_{x,y} \sum_{k\geq 1} \frac{1}{k^2} \frac{1+x}{(x+k)(x+k+1)} < 1.
\end{aligned}$$

Pour obtenir la deuxième inégalité, on observe que le majorant de ρ est le barycentre des réels $\frac{1}{k^2}$, $k \geq 1$, par les poids $\frac{1+x}{(x+k)(x+k+1)}$, dont la somme vaut 1. $\square$

Par récurrence, on déduit de (12.17),

$$m(P^n f) \leq \rho^n m(f) + (1 + \rho + \rho^2 + ... + \rho^{n-1}) C\|f\|_\infty, \ \forall n \geq 1.$$

D'où :

$$m(P^n f) \leq \rho^n m(f) + \frac{C}{1-\rho}\|f\|_\infty, \forall n \geq 1,$$

ce qui implique l'équicontinuité de la famille $\{P^n f, n \geq 0\}$, pour toute fonction f lipschitzienne.

Par le théorème d'Ascoli, si f est lipschitzienne, ceci implique que la suite des moyennes $(\frac{1}{n}\sum_{k=0}^{n-1} P^k f)$ a des valeurs d'adhérence, pour la convergence uniforme, qui sont des fonctions continues P-invariantes.

On montre facilement que ces fonctions sont constantes.

Lemme 12.3.4 *Si h est une fonction continue telle que $Ph = h$, alors h est constante.*

Preuve : Les ensembles $\{y : h(y) = \max_x h(x)\}$, $\{y : h(y) = \min_x h(x)\}$ sont des fermés invariants par le semi-groupe engendré par les transformations $x \to \frac{1}{x+k}$. Donc ils contiennent tous les deux 0, ce qui implique que $\max_x h(x) = \min_x h(x)$. $\square$

Il y a donc convergence uniforme des moyennes $\frac{1}{n}\sum_0^{n-1} P^k f$ vers la constante $\nu(f)$, pour toute fonction f continue, et unicité de la mesure P-invariante ν : Si μ est une mesure de probabilité P-invariante, on a

$$\mu(f) = \int_0^1 \frac{1}{n}\sum_{k=0}^{n-1} P^k f \, d\mu = \lim_n \int_0^1 \frac{1}{n}\sum_{k=0}^{n-1} P^k f \, d\mu = \nu(f).$$

Proposition 12.3.5 *Si f et g sont dans $L^2(\nu)$, on a :*

$$\lim_n \frac{1}{n}\sum_0^{n-1}\int f\ T^k\overline{g}\ d\nu = \lim_n \frac{1}{n}\sum_0^{n-1}\int P^k f\ \overline{g}\ d\nu = \nu(f)\nu(\overline{g}).$$

Le système $(E, \mathcal{E}, \nu, \tau)$ est donc ergodique.

Preuve : La convergence est établie pour f et g continues. Comme T (et P) sont des contractions de $L^2(\nu)$, la propriété s'étend par densité à tout $L^2(\nu)$. $\square$

Notons que l'on peut en fait montrer la convergence de la suite des itérées $(P^n f)$, pour toute fonction f lipschitzienne, ce qui implique le mélange fort pour la transformation τ.

Corollaire 12.3.6 *(Paul Lévy) Pour presque tout x,*

$$\lim_n ((q_n(x)x))^{\frac{1}{n}} = \gamma,$$

où

$$\gamma = \frac{1}{\ln 2}\int_0^1 \frac{\ln x}{1+x}\ dx = \frac{1}{\ln 2}\sum_{n\geq 1}\frac{(-1)^n}{n^2} = -\frac{1}{\ln 2}\frac{\pi^2}{12}.$$

Preuve : D'après la relation (12.13), il suffit d'appliquer le théorème ergodique à la fonction $\phi(x) = \ln x$.

$\square$

12.4 Mesures invariantes et opérateur de transfert

Après ces exemples, nous allons donner, dans un cadre général, une méthode de construction de mesures invariantes pour la classe des transformations possèdant des inverses locales contractantes. Cette méthode est basée sur l'étude asymptotique des itérés de l'opérateur de transfert.

Si P est un opérateur markovien sur $\mathcal{C}(E)$, inverse à gauche de l'opérateur T associé à τ, c'est-à-dire vérifie

$$PTf = f,\ \forall f \in \mathcal{C}(E),$$

alors toute mesure ν invariante par P est invariante par τ, d'après la relation :

$$\int f\nu(dx) = \int PTf\nu(dx) = \int Tf\nu(dx).$$

L'étude des mesures laissées invariantes par un tel opérateur P permet donc de construire des mesures invariantes pour le système dynamique (topologique) (E, τ).

Soit maintenant un opérateur Q positif opérant sur l'espace $\mathcal{C}(E)$ tel que :

$$Q(gTf) = Q(g)\,f,\ \forall f, g \in \mathcal{C}(E). \tag{12.18}$$

Alors, pour toute fonction $u \in \mathcal{C}(E)$, telle que $Q(u) = 1$, on obtient un opérateur markovien inverse à gauche de T en posant $P_u(f) := Q(uf)$ et cet opérateur vérifie encore la relation (12.18) :

$$P_u(gTf) = P_u(g)\,f,\ \forall f, g \in \mathcal{C}(E).$$

Relation avec les opérateurs adjoints

Si ν est une mesure invariante par P et si P vérifie (12.18), alors P est ν-adjoint de T :

$$\int f\ T\overline{g}\ \nu(dx) = \int P(f\ T\overline{g})\ \nu(dx) = \int P(f)\ \overline{g}\ \nu(dx).$$

Inversement, l'opérateur ν-adjoint de T au sens d'une mesure ν invariante par τ vérifie (12.18) sur les espaces $L^p(\nu)$ et donc aussi cette relation (12.18) sur l'espace $\mathcal{C}(E)$, pourvu qu'il applique $\mathcal{C}(E)$ dans lui-même.

• Inversion de τ par un opérateur

La discussion précédente pose le problème général de la construction d'un inverse à gauche pour T. Dans de nombreux exemples, on peut effectivement construire un opérateur P ayant les propriétés précédentes en termes d'it inverses locaux de τ.

Supposons qu'il existe une partition de E en un nombre fini de sous-ensembles mesurables $E_s, s \in I$, tels que la restriction de τ à E_s possède une application inverse, notée $s : x \to sx$, de E sur E_s. Si ϕ est une fonction positive sur E, on considère Q défini par

$$Q_\phi f(x) = \sum_{s \in I} \phi(sx)\ f(sx) = \sum_{y:\tau y = x} \phi(y) f(y).$$

Si τ possède localement un inverse continu, on obtient ainsi un opérateur positif sur $\mathcal{C}(E)$.

Dans le cas où l'ensemble des inverses est fini, on peut prendre ϕ identiquement égal à 1, et pour toute fonction u continue ≥ 0 sur E, telle que $\sum_s u(sx) = 1$, construire l'opérateur P_u :

$$P_u f(x) = \sum_{y:\tau y = x} u(y) f(y) = \sum_s u(sx) f(sx).$$

On a bien

$$P_u(fTg)(x) = \sum_s u(sx) f(sx) g(\tau(sx)) = g(x) P_u f(x),$$

et la relation (12.18) est vérifiée.

Un procédé pour passer d'un opérateur Q non markovien à un opérateur P markovien, est celui de la *relativisation* : si v est une fonction propre strictement positive pour Q, de valeur propre $\lambda > 0$, $Qv = \lambda v$, on définit l'opérateur P par :

$$Pf(x) = \frac{1}{\lambda v(x)} Q(fv)(x).$$

Il est clair que P est un opérateur markovien.

Remarque : La construction des opérateurs Q et P s'étend au cas d'une famille dénombrable d'inverses (on peut en particulier englober le cas de la transformation du développement en fraction continue).

Exemple : $\tau : x \to 2x \bmod 1$

Considérons la transformation $\tau : x \to 2x \bmod 1$. Soit P l'opérateur défini sur $\mathcal{C}(\mathbb{R}/\mathbb{Z})$ par

$$Pf(x) = \frac{1}{2}[f(\frac{x}{2}) + f(\frac{x}{2} + \frac{1}{2})].$$

Nous avons vu que cet opérateur est l'opérateur λ-adjoint de l'opérateur T de composition par τ, λ étant la mesure de Lebesgue sur l'espace $\mathbb{R}/\mathbb{Z}$.

En suivant la méthode générale que nous venons d'esquisser, nous pouvons définir d'autres opérateurs markoviens sur le cercle associés à la transformation τ. Notons Q l'opérateur (non markovien)

$$Qf(x) = f(\frac{x}{2}) + f(\frac{x}{2} + \frac{1}{2}).$$

Soit maintenant une fonction u, continue et positive, telle que $Qu = 1$, c'est-à-dire, $u(x)+u(x+\frac{1}{2}) = 1$. On définit un opérateur P_u par $P_u f = Q(uf)$, soit plus explicitement :

$$P_u f(x) = u(\frac{x}{2})f(\frac{x}{2}) + u(\frac{x}{2} + \frac{1}{2})f(\frac{x}{2} + \frac{1}{2}).$$

L'opérateur P_u vérifie la condition (12.18), ce qui implique que toute mesure invariante par P_u est invariante par τ.

Fractions continues : l'opérateur utilisé plus haut est du type P_u, avec

$$u(x) = x^2 \frac{1 + \{\frac{1}{x}\}}{1 + x}.$$

Le facteur $\frac{1+\{\frac{1}{x}\}}{1+x}$ provient de la *relativisation* de l'opérateur Q adjoint pour la mesure de Lebesgue, qui est de la forme

$$Qf(x) = \sum_{k=1}^{\infty} f(\frac{1}{x+k}) \frac{1}{(x+k)^2}.$$

associé à la fonction $\phi(x) = x^2$. La fonction $v(x) = \frac{1}{x+1}$ est Q-invariante.

On peut également considérer la famille des opérateurs de la forme

$$Q_\alpha f(x) = \sum_{k=1}^{\infty} f(\frac{1}{x+k}) \frac{1}{(x+k)^\alpha}$$

associé à la fonction $\phi_\alpha(x) = x^\alpha$, pour $\alpha > 1$.

Exemple : Opérateur de transfert pour le shift

Si ω est un point de $\Omega^+ = I^{\mathbb{N}}$ et $a \in I$ un symbole, on note $a\omega$ le mot obtenu en plaçant a devant ω. Par la construction décrite précédemment, on obtient sur l'espace $\mathcal{C}(\Omega^+)$, **l'opérateur de transfert** $P_u f$ associé à une fonction continue $u > 0$. Il est donc défini par

$$P_u f(\omega) = \sum_{\omega' : \theta\omega' = \omega} u(\omega') f(\omega') = \sum_{a \in I} u(a\omega) f(a\omega).$$

Cette construction s'étend aux "sous-décalages" (sub-shift) de type fini.

• Mesures quasi-invariante et opérateur adjoint

Supposons maintenant que la transformation τ opère sur l'espace mesuré $(E, \mathcal{E}, \mu)$, en laissant la mesure μ quasi-invariante ($\tau\mu \ll \mu$), mais non nécessairement invariante. Un problème est alors la recherche d'une mesure invariante ν équivalente à μ.

Notons que, dans cette situation, nous avons une notion d'*ergodicité* de la classe de μ, sous l'action de τ (si A est un ensemble mesurable tel que $\tau^{-1}A \subset A$, alors A, ou son complémentaire, est μ-négligeable).

On peut opérer par dualité sur les mesures appartenant à la classe d'équivalence de μ. Si ν est une mesure de probabilité équivalente à μ, son image est encore une mesure de probabilité équivalente. On peut peut définir un opérateur P opérant sur les densités de ces mesures. Pour $\nu = f\mu$, $f \in L^1(\mu)$, on pose :

$$Pf = \frac{d\tau\nu}{d\mu} = \frac{d\tau(f\mu)}{d\mu}.$$

On a donc :

$$\int_E Pf\ \overline{\varphi}\ \mu(dx) = \int_E f\ T\overline{\varphi}\ \mu(dx), \forall f \in L^1(\mu), \forall \varphi \in L^\infty(\mu).$$

Il est clair que P est un opérateur positif tel que $\int_E Pf\ \mu(dx) = \int_E f\ \mu(dx)$. C'est une contraction positive de $L^1(\mu)$, Dans le cas où la mesure μ est invariante par τ, on a $P1 = 1$ et l'opérateur P est un opérateur **markovien** sur $L^1(\mu)$.

Dans le cas général où μ est seulement quasi-invariante, l'existence d'une mesure ν invariante par τ dans la classe de μ équivaut à l'existence d'une fonction ψ positive dans $L^1(\mu)$ (la densité $\frac{d\nu}{d\mu}$) telle que $P\psi = \psi$. En effet, si une solution $\psi \geq 0$ existe, alors la mesure ν définie par $d\nu = \psi\, d\mu$ est la mesure invariante cherchée.

Expression de l'opérateur adjoint

Supposons qu'il existe une famille S finie ou dénombrable de transformations et une décomposition de E en ensembles E_s, $s \in S$, telles que la restriction de τ à E_s soit une bijection de E_s sur E, d'inverse notée $x \to sx$. On a :

$$\begin{aligned}
\int_E f\,T\phi\ \mu(dx) &= \sum_s \int_{E_s} f(x)\phi(\tau x)\ \mu(dx) \\
&= \sum_s \int_E f(sx)\phi(x)\,\frac{d\,s^{-1}\mu}{d\mu}(x)\ \mu(dx) \\
&= \int_E Pf(x)\,\phi(x)\ \mu(dx),
\end{aligned}$$

avec

$$Pf(x) = \sum_s f(sx)\frac{d\,s^{-1}\mu}{d\mu}(x).$$

En définissant u sur E_s par $u(sx) = \frac{d\,s^{-1}\mu}{d\mu}(x)$, on obtient un opérateur adjoint de la forme $P = P_u$.

Les densités θ des mesures de probabilité τ-invariantes, absolument continues par rapport à μ, sont les solutions θ dans $L^1(\mu)$ de l'équation $P\theta = \theta$.

Par exemple, si nous prenons pour transformation τ la transformation $x \to \{\frac{1}{x}\}$ des fractions continues, les inverses locaux sont les applications $x \to \frac{1}{x+k}$ et l'ensemble S est indexé par $\mathbb{N}$. L'opérateur adjoint, par rapport à la mesure de Lebesgue sur $[0,1]$, est l'opérateur P défini par :

$$Pf(x) = \sum_{k\in\mathbb{N}} f(\frac{1}{x+k})\frac{1}{(x+k)^2}.$$

Ici, les transformations $s \in S$ sont les transformations $x \to \frac{1}{x+k}$, indexées par $k \in \mathbb{N}$, appliquant $]0,1[$ sur $]\frac{1}{k+1}, \frac{1}{k}[$, d'inverse $x \to \frac{1}{x} - k$, de $]\frac{1}{k+1}, \frac{1}{k}[$ sur $]0,1[$ et par changement de variable nous avons : $\displaystyle\int_{\frac{1}{k+1}}^{\frac{1}{k}} f(\frac{1}{x} - k)\ dx = \int_0^1 f(y)\frac{1}{(y+k)^2}\ dy.$

La fonction u est la fonction $u(y) = y^2$ et la fonction invariante donnant la densité de la mesure τ-invariante est $y \to \frac{1}{y+1}$.

12.5 Noyau formé d'applications contractantes

Nous considérons maintenant le cas d'un "noyau de transition" non nécessairement normalisé, construit à partir d'une famille d'applications s, et nous montrons comment on peut se ramener au cas markovien quand les transformations s sont contractantes.

Supposons donné un espace E compact et une famille S finie d'applications continues $s : x \to sx$ de E dans lui-même.

Considérons une fonction $\phi > 0$ et continue sur E. Formons l'opérateur Q défini par

$$Qf(x) = \sum_{s \in S} \phi(sx) f(sx).$$

On peut plus généralement considérer une famille $(x \to \phi(s,x), s \in S)$ de fonctions continues $\phi(s,.)$ strictement positives définies sur E et former l'opérateur

$$Qf(x) = \sum_{s \in S} \phi(s,x) f(sx).$$

On ne suppose pas réalisée la condition de normalisation $\sum_{s \in S} \phi(sx) = 1$.

L'opérateur Q envoie $\mathcal{C}(E)$ dans lui même, conserve le cône des fonctions positives, ainsi que le cône $\mathcal{M}^+$ des mesures de Radon positives sur lequel il opère par dualité. Pour construire à partir de Q un opérateur markovien, nous allons montrer, sous certaines conditions, l'existence d'une fonction propre strictement positive pour l'opérateur Q^* dual de Q dans l'espace des mesures de Radon sur E. Notons d'abord un résultat général sur les noyaux de transition opérant sur les fonctions continues.

Lemme 12.5.1 *Si E est un espace métrique compact et Q un opérateur positif sur l'espace des fonctions continues, il existe une mesure de probabilité sur E et un scalaire $\lambda > 0$ tel que $Q^*\mu = \lambda\mu$.*

Preuve : Le cône $\mathcal{M}^+$ des mesures de Radon positives sur E est à base compacte et Q^* est opérateur continu sur ce cône. D'après le théorème de Schauder-Tychonov, il existe une mesure μ non nulle, que l'on peut normaliser, et un scalaire $\lambda > 0$ tels que $Q^*\mu = \lambda\mu$. □

Nous allons voir qu'en faisant des hypothèses de contraction sur les applications s de S et de régularité sur $(\phi(s,.)$, on peut étudier le comportement de la suite (Q^n), grâce à l'existence d'un fonction propre pour Q. Pour cela, nous allons donner une condition, reliant une propriété de contraction des transformations s et la régularité des poids $\phi(s,x)$, qui assure l'existence d'un élément propre $v > 0$ dans $\mathcal{C}^+(E)$. Nous nous ramènerons ensuite par relativisation à un noyau markovien P. Sous nos hypothèses, nous pourrons montrer la convergence de la suite des itérés $(P^n f)$ pour f continue, ce qui impliquera en particulier l'unicité de la mesure P-invariante.

Notations - Hypothèses 12.5.2 Pour g continue sur E et $\varepsilon > 0$, posons

$$\mathcal{V}(g,\varepsilon) = \sup_{x,y \in E, d(x,y) < \varepsilon} |g(x) - g(y)|.$$

Soient $(\eta_k)_{k \geq 0}$ et $(\varepsilon_k)_{k \geq 0}$ deux suites de réels strictement positifs, tendant vers 0.

Définissons le cône $\mathcal{F} = \mathcal{F}((\eta_k),(\varepsilon_k))$ par

$$\mathcal{F}((\eta_k),(\varepsilon_k)) = \{g \in \mathcal{C}^+(E) : \forall k,\ d(x,y) \leq \eta_k \Rightarrow \frac{g(x)}{g(y)} \leq e^{\varepsilon_k}\}.$$

Soit $x_1,\dots,x_r$ des points de E tels que les boules $B(x_i,\eta_0)$ recouvrent E. Ce cône $\mathcal{F}$ est à base compacte, d'après le théorème d'Ascoli. En effet, la section $\mathcal{F}\cap\{g : \sum_{i=1}^r g(x_i) = 1\}$ est un ensemble équicontinu et borné de fonctions continues.

Supposons qu'on ait la propriété de contraction suivante :

$$d(x,y) \leq \eta_k \Rightarrow d(sx,sy) \leq \eta_{k+1}. \tag{12.19}$$

Supposons par ailleurs que les fonctions $(\phi(s,.)$ vérifient la condition de régularité suivante :

$$\sum_{k\geq 0} \sup_{s\in S} \mathcal{V}(\mathrm{Log} u_s, \eta_k) < +\infty. \tag{12.20}$$

Lemme 12.5.3 *Si (ε_k) est définie par $\varepsilon_k = \sum_{j=k}^{\infty} \sup_{s\in S} \mathcal{V}(\mathrm{Log} u_s, \eta_k) < +\infty$, le cône $\mathcal{F}((\eta_k),(\varepsilon_k))$ est laissé invariant par l'opérateur Q.*

Preuve : Le lemme résulte de la majoration suivante. Pour toute fonction $g > 0$, on a :

$$d(x,y) \leq \eta_k \Rightarrow f(sx) \leq \exp(\mathcal{V}(\mathrm{Log} f), \eta_{k+1}) f(sy).$$

Ceci implique :

$$\begin{aligned} Qf(x) &= \sum_s \phi(s,x) f(sx) \\ &\leq \exp(\mathcal{V}(\mathrm{Log} f, \eta_{k+1}) + \sup_s \mathcal{V}(\mathrm{Log}\phi(s,.), \eta_k)) \sum_s \phi(s,y) f(sy) \\ &= \exp(\mathcal{V}(\mathrm{Log} f, \eta_{k+1}) + \sup_s \mathcal{V}(\mathrm{Log}\phi(s,.), \eta_k))\, Qf(y)\ \leq e^{\varepsilon_k}\, Qf(y). \end{aligned}$$

Si f est dans le cône $\mathcal{F}((\eta_k),(\varepsilon_k))$, Qf y est aussi, puisqu'on a la majoration :

$$\exp(\mathcal{V}(\mathrm{Log} f, \eta_{k+1}) + \sup_s \mathcal{V}(\mathrm{Log}\phi(s,.), \eta_k)) \leq \varepsilon_{k+1} + \sup_s \mathcal{V}(\mathrm{Log}\phi(s,.), \eta_k))\varepsilon_k.$$

□

Proposition 12.5.4 *Si, pour une suite strictement positive (η_k) tendant vers 0, on a : $d(x,y) < \eta_k \Rightarrow d(sx,sy) < \eta_{k+1}$, et $\sum_k \sup_s \mathcal{V}(\mathrm{Log}\phi(s,.), \eta_k) < \infty$, alors il existe dans $\mathcal{C}^+(E)$ une fonction propre v de valeur propre $\lambda > 0$: $Qv = \lambda v$.*

Preuve : D'après le lemme, les conditions de la proposition assurent qu'il existe dans $\mathcal{C}^+(E)$ un cône $\mathcal{F}$ invariant par Q à base compacte. D'après le théorème de Schauder-Tychonov, Q laisse invariante une demi-droite de ce cône.

On en déduit l'existence d'une fonction propre continue positive v, de valeur propre $\lambda > 0 : Qv = \lambda v$.
□

La proposition 12.5.4 s'applique au cas où les applications $x \to sx$ sont contractantes dans le sens suivant : il existe $c < 1$ et $\eta_0 > 0$ tels que

$$d(x,y) < \eta_0 \Rightarrow d(sx,sy) < cd(x,y). \tag{12.21}$$

Définition 12.5.5 L'action de S sur E est dite **proximale**, si les fermetures des orbites de deux points quelconques sous l'action du semi-groupe engendré par S ont un point commun.

Sous l'hypothèse de proximalité, on vérifie facilement que les seules fonctions continues P-invariantes sont les constantes (considérer, comme nous l'avons fait dans l'exemple des fractions continues, les compacts formés des points où la fonction atteint son maximum et son minimum).

Dans la situation envisagée ici, la proximalité résulte de la propriété de contraction des applications $x \to sx$. Par contre, dans le cas où les "poids" $\phi(sx)$ peuvent s'annuler (i.e. si ϕ n'est pas > 0 en tout point), l'existence de "transitions" $x \to sx$ interdites, peut être un obstacle à la proximalité. Nous ne traiterons pas ce cas plus général ici.

En définissant l'opérateur P par :

$$Pf(x) = \frac{1}{\lambda v(x)} Q(fv)(x),$$

on obtient un opérateur markovien. On a, pour tout entier $n \geq 1$:

$$P^n f(x) = \frac{1}{\lambda^n v(x)} Q^n(fv)(x),$$

Pour comparer la régularité d'une fonction > 0 et celle de son logarithme, nous pouvons utiliser le lemme élémentaire suivant :

Lemme 12.5.6 *Pour tout réel $a > 0$ et tout réel b vérifiant pour un réel non nul M, $M^{-1} \leq y \leq M$, on a les implications :*

$$\begin{aligned} |a-b| &\leq \frac{1}{2M} \Rightarrow |\mathrm{Log} a - \mathrm{Log} b| \leq 2M|a-b|, \\ |\mathrm{Log} a - \mathrm{Log} b| &\leq 1 \Rightarrow |a-b| \leq eM|\mathrm{Log} a - \mathrm{Log} b|. \end{aligned}$$

Preuve : Pour a et b strictement positifs, on a :

$$\begin{aligned} |1-\frac{a}{b}| &\leq \frac{1}{2} \Rightarrow |\mathrm{Log} a - \mathrm{Log} b| \leq 2|1-\frac{a}{b}|, \\ |\mathrm{Log} a - \mathrm{Log} b| &\leq 1 \Rightarrow |1-\frac{a}{b}| \leq eM|\mathrm{Log} a - \mathrm{Log} b|. \end{aligned}$$

□

On en déduit :

Lemme 12.5.7 *Soit f une fonction vérifiant $M^{-1} \leq f(x) \leq M$, $\forall x \in E$. On a, pour tout $\alpha > 0$, les implications :*

$$\begin{aligned} \mathcal{V}(f,\alpha) &\leq \frac{1}{2M} \Rightarrow \mathcal{V}(\mathrm{Log} f,\alpha) \leq 2M\mathcal{V}(f,\alpha), \\ \mathcal{V}(\mathrm{Log} f,\alpha) &\leq 1 \Rightarrow \mathcal{V}(f,\alpha) \leq eM\mathcal{V}(\mathrm{Log} f,\alpha). \end{aligned}$$

Par relativisation, on se ramène au cas markovien. La fonction v fournie par la méthode précédente a une régularité fixée par la vitesse de convergence vers 0 du reste de la série donnant la régularité de ϕ. Il est clair que, sous l'hypothèse de contraction (C_1) et si ϕ est höldérienne, la fonction v est également höldérienne.

On peut montrer, en utilisant les résultats de la suite du chapitre, que l'opérateur P relativisé, construit à partir de Q et de la fonction propre v, a la propriété que pour toute fonction continue f la famille $\{P^n f, n \geq 0\}$ est équicontinue et bornée.

Proposition 12.5.8 *Sous les hypothèses de la proposition (12.5.4), pour toute fonction continue f sur E, la suite $(P^n f)$ converge uniformément vers la constante $\nu(f)$, où ν est une mesure de probabilité P-invariante sur E. La valeur propre λ est la plus grande valeur propre de l'opérateur Q opérant dans le cône des mesures bornées sur E et μ est l'unique mesure de probabilité vérifiant l'équation $Q\mu = \lambda\mu$.*

Remarque 12.5.9

En prenant v normalisée, $\mu(v) = 1$, la mesure de probabilité P-invariante ν est la mesure de densité v par rapport à μ : $d\nu = v\ d\mu$, puisque l'on a :

$$\begin{aligned}\int Pf\ d\nu &= \frac{1}{\lambda}\int \frac{1}{v}Q(fv)\ d\nu = \frac{1}{\lambda}\int Q(fv)\ d\mu \\ &= \frac{1}{\lambda}\int fv\ dQ^*\mu = \int f\frac{1}{v}\ d\mu = \int f\ d\nu.\end{aligned}$$

On a :

$$\lim_n \lambda^{-n}Q^n(f) = v\ \lim_n P^n(f/v) = v\nu(f/v) = v\mu(f).$$

Exemple du décalage et des fonctions höldériennes

Nous nous plaçons maintenant dans le cas particulier du décalage sur un espace produit. On pourrait restreindre le décalage à certains sous-ensembles invariants par θ (sous-shift de type fini transitifs).

Prenons comme espace compact E l'espace produit (unilatère) $\Omega^+ = \{1, ..., d\}^{\mathbb{N}}$ sur un alphabet fini $A = \{1, ..., d\}$ et pour transformation τ le décalage noté θ.

Les coordonnées d'un point x de Ω^+ sont notées $x_i, i \in \mathbb{N}$. Etant donné un point $x \in \Omega^+$ et un entier m, on note $C_m(x)$ le "cylindre" d'ordre m déterminé par x, c'est-à-dire l'ensemble des points $y \in \Omega^+$ tels que $y_i = x_i, i = 0, ..., m-1$.

On définit sur Ω^+ une métrique d par $d(x, y) = 2^{-N}$, où N est le plus grand entier ≥ 0 tel que $x_i = y_i, \forall\ 0 \leq i < N$.

Soit f une fonction définie sur Ω^+ à valeurs réelles. On définit

$$\mathcal{V}(f, 2^{-k}) = \sup\{|f(x) - f(y)|, x, y \in \Omega^+\ : x_i = y_i, 0 \leq i \leq k\}.$$

Pour $\alpha \in]0, 1[$, on considère l'espace $\mathcal{H}^\alpha(X)$ des fonctions f (höldériennes) sur X, telles que $\mathcal{V}(f, 2^{-k}) \leq c\alpha^k, \forall k \in \mathbb{N}$, pour une constante c.

Opérateur de transfert associé à f

Si x est un point de Ω^+ et a un symbole, on note ax le mot obtenu en plaçant a devant x. Sur l'espace $\mathcal{C}(\Omega^+)$, l'opérateur de transfert (appelé aussi opérateur de Ruelle-Perron-Frobenius) R_f associé à f est défini par

$$R_f\phi(x) = \sum_{y\ :Ty=x} e^{f(y)}\phi(y) = \sum_{a\in A} e^{f(ax)}\phi(ax).$$

Les applications $x \to ax$, $a \in A$, vérifient la condition de contraction :

$$d(ax, ay) \leq \frac{1}{2}d(x, y).$$

12.6 Itérés d'un opérateur markovien, série $\sum P^n f$

Soient E un espace compact et une famille S finie d'applications continues $s : x \to sx$ de E dans lui-même. Partons d'un opérateur Q de la forme

$$Qf(x) = \sum_{s \in S} \phi(s,x) f(sx)$$

non normalisé, nous avons vu comment, sous des hypothèses convenables, on peut se ramener au cas markovien. Nous supposons maintenant donné un opérateur **markovien**, que nous noterons P, de la forme

$$Pf(x) = \sum_{s \in S} u(sx) f(sx),$$

où u est une fonction continue strictement positive, vérifiant $\sum_s u(sx) = 1$, $\forall x \in E$.

Dans la suite, on supposera que S est fini. On note $C = Card(S)$.

On fixe sur E une mesure de probabilité P-invariante ν (les hypothèses sur S et la condition imposée plus loin à u impliquent en fait l'unicité de ν).

Notations - Hypothèses 12.6.1

Pour tout $\alpha > 0$ et toute fonction f sur E, on note comme en 12.5.2 :

$$\mathcal{V}(f, \alpha) = \sup_{x,y : d(x,y) < \alpha} |f(x) - f(y)|.$$

On considère comme précédemment une suite décroissante $(\eta_k)_{k \geq 0}$ telle que $\lim_k \downarrow \eta_k = 0$. On convient de prendre $\eta_0 = diam(E)$. Pour une fonction f sur E, on définit, pour $p \geq 0$:

$$m_p(f) = \sum_{k \geq p} \mathcal{V}(f, \eta_k),$$

$$m(f) = m_0(f) = \sum_{k \geq 0} \mathcal{V}(f, \eta_k).$$

On note $M(f)$ la "série des restes" :

$$M(f) = \sum_{k \geq 0} m_k(f).$$

D'autre part, on fait sur l'action de S (cf. 12.5.2), l'hypothèse de contraction suivante :

$$d(x,y) \leq \eta_k \Rightarrow d(sx, sy) \leq \eta_{k+1}, \tag{12.22}$$

et sur u l'hypothèse que la série des restes des variations converge, i.e. :

$$M(u) = \sum_{k \geq 0} m_k(u) < +\infty.$$

Lemme 12.6.2 *Sous l'hypothèse de contraction (12.22), on a les majorations :*

$$\begin{aligned}
\mathcal{V}(P^n f, \eta_k) &\leq \mathcal{V}(f, \eta_{k+n}) + C \sum_{j=1}^{n} \mathcal{V}(u, \eta_{k+j}) \|P^{n-j} f\|_\infty, \ \forall n, k \geq 0 &(12.23)\\
m_p(P^n f) &\leq m_{p+n}(f) + C \sum_{j=1}^{n} m_{p+j}(u) \|P^{n-j} f\|_\infty, \ \forall n, p \geq 0. &(12.24)
\end{aligned}$$

<u>Preuve</u> : Compte-tenu de la normalisation ($\sum_s u(sx) = 1$) et de la condition de contraction, on a, pour toute fonction g sur E et pour tout couple (x, y) de points de E vérifiant $d(x, y) \leq \eta_k$:

$$\begin{aligned}
|Pg(x) - Pg(y)| &\leq \sum_s |u(sx)g(sx) - u(sy)g(sy)| \\
&\leq \sum_s u(sx)|g(sx) - g(sy)| + \sum_s |u(sx) - u(sy)||g(sy)| \\
&\leq \mathcal{V}(g, \eta_{k+1}) + C\mathcal{V}(u, \eta_{k+1})\|g\|_\infty;
\end{aligned}$$

d'où :

$$\mathcal{V}(Pg, \eta_k) \leq \mathcal{V}(g, \eta_{k+1}) + C\mathcal{V}(u, \eta_{k+1})\|g\|_\infty.$$

En appliquant cette inégalité à $P^n f$ et en itérant, on obtient l'inégalité (12.23) :

$$\mathcal{V}(P^n f, \eta_k) \leq \mathcal{V}(f, \eta_{k+n}) + C \sum_{j=1}^{n} \mathcal{V}(u, \eta_{k+j}) \|P^{n-j} f\|_\infty, \ \forall n, k \geq 0.$$

L'inégalité (12.24) s'en déduit par sommation en k.

□

On en déduit immédiatement, par sommation, le corollaire suivant :

Corollaire 12.6.3 *On a, pour toute fonction f sur E,*

$$\mathcal{V}(P^n f, \eta_k) \leq \mathcal{V}(f, \eta_k) + C m_{k+1}(u) \|f\|_\infty, \ \forall n, k \geq 0, \tag{12.25}$$

$$m(P^n f) \leq m_n(f) + CM(u)\|f\|_\infty \leq m(f) + CM(u)\|f\|_\infty \ \forall n \geq 0. \tag{12.26}$$

Proposition 12.6.4 *Pour toute fonction f continue, sous la condition $m(u) < +\infty$, la famille $\{P^n f, n \geq 0\}$ est équicontinue et la suite des itérées $(P^n f)$ converge uniformément vers $\nu(f)$, où ν est l'unique mesure P-invariante sur E.*

<u>Preuve</u> : L'équicontinuité résulte de (12.25). Soit ν une mesure de probabilité P-invariante. Soit f une fonction continue telle que $\nu(f) = 0$. Pour obtenir la proposition, nous montrons que toute valeur d'adhérence h, pour la convergence uniforme, de la suite $(P^n f)_{n \geq 0}$ est nulle.

Le procédé diagonal, nous permet d'obtenir une suite d'entiers $(\varphi(n))_{n \geq 0}$ telle que, pour tout entier naturel k, la suite $P^{\varphi(n)-k} f$ converge uniformément vers une fonction h_k, avec $h_0 = h$.

Nous avons alors $Ph_k = h_{k-1}$, $\forall k \geq 1$. La suite $(\min(h_k))_{k \geq 0}$ des minima des fonctions h_k est donc décroissante. Elle est minorée. Soit α sa limite.

Montrons que toute valeur d'adhérence de la suite (h_k) est nulle. Soit h_∞ la limite d'une sous-suite uniformément convergente extraite de la suite (h_k). Pour tout entier $\ell \geq 0$, les fonctions $P^\ell h_\infty$ ont toutes le même minimum α. On obtient de même qu'elles ont le même maximum β.

Il existe donc des points x_ℓ et y_ℓ tels que $\forall s_1 \cdots s_\ell \in S^\ell$, $\forall \ell \geq 0$,

$$h(s_1 \cdots s_\ell x_\ell) = \alpha, \; h(s_1 \cdots s_\ell y_\ell) = \beta.$$

Comme $d(s_1 \cdots s_\ell x_\ell, s_1 \cdots s_\ell y_\ell) \leq \eta_\ell(x)$, on en déduit que $\alpha = \beta$. La fonction h_∞ est constante, donc nulle, car $\nu(h_\infty) = 0$.

Pour tout $\varepsilon > 0$, il existe donc k tel que $\|h_k\|_\infty \leq \varepsilon$. D'où : $\|h\|_\infty = \|P^k h_k\|_\infty \leq \varepsilon$. Ceci implique que h est identiquement nulle.

De la convergence uniforme de $(P^n f)$ vers $\nu(f)$, pour toute fonction continue f, on déduit qu'il existe une seule mesure de probabilité P-invariante. $\square$

Considérons le sous-espace V_0 de l'espace des fonctions continues défini par :

$$V_0 = \{f : \nu(f) = 0 \text{ et } m(f) < +\infty\}.$$

Ce sous-espace est stable par P compte-tenu de l'invariance de ν par P et de (12.26).

Pour $f \in V_0$, on a :

$$\begin{aligned} |f(x)| &= |\int_E (f(x) - f(y)) \, d\nu(y)| \\ &\leq \int_E |f(x) - f(y)| \, d\nu(y) \leq \mathcal{V}(f, \eta_0) \leq m(f); \end{aligned}$$

d'où :

$$\|f\|_\infty \leq m(f).$$

L'application $f \to m(f)$ définit donc une norme sur V_0, majorant la norme uniforme.

Tout ensemble borné de $(V_0, m(.))$ est relativement compact pour la norme uniforme. Il est en effet borné en norme uniforme et forme une famille équicontinue de fonctions, compte-tenu de l'inégalité :

$$\mathcal{V}(f, \eta_n) \leq \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{V}(f, \eta_k) \leq \frac{m(f)}{n}.$$

Sous la condition $M(u) < +\infty$, la famille $\{P^n f, m(f) \leq 1, \, n \geq 0\}$, pour $f \in V_0$, est équicontinue.

Proposition 12.6.5

$$\lim_n \sup_{\{g \in V_0 : m(g) \leq 1\}} \|P^n g\|_\infty = 0.$$

Preuve : Soit $B_1 = \{g \in V_0 : m(g) \leq 1\}$. D'après l'inégalité (12.26), nous avons

$$\sup_{\{g \in B_1, n \geq 0\}} m(P^n g) \leq 1 + CM(u),$$

et la famille $\{P^n g, g \in B_1, n \geq 0\}$ est équicontinue et bornée.

Pour prouver la proposition, il suffit de montrer que, pour toute suite (n_j) strictement croissante d'entiers et toute suite (f_j) de fonctions dans B_1 telles que la suite $(P^{n_j} f_j)_{j \geq 1}$ converge uniformément vers une limite h, on a $h = 0$.

La démonstration de la proposition 12.6.4 s'applique mot pour mot, en remplaçant f par f_j. $\square$

On en déduit la proposition suivante :

Corollaire 12.6.6 *Pour toute fonction $f \in V_0$, pour tout $\varepsilon > 0$, il existe $r \geq 0$ tel que*

$$\|P^{r+n} f\|_\infty = \varepsilon \; m(P^n f), \; \forall n \geq 0.$$

Théorème 12.6.7 *Pour toute fonction f vérifiant $M(f) < \infty$, il existe g telle que $m(g) < \infty$ et $f = g - Pg$.*

Preuve : Posons $\rho = C\varepsilon \; \sum_{j=1}^{\infty} m_j(u)$ et choisissons $\varepsilon > 0$ tel que l'on ait $\rho < 1$. Soit $r \geq 0$ donné par la proposition précédente.

On a, d'après l'inégalité (12.24) et en majorant $\|P^{n+r-j} f\|_\infty$ par $\varepsilon \; m(P^{n-j} f)$:

$$m(P^{n+r} f) \leq m_n(P^r f) + C\varepsilon \sum_{j=1}^{n} m_j(u) m(P^{n-j} f), \; \forall n \geq 0;$$

d'où :

$$\begin{aligned}
\sum_{n=0}^{N} m(P^{n+r} f) &\leq \sum_{n=0}^{N} m_n(P^r f) + C\varepsilon \sum_{n=0}^{N} \sum_{j=1}^{n} m_j(u) m(P^{n-j} f) \\
&= \sum_{n=0}^{N} m_n(P^r f) + C\varepsilon \sum_{j=1}^{N} m_j(u) \sum_{n=0}^{N-j} m(P^n f) \\
&\leq \sum_{n=0}^{N} m_n(P^r f) + C\varepsilon \; (\sum_{j=1}^{N} m_j(u))(\sum_{n=0}^{N} m(P^n f)) \\
&\leq \sum_{n=0}^{N} m_n(P^r f) + \rho \sum_{n=0}^{N} m(P^n f).
\end{aligned}$$

On a donc :

$$\sum_{n=0}^{N+r} m(P^n f) \leq \sum_{n=0}^{r-1} m(P^n f) + \sum_{n=0}^{N} m_n(P^r f) + \rho \sum_{n=0}^{N+r} m(P^n f).$$

d'où :

$$\begin{aligned}
\sum_{n=0}^{N+r} m(P^n f) &\leq \frac{1}{1-\rho} [\sum_{n=0}^{r-1} m(P^n f) + \sum_{n=0}^{N} m_n(P^r f)] \\
&\leq \frac{1}{1-\rho} [\sum_{n=0}^{r-1} m(P^n f) + M(f) + Crm(f)M(u)],
\end{aligned}$$

compte-tenu de la majoration

$$m_n(P^r f) \leq m_{n+r}(f) + C \sum_{j=1}^{r} m_{n+j}(u) m(f).$$

Ceci implique la convergence de la série $\sum_{n=0}^{\infty} m(P^n f) < \infty$ et donc la convergence :

$$\sum_{n=0}^{\infty} \|P^n f\|_\infty < \infty.$$

La somme $g = \sum_{n=0}^{\infty} P^n f$ est une fonction continue, et même régulière, vérifiant :

$$f = g - Pg.$$

□

Comme application, nous obtenons un théorème de la limite centrale pour les fonctions assez régulières.

Nous supposons qu'il existe une application surjective τ de E sur E laissant invariante l'unique mesure de probabilité P-invariante ν telle que l'opérateur P soit le dual de l'opérateur T de composition par τ ; c'est-à-dire que

$$\int_E Pf\ g\ d\nu = \int_E f\ g \circ \tau\ d\nu,$$

pour toutes fonctions f et g de $L^2(E, \nu)$.

Proposition 12.6.8 *Sous l'hypothèse $M(u) < +\infty$, pour toute fonction f vérifiant $\nu(f) = 0$ et $M(f) < \infty$, la fonction f est homologue à une différence de martingales renversées. En particulier la suite $(T^n f)$ dans le système dynamique $(E, \mathcal{E}, \nu, \tau)$ vérifie le TCL.*

Preuve : Nous avons vu qu'il existe, sous les hypothèses du théorème, une fonction continue g telle que $f = g - Pg$.

On peut alors écrire $f = g_1 + h - Th$, avec $g_1 = g - TPg$ et $h = TPg - Pg$. La fonction f est homologue à g_1. La relation $Pg_1 = 0$ montre que la suite $(T^n g_1)$ est une suite d'accroissements de martingales renversées. Cette suite est stationnaire et ergodique et vérifie donc le théorème de la limite centrale (théorème 11.2.2). □

12.7 Notion de pression et principe variationnel

Entropie topologique

Considérons un système dynamique topologique (E, τ), où E est un espace métrique compact et τ une application continue et surjective de E dans lui-même.

Notons $\mathcal{O}$ l'ensemble des recouvrements ouverts de E. Pour $\mathcal{U} \in \mathcal{O}$, on appelle **entropie du recouvrement** $\mathcal{U}$, la quantité $H(\mathcal{U}) = \mathrm{Log} N(\mathcal{U})$, où $N(\mathcal{U})$ est le nombre minimum d'ouverts de $\mathcal{U}$ recouvrant E.

Il existe sur $\mathcal{O}$ une structure d'ordre et de treillis. De plus τ opère sur $\mathcal{O}$, par $\tau^{-1}\mathcal{U} = \{\tau^{-1}A, A \in \mathcal{U}\}$. On peut donc définir

$$h(\mathcal{U}, \tau) = \lim_n \frac{1}{n} H(\mathcal{U} \vee \tau^{-1}\mathcal{U} \cdots \vee \tau^{-n+1}\mathcal{U}).$$

Définition 12.7.1 On appelle **entropie topologique** du système (E, τ) le nombre

$$h(\tau) = \sup_{\mathcal{U} \in \mathcal{O}} h(\mathcal{U}, \tau).$$

Notons $\mathcal{M}(\tau)$ l'ensemble des mesures de probabilité τ-invariantes sur E. Pour toute mesure invariante μ sur E, on note $h_\mu(\tau)$ l'entropie du système mesuré (E, μ, τ).

Les résultats suivants, dus respectivement à L. Goodwyn et T.N.T. Goodman, relient l'entropie topologique $h(\tau)$ et les entropies métriques (i.e. au sens de la mesure) $h_\mu(\tau)$.

Théorème 12.7.2

$$h_\mu(\tau) \leq h(\tau), \ \forall \mu \in \mathcal{M}(\tau),$$

$$\sup_{\mu \in \mathcal{M}(\tau)} h_\mu(\tau) = h(\tau).$$

Considérons maintenant un système (E, τ) inversible et **expansif**, c'est-à-dire tel qu'il existe un nombre $\delta > 0$ vérifiant :

$$\sup_{n \in \mathbb{Z}} d(\tau^n x, \tau^n y) \geq \delta, \forall x, y \in E, x \neq y.$$

Sous l'hypothèse d'expansivité, le système possède une partition "régulière" génératrice, ce qui assure que l'entropie définit une fonction semi-continue supérieurement (scs) $\mu \to h_\mu(\tau)$. La borne supérieure, dans le théorème 12.7.2 est donc atteinte pour une mesure ν (dite mesure d'entropie maximale). (Ceci est faux en général, sans l'hypothèse d'expansivité).

Notion de pression

Généralisant des notions classiques en mécanique statistique, on peut définir la notion de pression et d'état d'équilibre pour un système dynamique topologique (E, τ) et une fonction continue f sur E. La "pression", quand on choisit f identiquement nulle, redonne l'entropie topologique.

Soit $\mathcal{U}$ un recouvrement ouvert de E par des boules $B_1, ..., B_l$ de rayon $< \delta$. On peut définir, pour $f \in C(E)$, la quantité

$$Z_n(f) = \sum_{i_0, ..., i_{n-1} \in \{1, ..., l\}^{\mathbb{N}}} \exp \sup \{\sum_0^{n-1} f(\tau^i x) \ : \tau^k x \in B_{i_k}, k = 0, ..., n-1\}.$$

La pression est définie comme la limite

$$P(f) = \limsup_n \frac{1}{n} \mathrm{Log} Z_n(f).$$

La fonction $f \to P(f)$ est positive ou nulle, convexe, sous additive, et lipschitzienne :

$$|P(f) - P(g)| \leq \|f - g\|_\infty.$$

Le principe variationel, contenu dans l'énoncé du théorème suivant, étend un résultat classique en mécanique statistique et se réduit, dans le cas où f est identiquement nulle, aux théorèmes de Goodwyn et Goodman.

Théorème 12.7.3

$$P(f) = \sup_{\mu \in \mathcal{M}(\tau)} (h_\mu(\tau) + \int f \ d\mu). \tag{12.27}$$

Les états, ou mesures d'équilibre, sont les mesures $\mu \in \mathcal{M}(\tau)$ telles que $h(\mu) + \int f \, d\mu = P(f)$. Pour un système expansif, la quantité figurant à droite dans (12.27 étant scs, il existe au moins une mesure invariante réalisant la borne supérieure, et donc au moins une mesure d'équilibre associée à f.

Nous allons préciser ce théorème, en nous plaçant dans le cas où la transformation est le décalage sur un espace $I^{\mathbb{N}}$, I étant un espace d'états fini, la fonction f étant supposée **régulière** i.e. höldérienne. Rappelons la définition.

On définit sur Ω^+ une métrique d par $d(x,y) = 2^{-N}$, où N est le plus grand entier ≥ 0 tel que $x_i = y_i, \forall\ 0 \leq i < N$.

Pour une fonction f définie sur Ω^+ à valeurs réelles, on pose

$$\mathcal{V}(f, 2^{-k}) = \sup\{|f(x) - f(y)|, x, y \in \Omega^+ \ : x_i = y_i, 0 \leq i \leq k\}.$$

Une fonction f est dite höldérienne s'il existe $\alpha \in]0,1[$ et une constante C, tels que :

$$\mathcal{V}(f, 2^{-k}) \leq C\alpha^k, \forall k \in \mathbb{N}.$$

12.8 Opérateur de transfert et mesures de Gibbs

Pour des fonctions f régulières, on peut construire, grâce à l'opérateur de transfert, la mesure d'équilibre (mesure de Gibbs) donnant le maximum dans l'expression (12.27) du principe variationnel.

Notons Q_f l'opérateur défini par

$$Q_f\phi(x) = \sum_{y \ :\tau y = x} \exp(f(y))\phi(y).$$

Nous utilisons la proposition 12.5.4 qui fournit une fonction propre $v > 0$ continue (et ayant même une régularité dépendant de celle de f) et une valeur propre que nous noterons ici $\exp(\rho(f))$, pour l'opérateur Q_f. Quand f est höldérienne, v est aussi höldérienne.

Sur l'espace produit Ω^+, désignons par (X_n) les applications coordonnées. Avec les notations utilisées dans le chapitre sur l'entropie, désignons par α la partition génératrice correspondant à la première coordonnée. Si $A \in I^n$ est un bloc de longueur n, notons $C_n(A)$ le "cylindre" déterminé par ce bloc. Quand A varie dans I^n, $C_n(A)$ décrit les atomes de la partition $\alpha_0^{n-1} = \alpha \vee \tau^{-1}\alpha \cdots \vee \tau^{-n+1}\alpha$ associée aux n premières coordonnées.

Pour tout point $x \in \Omega^+$, nous noterons également $C_n(x)$ le cylindre déterminé par le bloc des n premières coordonnées du point x, autrement dit l'atome de la partition α_0^{n-1} contenant x.

Théorème 12.8.1 *Si f est höldérienne, il existe une mesure de probabilité ν_f sur Ω^+ invariante par le décalage τ telle que l'on ait, pour des constantes c_1, c_2 :*

$$c_1 \leq \frac{\nu_f(C_n(x))}{\exp(-n\rho(f) + \sum_0^{n-1} f(\tau^k x))} \leq c_2, \forall n \in \mathbb{N}, \forall x \in \Omega^+, \tag{12.28}$$

où $e^{\rho(f)}$ est la valeur propre positive de l'opérateur Q_f.

Preuve : Soit v la fonction propre de l'opérateur Q_f. Posons $v = \exp(w)$. On peut écrire $f = \rho(f) + Tw - w + g$, avec $\sum_{a\in A} \exp g(ax) = 1$.

Considérons l'opérateur "relativisé" P_g déduit de Q_f :

$$\begin{aligned} P_g\phi(x) &= \sum_{y\ :\tau y = x} e^{-\rho(f)} \frac{\exp(w(y))}{\exp(Tw(y))} \exp(f(y))\phi(y) \\ &= \sum_{y\ :\tau y = x} \exp(g(y))\phi(y). \end{aligned}$$

L'opérateur P_g est markovien, possède une unique mesure invariante ν_f et est l'adjoint de T, par rapport à ν_f :

$$\int \phi\ T\psi\ d\nu_f = \int P_g(\phi\ T\psi)\ d\nu_f = \int P_g\phi\ \psi\ d\nu_f.$$

La mesure ν_f est donc invariante par le décalage τ et vérifie, pour ϕ continue sur Ω^+ :

$$\mathbb{E}_{\nu_f}(\phi | X_n = x_n, n > 1) = \sum_{a\in A} \exp(g(ax_1...))\phi(ax_1...).$$

En particulier, on a :

$$\mathbb{E}_{\nu_f}(X_0 = x_0 | X_n = x_n, n \geq 1) = e^{g(x_0x_1...)}, \tag{12.29}$$

d'où

$$\mathbb{E}_{\nu_f}(X_0 = x_0, ..., X_{m-1} = x_{m-1} | X_n = y_n, n \geq m) = e^{\sum_{k=0}^{m-1} g(x_k...x_{m-1}y_my_{m+1}...)},$$

et en intégrant :

$$\nu_f(X_0 = x_0, ..., X_{m-1} = x_{m-1}) = \int e^{\sum_{k=0}^{m-1} g(x_k...x_{m-1}y_my_{m+1}...)}\ d\nu_f(y).$$

La régularité de g permet alors d'écrire

$$\nu_f(X_0 = x_0, ..., X_{m-1} = x_{m-1}) = e^{\sum_{k=0}^{m-1} g(x_k...x_{m-1}x_mx_{m+1}...)} \int e^{[\ \ ]}\ d\nu_f(y),$$

où le crochet dans l'intégrale est la différence

$$-\sum_{k=0}^{m-1} g(x_k...x_{m-1}x_mx_{m+1}...) + \sum_{k=0}^{m-1} g(x_k...x_{m-1}y_my_{m+1}...),$$

différence majorée en valeur absolue par $C = \sum_{k\geq 1} \mathcal{V}(g, 2^{-k})$. En exprimant g à l'aide de f et de w, on obtient donc l'encadrement

$$e^{-C} \leq \frac{\nu_f(C_n(x))}{\exp(-n\rho(f) + \sum_0^{n-1} f(\tau^k x) + w(x) - w(\tau^n x))} \leq e^C, \forall n \in \mathbb{N}, \forall x \in \Omega^+.$$

La quantité $w(x) - w(\tau^n x)$ restant bornée, on a bien le résultat annoncé. □

La mesure ν_f est appelée *mesure de Gibbs* associée à f. On peut maintenant préciser le principe variationnel.

A une fonction f, associons la "pression" définie comme la limite :

$$P(f) = \lim_n \frac{1}{n} \mathrm{Log} Z_n(f),$$

où $Z_n(f)$ s'écrit ici sous la forme

$$Z_n(f) = \sum_{(a_0,\dots,a_{n-1})\in I^n} \exp\{\sup_{x\ :x_i=a_i, 0\le i<n} (\sum_0^{n-1} f(\tau^k x))\}.$$

La limite existe car la suite (Z_n) est sous-additive. Pour les fonctions régulières, le principe variationnel est précisé par le théorème suivant.

Théorème 12.8.2 *Soit f höldérienne. La constante $\rho(f)$ du théorème 12.8.1. est égale à la pression $P(f)$. On a :*

$$h_\mu(\tau) + \int f\ d\mu \le P(f), \tag{12.30}$$

pour toute mesure μ invariante par τ et la mesure de Gibbs ν_f associée à f est l'unique mesure invariante par τ réalisant l'égalité dans (12.30).

Preuve : 1) D'après (12.28), on a :

$$\begin{aligned}
&c_2^{-1}\nu_f(X_0 = a_0, \dots, X_{m-1} = a_{m-1}) \\
&\le \exp(-m\rho(f)) \exp\{\sup_{x\ :x_i=a_i, 0\le i<m} (\sum_0^{m-1} f(\tau^k x))\} \\
&\le\ c_1^{-1}\nu_f(X_0 = a_0, \dots, X_{m-1} = a_{m-1}),
\end{aligned}$$

d'où :

$$c_2^{-1} \le e^{-m\rho(f)} Z_m(f) \le c_1^{-1}.$$

A partir de la relation (12.29), on obtient $h_{\nu_f}(\tau) = -\int g(x)\ d\nu_f(x)$, d'où, en revenant à l'expression de f $(f = \rho(f) + g + Tw - w)$, l'égalité :

$$\rho(f) = \nu_f(f) + h_{\nu_f}(\tau).$$

2) Notons u la fonction $x \to -x\ \mathrm{Log} x$. Par concavité de cette fonction, nous avons :

$$\begin{aligned}
&-\ \sum_{A\in I^n} \nu_f(C_n(A)) \frac{\mu(C_n(A))}{\nu_f(C_n(A))}\ \mathrm{Log} \frac{\mu(C_n(A))}{\nu_f(C_n(A))} \\
&\le\ u(\sum_{A\in I^n} \nu_f(C_n(A)) \frac{\mu(C_n(A))}{\nu_f(C_n(A))}) = u(1) = 0.
\end{aligned}$$

D'où :

$$-\sum_{A\in I^n} \mu(C_n(A)) \mathrm{Log} \mu(C_n(A)) \le -\sum_{A\in I^n} \mu(C_n(A)) \mathrm{Log} \nu_f(C_n(A)). \tag{12.31}$$

Désignons par α la partition génératrice correspondant à la première coordonnée et utilisons (12.28). Choisissons pour chaque $A \in I^n$ un point x_A dans $C_n(A)$. La somme $\sum_{A\in I^n} \mu(C_n(A))(S_n f(x_A))$ est une approximation de $\mu(f)$ et la régularité de f permet de contrôler l'erreur, quand on remplace cette somme par l'intégrale :

Nous avons, pour des quantités B_n et D_n bornées en n :

$$|\sum_{A\in I^n} \mu(C_n(A))(S_n f(x_A)) - n\mu(f)| < B_n;$$

$$\begin{aligned} H_\mu(\alpha_0^{n-1}) &\leq -\sum \mu(C_n(A))(-n\rho(f) + S_n f(x_A) + B_n) \\ &\leq n\rho(f) - \sum_{A \in I^n} \mu(C_n(A))(S_n f(x_A) + B_n). \end{aligned}$$

d'où :

$$H_\mu(\alpha_0^{n-1}) + n\mu(f) \leq n\rho(f) + D_n.$$

En divisant par n et en passant à la limite, on obtient l'inégalité :

$$h_\mu(\alpha, \tau) + \mu(f) \leq \rho(f).$$

Il y a égalité dans l'inégalité pour $\mu = \nu_f$, le membre de gauche étant alors égal à $\rho(f)$.

3) Montrons maintenant l'unicité de la mesure réalisant la borne supérieure dans (12.30). Nous utilisons un argument de Spitzer. Soit μ une mesure de probabilité invariante par le décalage. Considérons le processus sur (Ω^+, μ) formé des rapports de vraisemblance

$$V_n(x) = \frac{\nu_f(C_n(x))}{\mu(C_n(x))},$$

si $\mu(C_n(x)) \neq 0$, et $V_n(x) = 0$, si $\mu(C_n(x)) = 0$.

La suite (V_n) forme une martingale pour la filtration naturelle sur Ω^+ muni de la mesure μ. Cette suite converge, μ-presque partout, vers la dérivée de Radon-Nikodym de ν_f par rapport à μ, donc vers 0, si l'on choisit μ singulière par rapport à ν_f.

Soit $W_n = -\mathrm{Log} V_n$. On a :

$$\mathbb{E}_\mu(W_n) = \sum_{(a_0, \ldots, a_{n-1}) \in I^n} \mu(C_n(a_0, ..., a_{n-1})) \mathrm{Log} \frac{\mu(C_n(a_0, ..., a_{n-1}))}{\nu_f(C_n(a_0, ..., a_{n-1}))}.$$

La partie négative de W_n est d'espérance bornée, d'après la majoration.

$$\mathbb{E}_\mu(W_n^-) = \mathbb{E}_\mu((\mathrm{Log} V_n)^+) \leq \mathbb{E}_\mu((V_n - 1)^+) \leq \mathbb{E}_\mu(V_n) = 1.$$

On a donc

$$\lim_{n \to \infty} \mathbb{E}_\mu(W_n) = \infty.$$

Dans la relation (12.31), en utilisant le fait que l'entropie $h_\mu(\alpha, \tau)$ est la borne inférieure des entropies moyennes $\frac{1}{n} H_\mu(\alpha_0^{n-1})$, nous pouvons écrire :

$$h_\mu(\alpha, \tau) + \mu(f) \leq \frac{1}{n} H_\mu(\alpha_0^{n-1}) + \mu(f) \leq \rho(f) + \frac{1}{n}(D_n - B_n),$$

où, d'après l'argument de martingale, la suite (D_n) tend vers $-\infty$.

Nous avons donc $D_n - B_n < 0$, à partir d'un certain rang n_0. A partir de ce rang, l'inégalité est alors stricte, ce qui prouve que

$$h_\mu(\alpha, \tau) + \mu(f) < \rho(f).$$

□

Chapitre 13

Exercices

$\boxed{1}$

Isométries d'un espace métrique compact

On considère un espace métrique (E,d) et une **isométrie** τ de cet espace, c'est à dire une application de E dans E préservant la distance d.

1) Etude d'un exemple :

On prend pour E le cercle unité et pour τ une rotation τ_α d'angle α sur E.

a) Discuter la densité des orbites de τ_α suivant que $\frac{\alpha}{2\pi}$ est rationnel ou non.

b) Discuter l'unicité ou la non unicité des mesures invariantes suivant que $\frac{\alpha}{2\pi}$ est rationnel ou non.

c) Montrer que, dans tous les cas, l'espace E se décompose en une réunion de fermés minimaux, portant une unique mesure invariante par τ_α.

2) On considère maintenant le cas général.

a) Montrer que, si τ possède une orbite dense, toutes ses orbites $\{\tau^n x, n \geq 0\}$ sont denses.

b) Montrer que, si τ possède une orbite dense, il existe une unique mesure de probabilité sur E invariante par τ.

c) Montrer, comme pour le cercle, que, dans tous les cas, l'espace E se décompose en une réunion de fermés minimaux, portant une unique mesure invariante par τ.

$\boxed{2}$

Transformations distales

Etant donné un espace métrique (E,d), une transformation τ de E dans E est dite **distale**, si elle vérifie la condition :

$$\inf_n d(\tau^n x, \tau^n y) = 0 \;\Rightarrow\; x = y.$$

1) a) Montrer qu'une isométrie est une transformation distale.

b) Soient τ une transformation distale sur un espace (E,d) et φ une application continue de E dans le cercle unité $\mathbb{T}^1$.

On note σ_φ la transformation de $E \times \mathbb{T}^1$ dans lui même d éfinie par :

$$\sigma_\varphi(x,y) = (\tau x, \varphi(x)y).$$

Montrer que σ_φ est distale.

$\boxed{3}$

Propriétés spectrales

On considère un espace probabilisé $(E, \mathcal{A}, \mu)$ et deux transformations τ et σ de E dans E, $\mathcal{A}$-mesurables et laissant μ invariante.

1) a) On suppose σ ergodique. Montrer que si $\tau^k = \sigma$, pour un entier $k > 0$, alors τ est ergodique.

b) Donner des exemples simples de transformations σ ergodiques, telles que σ^k ne soit pas ergodique.

2) On suppose dans la suite que τ et σ commutent $(\tau\sigma = \sigma\tau)$ et que σ est ergodique.

a) Montrer que toute fonction propre pour σ est une fonction propre pour τ.

b) On prend pour système $(E, \mathcal{A}, \mu, \sigma)$ une rotation ergodique sur le cercle muni de la mesure de Lebesgue.

Montrer que τ est à spectre discret.

c) On prend pour $(E, \mathcal{A}, \mu, \sigma)$ un système à spectre continu (pas de fonctions propres autres que les constantes), par exemple un schéma de Bernoulli).

Montrer que τ est soit à spectre continu, soit non ergodique.

$\boxed{4}$

Etant donné un nombre réel α, on note (E, τ_α, μ) le système dynamique défini par la rotation $x \to \tau_\alpha(x) = x + \alpha \ mod\ 1$ sur le cercle unité $\mathbb{T}^1$ identifié à $\mathbb{R}/\mathbb{Z}$ et muni de la mesure de Lebesgue λ.

On note T l'opérateur de composition par τ_α.

1) Pour $f \in L^1(\lambda)$, expliciter la valeur de

$$\lim_N \frac{1}{N} \sum_0^{N-1} f(x + n\alpha),$$

en distinguant suivant que α est ou non irrationnel.

2) Soit φ une application mesurable de $\mathbb{T}^1$ dans $\mathbb{T}^1$.

On définit sur $Z = \mathbb{T}^1 \times \mathbb{T}^1$ une transformation en posant

$$\sigma_{\alpha,\varphi}(x, y) = (x + \alpha, y + \varphi(x)).$$

a) Montrer que σ laisse invariante la mesure produit $\lambda \times \lambda$ sur Z.

b) Soit $f(x, y)$ une fonction dans $L^1(\lambda \times \lambda)$. On note $\widehat{f}_n(y)$ les "coefficients de Fourier de f par rapport à la variable x" définis par :

$$\widehat{f}_n(y) = \int f(x, y) \exp(-2\pi i n x) dx.$$

Montrer que, si f est invariante par σ,

$$\widehat{f}_n(y + \alpha) = \exp(2\pi i n \varphi(y)) f_n(y).$$

3) [Critère d'ergodicité de $\sigma_{\alpha,\varphi}$]

On suppose dans la suite que α est irrationnel.

a) Montrer que $|\widehat{f}_n|$ est constante.

b) Montrer que $\sigma_{\alpha,\varphi}$ est non ergodique si et seulement s'il existe un entier $p > 0$ et une fonction u mesurable de module 1, tels que

$$u(y+\alpha) = \exp(-2\pi i p \varphi(y))u(y).$$

c) Pour p entier non nul, on note χ_p la fonction $\chi_p(y) = \exp(2\pi i p y)$.

En utilisant le théorème ergodique, montrer que la projection de χ_p sur les invariants de τ est de la forme $\theta_p(x)\chi_p(y)$, où θ est une fonction mesurable de période 1 telle que

$$\theta(x+\alpha) = \theta(x)\exp(-2\pi i p \phi(x)).$$

Retrouver à l'aide de ce résultat le critère d'ergodicité de précédent.

4) On prend pour φ la fonction $\varphi(x) = x$.

a) Exprimer la n-ième itérée de la transformation $\sigma_{\alpha,\varphi}$.

b) Montrer que $\sigma_{\alpha,\varphi}$ est ergodique.

c) En déduire que, pour $f \in L^1(\lambda)$, on a, pour presque tout y

$$\lim \frac{1}{N} \sum_0^{N-1} f(y + n(n-1)\alpha) = \int_0^1 f(y)\, dy.$$

5) [Unique ergodicité]

Soit μ une mesure invariante par τ sur $\mathbb{T}^2$. Pour ψ continue positive et telle que $\int_0^1 \psi(y)dy = 1$, on définit la mesure μ_ψ par

$$\int f(x,y)\, d\lambda_\psi(x,y) = \int\int f(x, y+u)\, du\, d\lambda(x,y).$$

a) Montrer que μ_ψ est invariante par τ et qu'elle est absolument continue par rapport à la mesure de Lebesgue λ.

b) En déduire que $\mu_\psi = \lambda$, pour tout ψ, puis que $\mu = \lambda$.

c) Montrer que τ est uniquement ergodique, et en déduire, pour α irrationnel, l'équirépartition de la suite $(n^2\alpha \; mod \; 1)_{n\geq 1}$.

6) [Etude spectrale]

On décompose l'espace $L^2(\mathbb{T}^2)$ en $H_0 + H_0^\perp$, où H_0 est le sous-espace des fonctions ne dépendant pas de la deuxième coordonnée et $H_0^\perp$ son orthogonal.

a) Quel est le spectre de la restriction de T au sous-espace H_0.

b) En utilisant les fonctions $\chi_{n,p} : (x,y) \to \exp(2\pi i(nx+py))$, pour $n,p \in \mathbb{Z}$, $p \neq 0$, former une base orthonormée $f_{i,k}$ de $H_0^\perp$ telle que $f_{i,k} = T^k f_{i,0}$.

c) En déduire que le spectre de la restriction de τ à $H_0^\perp$ est de Lebesgue, et que, pour $f, g \in H_0^\perp$, on $\lim_n \int T^k f \; g d\lambda = 0$.

$\boxed{5}$

Sur $E = \{0,1\}^{\mathbb{N}}$, on considère la structure de groupe définie de la façon suivante :

pour $x = (x_n), y = (y_n)$ dans E, $z = x + y$ est défini (en convenant que $r_{-1} = 0$) par :

$$\begin{cases} z_n = x_n + y_n + r_{n-1} \ mod\ 2, n \geq 0, \\ r_n = x_n + y_n + r_{n-1} \ div\ 2, n \geq 0. \end{cases}$$

Il s'agit donc de l'addition avec retenue, qui correspond à l'addition sur $\mathbb{Z}$ exprimée dans la représentation des entiers en base 2.

On pose $\beta = (1, 0, 0, ...)$.

Vérifier que les éléments 2β, 3β,... s'écrivent sous la forme : 2β = (0,1,0,0,...), 3β = (1,1,0,0,...), 4β = (0,0,1,0,...),

a) En raisonnant sur les cylindres, montrer que la mesure produit $\mu = \{\frac{1}{2}, \frac{1}{2}\}^{\mathbb{N}}$ est invariante par translation dans le groupe E.

b) Montrer que la transformation τ définie sur (E, μ) par $x \to \tau(x) = x + \beta$ est uniquement ergodique, d'unique mesure invariante μ.

c) Montrer que τ est à spectre discret et que ses valeurs propres sont toutes les racines 2^n-ième de l'unité.

d) Etant donnée une fonction f sur E, décrire un procédé de calcul de $\int f d\mu$ utilisant l'élément β du groupe E et ses "multiples" $n\beta$ (au sens du groupe).

$\boxed{6}$

Soit α un réel dans $]0, 1[$. Sur $E = [0, 1]$, on définit une transformation τ (transformation à deux pentes) par

$$x \to \tau(x) \ = \ \begin{cases} \alpha^{-1}x, \text{ si } x < \alpha, \\ (1-\alpha)^{-1}(x-\alpha), \text{ si } x > \alpha, \end{cases}$$

a) Montrer que la mesure uniforme λ sur $[0, 1]$ est invariante par τ.

b) Quelle transformation trouve-t-on pour $\alpha = 1/2$?

c) Montrer que la partition Γ de E en deux intervalles $C_1 = [0, \alpha[$ et $C_2 = [\alpha, 1[$ est génératrice pour τ.

(La situation considérée ici est celle d'un "développement en base α" généralisant le développement en base 2. En particulier on peut considérer les intervalles (analogues des intervalles dyadiques) homothétiques de C_1 et C_2 par des homothéties de rapport $\alpha^k(1-\alpha)^{n-k}$).

d) A l'aide du codage utilisant la partition Γ, montrer que (E, m, τ) est isomorphe au "shift" sur l'espace $E = \{0,1\}^{\mathbb{N}}$ muni de la mesure produit $(\alpha, (1-\alpha))^{\mathbb{N}}$.

e) En déduire que (E, λ, τ) est ergodique.

f) Calculer l'entropie de ce système.

$\boxed{7}$

On considère un espace d'états fini $\mathcal{E}$ de cardinal l et une matrice A de dimension $l \times l$, formée de 0 et de 1.

On suppose qu'il existe une puissance A^r de A ayant tous ses coefficients > 0. Dans l'espace produit $E = \mathcal{E}^{\mathbb{N}}$, on appelle **suite admissible** toute suite $\omega = (\omega_0, \omega_1, ..., \omega_n, ...)$ telle que

$$A(\omega_k, \omega_{k+1}) = 1, \text{ pour tout } k \geq 0.$$

1) Montrer que l'ensemble E_0 des suites admissibles est un fermé de E, invariant par le shift θ.

2) On appelle mot admissible de longueur $n+1$, toute suite finie $(\omega_0, \omega_1, ..., \omega_n)$ d'éléments de I tel que

$$A(\omega_k, \omega_{k+1}) = 1, \text{ pour } 0 \leq k < n.$$

On note $\varphi(n)$, (resp. $\varphi_i(n)$), le nombre de mots admissibles de longueur $n+1$ (resp. de longueur $n+1$ se terminant par i).

Etablir une formule de récurrence sur la suite $(\varphi_i(n))_{n\geq 1}$ et calculer $\lim_n \frac{1}{n} \ln(\varphi_i(n))$, à l'aide de la plus grande valeur propre de A.

3) On définit sur E_0 une mesure markovienne invariante par le shift θ de la façon suivante :

Soient U et V des vecteurs de coordonnées > 0 vérifiant :

$$AU = \lambda U, {}^t AV = \lambda V, \text{ et tels que } \sum_{i=1}^{\ell} U_i V_i = 1.$$

On pose $M(i,j) = \dfrac{1}{\lambda} \dfrac{U_j}{U_i} A(i,j), \; P(i) = U_i V_i$.

a) Vérifier que la mesure μ définie en restriction aux cylindres de E_0 par :

$$\mu(X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = P(i_0) M(i_0, i_1) ... M(i_{n-1}, i_n)$$

est invariante par le shift θ.

b) Calculer l'entropie $h_\mu(\tau)$ de (E_0, μ, θ).

4) Montrer que, pour toute mesure ν sur E_0 invariante par θ, l'entropie h_ν de (E_0, ν, τ) vérifie $h_\nu \leq \ln \lambda$.

(En fait, μ est l'unique mesure invariante sur E_0 telle que $h_\mu = \ln \lambda$).

$\boxed{8}$

On considère un système dynamique $(E, \mathcal{A}, \mu, \tau)$ et une partition finie α de E en r atomes. On note $\alpha_0^{n-1} = \alpha \vee ... \vee \tau^{-n+1} \alpha$.

On rappelle que $h = h(\alpha, \tau) = \lim_n \frac{1}{n} H(\alpha_0^{n-1})$ désigne l'entropie moyenne de α relativement à θ.

Pour γ réel, on note $\phi(n, \gamma)$ le nombre d'atomes de la partition α_0^{n-1} dont la mesure est supérieure $e^{-n\gamma}$, et $\theta(n)$ le nombre d'atomes de α_0^{n-1} de mesure > 0.

1) a) Montrer que la suite $(\ln \theta(n))_{n\geq 1}$ est sous-additive et que la limite $\beta = \lim_n \frac{1}{n} \ln \theta(n)$ existe.

b) Montrer que β majore l'entropie moyenne $h(\alpha, \tau)$.

c) En déduire que, pour que l'entropie $h(\tau)$ soit non nulle, il faut qu'il existe une partition α telle que le nombre d'atomes de α_0^{n-1} de mesure non nulle croisse exponentiellement.

2) a) Montrer que, pour tout $\gamma > h$, on a

$$\ln r \geq \liminf_n \frac{1}{n} \ln \phi(n, \gamma) \geq h.$$

(On pourra utiliser le théorème de Shannon-Mac Millan.)

b) En déduire que, pour que l'entropie $h(\tau)$ soit non nulle, il faut qu'il existe $\gamma < \infty$ et une partition α telle que le nombre d'atomes de α_0^{n-1} de mesure supérieure à $e^{-n\gamma}$ croisse exponentiellement.

3) On prend pour E le cercle unité muni de la tribu borélienne et de la mesure de Lebesgue, et pour τ une rotation ergodique.

Soit α une partition du cercle en intervalles.

a) En étudiant la croissance du nombre d'éléments de α_0^{n-1}, montrer que l'entropie $h(\alpha, \tau)$ est nulle.

b) En déduire que l'entropie $h(\tau)$ est nulle.

4) En prenant toujours pour E le cercle unité, on choisit maintenant pour τ une application continue surjective de E sur lui-même.

a) Existe-t-il une mesure μ invariante par τ et telle que le système (E, μ, τ) soit ergodique ?

b) Montrer que si τ est une bijection, son entropie est nulle, pour toute mesure invariante.

c) Donner des exemples, quand τ n'est pas injective, d'entropie non nulle.

$\boxed{9}$

On considère un espace probabilisé $(E, \mathcal{A}, \mu)$ et une application τ de E dans E qui est inversible, mesurable ainsi que son inverse et préserve μ (on dit que τ est un automorphisme de $(E, \mathcal{A}, \mu)$).

On suppose l'espace probabilisé $(E, \mathcal{A}, \mu)$ sans atomes, et tel qu'il existe dans $\mathcal{A}$ une famille, notée $\mathcal{A}_0$, dénombrable dense de sous-ensembles.

(On peut munir la tribu des sous-ensembles mesurables d'une distance en posant $d(A, B) = \mu(A \Delta B)$. Si deux sous-ensembles sont à une distance nulle, ils sont égaux à un ensemble de mesure nulle près.)

On suppose que τ vérifie la propriété (P) suivante :

il existe une constante K telle que

$$\limsup_n \mu(A \cap \tau^{-n} B) \leq K \mu(A) \mu(B), \ \forall A, B \in \mathcal{A}.$$

1) Montrer que la tribu des invariants du système $(E, \mathcal{A}, \mu, \tau)$ est finie.

2) On suppose dans la suite que le système $(E, \mathcal{A}, \mu, \tau^k)$ est ergodique, pour tout $k \geq 1$.

a) Montrer que tout facteur de $(E, \mathcal{A}, \mu, \tau)$ vérifie la condition (P).

b) Montrer que le système défini par une rotation sur le cercle ne peut vérifier la propriété (P).

c) En déduire que le système $(E, \mathcal{A}, \mu, \tau)$ est à spectre continu.

3) a) Montrer que l'on peut extraire de toute suite strictement croissante d'entiers une sous-suite strictement croissante (n_k) telle que, pour A et B quelconques dans $\mathcal{A}$, la suite $(\mu(A \cap \tau^{-n_k} B)_{k \in \mathbb{N}})$ ait une limite.

(On raisonnera dans $\mathcal{A}_0$, puis on montrera que cette convergence est encore vérifiée pour tout A et tout B dans $\mathcal{A}$.)

On note $\lambda(A \times B)$ la limite.

b) Montrer (ou admettre !) que λ peut être prolongée en une mesure sur l'espace produit $(E \times E, \mathcal{A} \times \mathcal{A})$, qui est absolument continue par rapport à $\mu \times \mu$ et invariante par $\tau \times \tau$.

4) En déduire que $\lambda(A \times B) = \mu(A)\mu(B),\ \forall A, B \in \mathcal{A}$.

5) Conclure que le système (E, τ) est fortement mélangeant, i.e. vérifie la propriété :

$$\lim_n \mu(A \cap \tau^{-n} B) = \mu(A)\mu(B),\ \forall A, B \in \mathcal{A}.$$

$\boxed{10}$

Sur la récurrence

On considère un espace mesuré $(E, \mathcal{A}, \mu)$ (μ n'étant pas nécessairement une mesure finie) et une transformation τ de E dans E, $\mathcal{A}$- mesurable et laissant μ invariante.

On fixe un ensemble C mesurable de mesure $\mu(C) > 0$. Pour $x \in C$, on définit
$t(x) = +\infty$, si $\tau^n x \notin C$, $\forall n \geq 1$, $= \inf\{n \geq 1 : \tau^n x \in C\}$, sinon.
$C_k = \{x \in C : t(x) = k\}$, $k = 1, 2, ...$,
$C_\infty = \{x : t(x) = \infty\}$.

Pour x dans C, $t(x)$ est donc le premier temps de retour (s'il existe) de x dans C.

1) Montrer que

$$C = C_\infty \cup \bigcup_{k \geq 1} C_k,$$

et que les ensembles $\tau^{-j} C_k$, $0 \leq j < k$, $k = 1, 2, ...$, sont deux à deux disjoints.

2) On suppose que $\mu(E) < \infty$.

a) Montrer que $\mu(C_\infty) = 0$, autrement dit, que presque tout point de C revient au moins une fois dans C.

b) Montrer que presque tout point de C revient une infinité de fois dans C.

3) On suppose dans la suite que $\mu(E) = 1$. On définit la transformation τ_C de C dans C (transformation induite) par

$$\tau_C x = \tau^{t(x)} x,\ x \in C.$$

Soit $\mathcal{A}_C$ la tribu des sous-ensembles de C $\mathcal{A}$-mesurables. On note μ_C la mesure normalisée définie par

$$\mu_C(B) = \frac{\mu(B)}{\mu(C)},\ B \in \mathcal{A}_C.$$

a) Montrer que τ_C définit une transformation préservant la mesure sur l'espace probabilisé $(C, \mathcal{A}_C, \mu_C)$.

b) Montrer que, si τ est ergodique, τ_C est aussi ergodique.

c) Montrer qu'inversement, si τ_C est ergodique, la restriction de τ à l'ensemble $\bigcup_{n \geq 0} \tau^{-n} C$ est ergodique.
Peut-on en déduire que τ est ergodique ?

4) On appelle "temps moyen de retour" dans C la quantité définie par

$$\frac{1}{\mu(C)} \int_C t(x) \mu(dx).$$

Montrer que, dans le cas où τ est ergodique, le temps moyen de retour dans C est égal à $\dfrac{1}{\mu(C)}$.

$\boxed{12}$

On considère un système dynamique $(E, \mathcal{A}, \mu\tau)$. Soit T l'opérateur isométrique associé à τ dans $L^2(\mu)$.

En appliquant le théorème de Von Neuman, montrer que pour tout λ de module 1 :

$$\frac{1}{N}\sum_0^{N-1} \lambda^{-n} T^n f \to P_\lambda f,$$

où P_λ est le projecteur orthogonal sur le sous-espace de $L^2(\mu)$ formé des fonctions propres de valeurs propres λ.

Si le spectre de T est continu, on a $P_\lambda = 0, \forall \lambda$. Dans le cas général, on a $P_\lambda = 0$, sauf pour une famille au plus dénombrable de valeurs de λ.

$\boxed{13}$

Soit (E, μ, τ) un système dynamique ergodique, où τ est inversible et μ est une probabilité sur E, τ-invariante.

Soit $\omega \to a(\omega)$ une variable aléatoire positive bornée non concentrée sur $[0, 1]$ et telle que $\int \ln a(\omega)\ d\mu(\omega) < 0$. On définit une suite bilatère de variables aléatoires $(X_n)_{n\geq 0}$ par $X_0 = 0$, et pour $n \geq 0$, $X_{n+1} = a_n X_n + 1$, où $a_n = a \circ \tau^n$.

1) Exprimer X_n et $X_n \circ \tau^{-n}$, $n > 0$, à l'aide des a_k, $k \in \mathbb{Z}$.

Montrer que la suite $(X_n \circ \tau^{-n})_{n\geq 0}$ converge presque partout vers une variable aléatoire notée Z.

Montrer que la suite $(X_n)_{n\geq 0}$ converge en loi.

2) On suppose que les variables aléatoires (a_n) sont indépendantes. Soit α le nombre réel > 0 défini par $\int a^\alpha(\omega)\ dm(\omega) = 1$.

Montrer que l'on a, pour $0 < \beta < \alpha$, $\mathbb{E}(Z^\beta) < \infty$ et pour $\beta > \alpha$, $\mathbb{E}(z^\beta) = \infty$.

3) On suppose E compact et τ continue uniquement ergodique, μ étant l'unique mesure invariante sur E par τ.

a) Montrer (cf. cours) que, pour toute fonction continue f, la suite

$$\left(\frac{1}{n}\sum_0^{n-1} f \circ \tau^k(x)\right)_{n\geq 1}$$

converge vers $\mu(f)$ uniformément en x.

b) Montrer que, si a est continue et strictement positive, Z est continue.

4) On considère le produit $E \times \mathbb{R}$ et la transformation $\widetilde{\tau}$ définie sur $E \times \mathbb{R}$ par $\widetilde{\tau}(\omega, x) = (\tau\omega, ax+1)$. On suppose E compact et τ continue.

En procédant par convergence faible d'une suite convenable de probabilités, montrer l'existence d'une probabilité $\widetilde{\mu}$ sur $E \times \mathbb{R}$, invariante par $\widetilde{\tau}$ et de projection μ sur E.

Montrer que si τ est mélangeante pour μ, $\widetilde{\tau}$ l'est aussi pour $\widetilde{\mu}$.

$\boxed{14}$

I) On note $\Omega = \{0, 1\}^{\mathbb{N}}$ l'ensemble des suites de "0" et de "1" (muni de la topologie produit) et θ la transformation (shift ou décalage) sur Ω définie par : $\theta(\omega)_k = \omega_{k+1}, k = 0, 1, ...$

On appelle **suite admissible** dans Ω les suites ω de 0 et de 1 telles que $\omega_k\omega_{k+1} = 0, \forall k \geq 0$. (Ce sont donc les suites telles qu'un "1" soit toujours suivi d'un "0".) On note Ω_0 l'ensemble des suites admissibles.

1) Montrer que Ω_0 est un fermé de Ω invariant par θ.

2) On note $C_n(0)$ (resp. $C_n(1)$) le nombre de suites finies admissibles de longueur n se terminant par un 0 (resp. par 1), c_n le nombre de suites finies admissibles de longueur n.

Montrer que le vecteur C_n vérifie la relation de récurrence : $C_{n+1} = MC_n$, où M est la matrice

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

et en déduire la valeur de $h = \lim_n \frac{1}{n} \log c_n$.

3) Construire une mesure markovienne m portée par Ω_0 invariante par le shift θ dont l'entropie est h.

II) On note β la racine > 1 de $t^2 - t - 1 = 0$ (le nombre d'or !) et τ la transformation $x \to \beta x \; mod \, 1$ de l'intervalle $E = [0, 1]$ dans lui-même. On désigne par α la partition de E en deux intervalles : $J_0 = [0, \beta^{-1}[$, $J_1 = [\beta^{-1}, 1[$.

1) Vérifier que la transformation τ possède une mesure μ invariante équivalente à la mesure uniforme λ, dont la densité γ est donnée par :

$$\gamma(x) = \frac{5 + 3\sqrt{5}}{10}, \text{ pour } 0 \leq x < \beta^{-1}, \text{ et } \frac{5 + \sqrt{5}}{10}, \text{ pour } \beta^{-1} < x < 1.$$

2) On considère le codage associé à τ et la partition α, c'est à dire l'application $\phi : E \to \Omega, x \to \phi(x) = (\phi_k(x))_{k \geq 0}$, avec

$$\tau^k x \in J_{\phi_k(x)}, k \geq 0.$$

a) Montrer que la suite $\phi(x)$ associée à tout point x est une suite admissible, au sens de la définition précédente. (On montrera que $\tau J_0 = E$, $\tau J_1 = J_0$.)

b) Inversement, à toute suite admissible ω, on associe le nombre réel

$$\psi(\omega) = \sum_0^\infty \beta^{-(k+1)} \omega_k.$$

Montrer que $0 \leq \psi(\omega) \leq 1$, et que ψ est l'application inverse de ϕ.

c) Montrer que α est une partition génératrice pour la transformation τ.

3) a) Montrer que les sytèmes (E, μ, τ) et (Ω_0, m, θ) sont conjugués. (On admettra que μ est la mesure image de m par ψ.)

b) Quelle est l'entropie du sytème dynamique (E, μ, τ) ?

c) Montrer que ce système est ergodique et à spectre continu.

En procédant par convergence faible d'une suite convenable de probabilités, montrer l'existence d'une probabilité $\widetilde{\mu}$ sur $E \times \mathbb{R}$, invariante par $\widetilde{\tau}$ et de projection μ sur E.

Exercice I

On considère un système dynamique ergodique $(E, \mathcal{A}, \mu, \tau)$, où τ est inversible et μ est une probabilité sur E invariante par τ.

Soit a une fonction > 0 sur Ω telle que

$$\int \ln a(x)\ d\mu(x) < 0.$$

On définit une suite de fonctions $(\phi_n)_{n\geq 0}$ par récurrence :

$$\phi_0 = 1, \text{ et pour } n \geq 0,\ \phi_{n+1}(x) = a(\tau^n x)\phi_n(x) + 1, x \in E.$$

I) 1) Exprimer ϕ_n et $\phi_n \circ \tau^{-n}$, $n > 0$, à l'aide de a et des itérées de τ.

2) Montrer que la suite $(\phi_n \circ \tau^{-n})_{n\geq 0}$ converge presque partout vers une limite ψ.

3) Montrer que la suite $(\phi_n)_{n\geq 0}$ converge en loi.

II) On suppose que E est compact et τ continue, et que τ possède une unique mesure invariante, la mesure μ, sur E.

1) a) Donner un exemple d'une telle situation (une transformation sur un compact possédant une unique mesure invariante).

b) Montrer que, pour toute fonction continue f sur E, la suite

$$(\frac{1}{n}\sum_0^{n-1} f(\tau^k(x))_{n\geq 1}$$

converge vers $\mu(f)$ uniformément en x.

c) Montrer que, si a est continue, la fonction ψ construite en I) est continue.

2) On considère la transformation $\widetilde{\tau}$ définie sur le produit $E \times \mathbb{R}$ par

$$\widetilde{\tau}(x, y) = (\tau x, a(x)y + 1).$$

a) Expliciter les itérées de la transformation $\widetilde{\tau}$.

b) Montrer que l'équation fonctionnelle

$$Z(\tau x) = a(x)Z(x) + 1$$

possède une solution.

Montrer que cette solution est unique.

c) A l'aide du I), mettre en évidence une mesure sur $E \times \mathbb{R}$ invariante par $\widetilde{\tau}$, dont la projection sur E est μ.

d) Il y a-t-il d'autres mesures de probabilité sur $E \times \mathbb{R}$ invariantes par $\widetilde{\tau}$?

e) Décrire le comportement des itérées de la transformation $\widetilde{\tau}$.

$\boxed{3}$

Soit (E, τ) un système dynamique topologique, i.e. la donnée d'un espace métrique compact (E, d) et d'une transformation continue τ de E dans lui-même.

On définit la notion de produit de deux sytèmes dynamiques, et de facteur d'un système dynamique. Si un système (F, σ) est facteur à la fois de (E_1, τ_1) et de (E_2, τ_2), on dit que c'est un facteur commun aux deux sytèmes.

Montrer que si un système est un facteur commun de deux systèmes (E_1, τ_1) et de (E_2, τ_2), qui n'ont pas de valeurs propres communes, alors il est à spectre continu.

$\boxed{14}$

Codage et entropie

a) Montrer qu'une transformation distale possède une partition génératrice au sens fort.

b) En déduire que, pour toute mesure invariante, l'entropie d'une transformation distale est nulle.

****

Problème 1

On considère la transformation $\tau : x \to 2x \bmod 1$ sur le cercle $E = \mathbb{T}^1$ identifié à $\mathbb{R}/\mathbb{Z}$ muni de la tribu borélienne et de la mesure de Lebesgue notée λ. Les exponentielles $x \to e^{2\pi i p x}$, pour $p \in \mathbb{Z}$, sont notées e_p.

Dans la suite, on désigne par α un nombre réel fixé, par $a(x)$ la fonction définie par

$$a(x) = \begin{cases} 0, \text{ pour } x \in [0, \frac{1}{2}[, \\ 1, \text{ pour } x \in [\frac{1}{2}, 1[. \end{cases}$$

On rappelle que l'on a $x = \sum_{k \geq 0} a(\tau^k x) 2^{-(k+1)}$, pour tout $x \in (0, 1[$, et que, sous la probabilité λ, les variables aléatoires $\{a \circ \tau^k, k \geq 0\}$ sont indépendantes. Pour $n \geq 1$, on désigne par $S_n(x)$ la somme des n premiers chiffres (ici n premiers bits) du développement de $x \in [0, 1[$ en base 2. On a donc $S_n(x) = \sum_{k=0}^{n-1} a(\tau^k x)$. On pose $S_0(x) = 0$.

Pour toute fonction borélienne ϕ, on construit une extension du système (E, λ, τ) en définissant sur le produit $E \times E$, muni de la mesure produit $\lambda \times \lambda$, la transformation τ_ϕ par :

$$\tau_\phi : (x, y) \to (\tau x, y + \phi(x) \bmod 1).$$

I.

a) En utilisant l'analyse de Fourier, donner à l'aide d'une famille d'équations fonctionnelles une condition nécessaire et suffisante sur ϕ pour que le système $(E \times E, \lambda \times \lambda, \tau_\phi)$ soit ergodique,

On prend maintenant pour ϕ la fonction $\phi_\alpha : x \to \alpha a(x)$. On note simplement τ_α la transformation τ_{ϕ_α}. On a donc :

$$\tau_\alpha(x, y) = (2x \bmod 1, y + \alpha a(x) \bmod 1).$$

b) Montrer l'ergodicité du système $(E \times E, \lambda \times \lambda, \tau_\alpha)$, pour α irrationnel.

c) Déterminer la tribu des invariants dans le cas où α est un rationnel, $\alpha = \frac{p_0}{q_0}$, ($\frac{p_0}{q_0}$ irréductible).

d) Montrer que, pour tout α irrationnel, la suite $(\alpha S_n(x) \bmod 1)_{n \geq 0}$ est équirépartie, pour presque tout x, i.e. vérifie pour tout intervalle $[a, b[\subset [0, 1[$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} 1_{[a,b[}(\alpha S_n(x) \bmod 1) = b - a, \text{ pour presque tout } x.$$

II.

Sur $E = \mathbb{T}^1$, on note $\mathcal{F}_0$ la tribu triviale et, pour $n \geq 1$, $\mathcal{F}_n$ la tribu engendrée par les intervalles dyadiques d'ordre n. On note $(\mathcal{F}_n)_{n\geq 0}$ la filtration croissante formée par les tribus $\mathcal{F}_n$.

a) Montrer, pour tout $p \in \mathbb{Z}$ et tout $n \geq 1$, la relation :

$$\mathbb{E}(e_p(\alpha S_n)|\mathcal{F}_{n-1}) = \gamma_p e_p(\alpha S_{n-1}),$$

où $\gamma_p = e^{i\pi p\alpha}\cos(\pi p\alpha)$.

b) Soit f une fonction continue sur le cercle d'intégrale nulle, de série de Fourier $f = \sum_{p\in\mathbb{Z}} c_p e_p$, telle que la série $\sum_{p\in\mathbb{Z}} |\frac{c_p}{1-\gamma_p}|$ converge.

On note $M_0 = \sum_{p\in\mathbb{Z}} \frac{c_p}{1-\gamma_p}$ et, pour $n \geq 1$,

$$M_n = \sum_{k=0}^{n-1} f(\alpha S_k) + \sum_{p\in\mathbb{Z}} \frac{c_p}{1-\gamma_p} e_p(\alpha S_n).$$

Montrer que le processus $(M_n)_{n\geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n\geq 0}$.

c) Montrer qu'il existe un entier $r \geq 1$ tel que, pour f de classe C^r d'intégrale nulle sur le cercle et pour presque tout α, la suite $(f(\alpha S_n))_{n\geq 0}$ vérifie le théorème de la limite centrale.

(On utilisera le fait que, pour tout $\epsilon > 0$, pour presque tout α réel, il existe une constante $C = C(\alpha)$ telle que $|\alpha - \frac{p}{q}| \geq \frac{C}{q^{2+\epsilon}}$, pour toute fraction irréductible $\frac{p}{q}$).

Problème 2

Soient $(E, \mathcal{E}, \mu)$ un espace probabilisé et, pour un entier $d \geq 1$, V une application mesurable de $(E, \mathcal{E})$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. On désigne par $(\Omega, \mathcal{F}, \mathbb{P})$ l'espace produit $\otimes_{\mathbb{N}^*}(E, \mathcal{E}, \mu)$, par $(X_k)_{k\geq 1}$ les applications coordonnées de l'espace produit $\Omega = E^{\mathbb{N}^*}$ et par $\theta : (\omega_1, \omega_2, \cdots) \to (\omega_2, \omega_3, \cdots)$ le décalage sur Ω. On note ν l'image de μ par V.

Soit $\widetilde{\theta}$ la transformation de $\Omega \times \mathbb{R}^d$ définie par

$$(\omega, x) \in \Omega \times \mathbb{R}^d \to (\theta\omega, V(X_1(\omega)) + x).$$

1) Si λ désigne la mesure de Lebesgue sur $\mathbb{R}^d$, montrer que la mesure positive de masse infinie $\mathbb{P} \times \lambda$ est $\widetilde{\theta}$-invariante.

2) Soit F une fonction mesurable bornée de $(\Omega \times \mathbb{R}^d, \mathcal{F} \times \mathcal{B}(\mathbb{R}^d))$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ invariante par $\widetilde{\theta}$. On note h_F la fonction définie par

$$h_F(x) = \int_\Omega F(\omega, x)\ \mathbb{P}(d\omega), x \in \mathbb{R}^d.$$

a) Montrer que h_F est ν-harmonique (i.e. vérifie $h_F(x) = \displaystyle\int_{\mathbb{R}^d} h_F(x + y)\ \nu(dy)$).

b) Montrer que l'on a, $\mathbb{P}$-p.s.,

$$\mathbb{E}[F(., x)|\sigma(X_1, \cdots, X_n)] = h_F(x + V(X_1) + \cdots + V(X_n)),\ \forall n \geq 1.$$

3) On admettra le résultat suivant : *Une fonction continue h sur $\mathbb{R}^d$ est ν-harmonique si, et seulement si, elle admet pour période tout élément du support de la mesure ν.*

a) Donner une condition nécessaire et suffisante pour que toute fonction continue ν-harmonique soit constante.

b) Sous cette condition, montrer que toute fonction borélienne sur $\mathbb{R}^d$ ν-harmonique est λ-p.s. constante. (On utilisera une méthode de régularisation par une identité approchée).

c) Sous la condition du a), montrer que la tribu des évènements $\widetilde{\theta}$-invariants est triviale.

Exercice

I

Soit I un "alphabet" fini. Pour $a, b \in I$, on pose $\delta(a,b) = 0$, si $a = b$, $= 1$, si $a \neq b$. On considère l'espace produit $\Omega = I^{\mathbb{N}}$, muni de la tribu borélienne produit $\mathcal{F}$ et de la distance d définie par $d(x,y) = \sum_{n \geq 0} 2^{-n}\delta(x_n, y_n)$, pour $x = (x_n)_{n \geq 0}$ et $y = (y_n)_{n \geq 0}$.

On note $\mathcal{C}(\Omega)$ l'espace des fonctions continues sur l'espace métrique (Ω, d). Pour un réel $\alpha \in]0,1]$, une fonction f définie sur Ω est dite α-höldérienne, s'il existe une constante K telle que $|f(x) - f(y)| \leq Kd(x,y)^\alpha$, $\forall x, y \in \Omega$. On note $\mathcal{H}_\alpha(\Omega)$ l'espace des fonctions α-höldériennes sur Ω.

Pour $n \geq 0$, on note X_n les applications coordonnées, θ le décalage, T l'opérateur de composition par θ : $Tf(x) = f(\theta x)$.

1) Soit μ une mesure θ-invariante sur $(\Omega, \mathcal{F})$. On note $\mathbb{E}_\mu$ l'espérance par rapport à μ, P_μ l'opérateur dual de l'opérateur T opérant sur $L^2(\mu)$.

Montrer que, pour toute fonction f dans $L^1(\mu)$, $TP_\mu f$ est une version de $\mathbb{E}_\mu[f|\theta^{-1}\mathcal{F}]$.

2) On se donne une matrice M indexée par $I \times I$ et un vecteur de probabilité π sur I. On suppose M stochastique (i.e. vérifiant : $M(a,b) \geq 0$, $\forall a, b \in I$ et $\sum_{b \in I} M(a,b) = 1$, $\forall a \in I$). On construit la mesure $\mathbb{P}_\pi$ sur Ω telle que

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

a) Montrer que $\mathbb{P}_\pi$ est θ-invariante si, et seulement si, π est un vecteur invariant par la matrice ${}^tM$ transposée de M.

On suppose dans toute la suite que le vecteur π est un vecteur invariant par la matrice ${}^tM$.

b) Calculer l'opérateur $P_{\mathbb{P}_\pi}$, noté plus simplement P, dual de l'opérateur T opérant sur $L^2(\mathbb{P}_\pi)$. [On pourra utiliser la représentation d'une fonction f quelconque sur Ω sous la forme :

$$f(x) = \sum_{a \in I} 1_{\{X_0 = a\}}(x) f(a\theta x).]$$

c) Montrer que P défini un opérateur markovien sur l'espace $\mathcal{C}(\Omega)$ des fonctions continues sur Ω.

3) Pour $f \in \mathcal{H}_\alpha(\Omega)$ on pose

$$m(f) = \sup\{\frac{|f(x) - f(y)|}{d(x,y)^\alpha},\ x, y \in \Omega, x \neq y\}.$$

a) Montrer que, pour toute fonction $f \in \mathcal{H}_\alpha(\Omega)$, on a :

$$m(Pf) \leq \rho\, m(f) + C\|f\|_\infty,$$

où ρ et C sont des constantes positives, avec $0 < \rho < 1$.

b) Montrer que, pour toute fonction $f \in \mathcal{H}_\alpha(\Omega)$, la famille $\{P^n(f), n \geq 0\}$ est équicontinue.

c) En déduire la convergence uniforme de la suite $(P^n(f))_{n\geq 0}$ vers $\int_\Omega f(x)\, d\mathbb{P}(x)$, pour toute fonction $f \in \mathcal{C}(\Omega)$.

d) Montrer que le système dynamique $(\Omega, \mathcal{F}, \mathbb{P}_\pi, \theta)$ vérifie la propriété suivante :

$$\lim_n \int_\Omega T^n f(x)\; g(x)\; d\mathbb{P}_\pi(x) = (\int_\Omega f(x)\; d\mathbb{P}_\pi(x))\; (\int_\Omega g(x)\; d\mathbb{P}_\pi(x))$$

et en déduire qu'il est ergodique.

4) Si f est une fonction à valeurs réelles définie sur I telle que $\sum_{a\in I} f(a)\pi(a) = 0$, établir un théorème de la limite centrale pour la suite de v.a. $(f(X_n))_{n\geq 0}$.

II

On considère maintenant une matrice A indexée par $I \times I$, à coefficients > 0. On peut regarder A comme une fonction dépendant des deux premières coordonnées, définie sur Ω par $x = (x_n)_{n\geq 0} \to A(x) = A(x_0, x_1)$.

Si μ est une mesure θ-invariante sur $(\Omega, \mathcal{F})$, on note $h_\mu(\theta)$ l'entropie du système $(\Omega, \mathcal{F}, \mu, \theta)$.

On admettra qu'elle est donnée par la partition $\mathcal{P}$ de Ω définie par $\mathcal{P} = \{(X_0 = a), a \in I\}$ et on utilisera pour son calcul la formule

$$h(\mathcal{P}, \theta) = H(\mathcal{P}|\theta^{-1}\mathcal{F}),$$

où $H(\mathcal{P}|\theta^{-1}\mathcal{F})$ est l'entropie conditionnelle de la partition $\mathcal{P}$ par rapport à la sous-tribu $\theta^{-1}\mathcal{F}$.

1) Montrer que A possède un unique vecteur de probabilité v propre (i.e. tel que $Av = \lambda v$). On note $\lambda > 0$ la valeur propre correspondante. Montrer que ${}^tA$, sa transposée, possède également un unique vecteur de probabilité propre noté w et que ${}^tAw = \lambda w$. Montrer que les coordonnées de v et de w sont > 0.

2) On note M la matrice indexée par $I \times I$ définie par

$$M(a, b) = \frac{v(b)}{\lambda v(a)} A(a, b),\;\; a, b \in I.$$

a) Montrer que M est stochastique.

b) Déterminer un vecteur de probabilité π invariant par la transposée de M. On note $\mathbb{P}_\pi$ la mesure de probabilité associée à π et M construite comme en II.2.

c) Montrer la relation

$$\int_\Omega \ln A(x)\; d\mathbb{P}_\pi(x) + h_{\mathbb{P}_\pi}(\theta) = \log \lambda.$$

===========

Problème 1

Soit μ une mesure de probabilité sur les boréliens de $\mathbb{R}$. On appelle $\mathcal{H}_\mu$ l'espace des fonctions $f : \mathbb{R} \to \mathbb{R}$ boréliennes bornées, solutions de l'équation fonctionnelle

$$\forall x \in \mathbb{R},\;\; f(x) = \int_\mathbb{R} f(x + t)\; d\mu(t).$$

On note λ la mesure de Lebesgue sur les boréliens de $\mathbb{R}$. On désigne par $(Y_k)_{k\geq 1}$ une suite de v.a.r. indépendantes, de loi commune μ, définies sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

Pour tout réel x et tout $\omega \in \Omega$, on pose :

$$S_0(x, \omega) = x \text{ et } \forall n \geq 1, \; S_n(x, \omega) = x + Y_1(\omega) + \dots Y_n(\omega).$$

On appelle $\mathcal{F}_0$ la tribu triviale $\{\emptyset, \Omega\}$ et, pour tout $n \geq 1$, $\mathcal{F}_n$ la tribu engendrée par les v.a.r. $\{Y_k : 1 \leq k \leq n\}$.

Soit h un élément de $\mathcal{H}_\mu$.

1) Pour tout réel x, montrer que le processus $(M_n(x, \omega) = h(S_n(x, \omega)))_{n \geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_n)_{n \geq 0}$. En déduire que ce processus converge $\mathbb{P}-$p.s. vers une v.a.r. $Z(x, \cdot)$ vérifiant $h(x) = \int_\Omega Z(x, \omega) \, d\mathbb{P}(\omega)$.

2) Pour tout réel x, on note

$$V_n(x, \omega) = \sum_{k=1}^{n} \mathbb{E}\Big[(M_k(x, \cdot) - M_{k-1}(x, \cdot))^2 | \mathcal{F}_{k-1}\Big](\omega) \qquad (n \geq 1)$$

le processus croissant associé à la martingale de carré intégrable $(M_n(x, \cdot))_{n \geq 0}$.

a) Montrer que, pour tout réel x, $\lim\limits_{n \uparrow +\infty} \uparrow \mathbb{E}[V_n(x, \omega)] < +\infty$.

b) En déduire que, pour tout réel x,

$$\int_{\mathbb{R}} [h(t + S_n(x, .)) - h(S_n(x, .))]^2 \, d\mu(dt) \xrightarrow[n \to +\infty]{\mathbb{P}-\text{p.s.}} 0,$$

puis que $Z(t + x, \omega) = Z(x, \omega)$, pour $\mu \times \mathbb{P}$-presque tout $(t, \omega) \in \mathbb{R} \times \Omega$.

3) a) Montrer que, pour tout $x \in \mathbb{R}$, on a $h(x + t) = h(x)$, pour μ-presque tout t.

b) Donner une condition nécessaire et suffisante pour que les fonctions **continues** de $\mathcal{H}_\mu$ soient constantes.

c) Lorsque cette condition nécessaire et suffisante est satisfaite, montrer que les fonctions de $\mathcal{H}_\mu$ sont λ-presque sûrement constantes.

4) On choisit $\mu = \frac{1}{2}(\delta_1 + \delta_{\sqrt{2}})$ et on note H le sous-groupe $\mathbb{Q} + \mathbb{Q}\sqrt{2}$.

a) Déterminer les fonctions continues de $\mathcal{H}_\mu$.

b) Soit B un borélien de $\mathbb{R}$. Que peut-on dire du borélien BH de $\mathbb{R}$?

Problème 2

On considère le transformation $\tau : x \to 2x \bmod 1$ sur l'espace $E = \mathbb{R}/\mathbb{Z}$ muni de la tribu borélienne et de la mesure de Lebesgue notée λ.

Etant donnée une fonction ϕ continue 1-périodique, on construit une extension du système (E, λ, τ) en définissant sur le produit $E \times E$, muni de la mesure produit $\lambda \times \lambda$, la transformation τ_ϕ par :

$$\tau_\phi : (x, y) \to (\tau x, y + \phi(x) \bmod 1).$$

1) a) En utilisant l'analyse de Fourier, donner une condition nécessaire et suffisante sur ϕ pour que le système $(E \times E, \lambda \times \lambda, \tau_\phi)$ soit ergodique.

On exprimera cette condition à l'aide d'une famille d'équations fonctionnelles notées $(*)$.

b) Montrer que l'ergodicité de τ_ϕ pour la mesure $\lambda \times \lambda$ entraîne l'unicité de la mesure invariante par τ_ϕ qui se projette sur λ.

2) On note ψ la fonction $\psi(x) = \exp(2\pi i\phi(x))$ et P_ψ l'opérateur défini par $P_\psi f = P(\psi f)$, $f \in L^1(\lambda)$, où P est l'opérateur adjoint de l'opérateur de composition par τ dans $L^2(\lambda)$.

a) Montrer que P_ψ est une contraction pour les normes L^1 et L^∞.

On suppose maintenant dans toute la suite que ϕ est lipschitzienne.

b) Pour f lipschitzienne, montrer que la suite des itérées $(P_\psi^n f)_{n\geq 0}$ est équicontinue.

c) En déduire que toute solution de l'une des équations $(*)$ est presque partout égale à une fonction lipschitzienne.

3) a) Etablir un critère d'ergodicité du système $(E \times E, \lambda \times \lambda, \tau_\phi)$ à l'aide d'une condition sur les points périodiques de τ.

b) A-t-on l'ergodicité de $(E \times E, \lambda \times \lambda, \tau_\phi)$ quand on prend $\phi(x) = \cos(2\pi x)$?

****

Les deux exercices dont indépendants. La question d) du premier exercice est plus difficile que les précédentes.

Exercice 1 On considère le système dynamique $(\mathbb{T}^1, \mathcal{B}, \tau, \lambda)$ défini par la transformation $x \to \tau x = 2x \ mod\ 1$ sur le cercle unité $\mathbb{T}^1$ identifié à $\mathbb{R}/\mathbb{Z}$, muni de sa tribu borélienne $\mathcal{B}$ et de la mesure de Lebesgue λ.

Soit φ une application mesurable de $\mathbb{T}^1$ dans $\mathbb{T}^1$. On définit sur $\mathbb{T}^1 \times \mathbb{T}^1$ muni de la mesure produit $\lambda \times \lambda$ une transformation en posant

$$\sigma_\varphi(x, y) = (\tau x, y + \varphi(x) \ mod\ 1).$$

a) Montrer que σ_φ laisse invariante la mesure $\lambda \times \lambda$ sur

$\mathbb{T}^1 \times \mathbb{T}^1$.

b) Soit f une fonction dans $L^1(\lambda \times \lambda)$ non presque partout égale à une constante. Pour $n \in \mathbb{Z}$, on note

$$\widehat{f_n}(x) = \int f(x, y)\exp(-2\pi i n y)dy.$$

Montrer que, si f est invariante par σ, il existe $n \in \mathbb{Z} - \{0\}$ tel que

$$\widehat{f_n}(\tau x) = \exp(2\pi i n\varphi(x))\ \widehat{f_n}(x),$$

avec $\widehat{f_n}$ non presque partout nulle et qu'une telle fonction $\widehat{f_n}$ est de module constant (presque partout) non nul.

c) On prend $\varphi(x) = x$. Montrer que σ_φ est ergodique.

d) (Question facultative) Pour $\epsilon \in]0, 1[$, on note I_ϵ un intervalle fermé du cercle de longueur ϵ et I_ϵ^c son complémentaire. Soit $\mathcal{P}_\epsilon$ la partition de $\mathbb{T}^1 \times \mathbb{T}^1$ en les sous-ensembles $([0, \frac{1}{2}[\times I_\epsilon, [0, \frac{1}{2}[\times I_\epsilon^c, [\frac{1}{2}, 1[\times I_\epsilon, [\frac{1}{2}, 1[\times I_\epsilon^c$.

Montrer que $\mathcal{P}_\epsilon$ est génératrice pour la transformation σ.

Quelle est l'entropie du système $(\mathbb{T}^1 \times \mathbb{T}^1, \lambda \times \lambda, \sigma)$?

Exercice 2

On considère un système dynamique mesuré $(E, \mathcal{E}, \tau, \mu)$. On note T l'opérateur de composition par τ, P l'opérateur adjoint de T dans $L^2(\mu)$, I l'opérateur identité.

1) a) Montrer que $\int PTf\,\overline{g}\;d\mu = \int Tf\,T\overline{g}\;d\mu = \int f\,\overline{g}\;d\mu$, pour $f,g \in L^2(\mu)$, puis que $PT = I$.

Montrer que T^nP^n, pour $n \geq 1$ est l'opérateur d'espérance conditionnelle par rapport à la tribu $\tau^{-n}\mathcal{E}$:

$$TPf = \mathbb{E}(f|\tau^{-1}\mathcal{E}).$$

b) On note $\mathcal{E}_\infty$ la tribu $\bigcap_n \tau^{-n}\mathcal{E}$. Montrer que

$$\lim_n \|T^nP^nf - \mathbb{E}(f|\mathcal{E}_\infty)\|_2 = 0, \forall f \in L^2(\mu).$$

2) (On note $\{y\}$ la partie fractionnaire d'un réel y.) On prend comme système dynamique la transformation de Gauss $\tau x = \{\frac{1}{x}\}$ sur l'intervalle $]0,1[$ muni de sa tribu borélienne notée $\mathcal{B}$ et de la mesure τ-invariante μ absolument continue par rapport à la mesure de Lebesgue, de densité $\dfrac{1}{\ln 2}\dfrac{1}{1+x}$.

On rappelle que l'opérateur P adjoint dans $L^2(\mu)$ de l'opérateur T de composition par τ a l'expression suivante :

$$Pf(x) = \sum_{k=1}^{\infty} f(\frac{1}{x+k})\frac{1+x}{(x+k)(1+x+k)}.$$

Cet opérateur laisse invariante la mesure μ, envoie l'espace des fonctions continues sur $[0,1]$ dans lui-même et vérifie (rappel) : pour toute fonction f lipschitzienne sur $[0,1]$, la famille $(P^nf)_{n\geq 0}$ est équicontinue.

Soit f une fonction lipschitzienne sur $[0,1]$. On se propose de montrer que

$$\lim_n \|P^nf - \mu(f)\|_\infty = 0. \tag{13.1}$$

a) Soit h une valeur d'adhérence de la suite (P^nf) pour la norme uniforme. Montrer que $T^rP^rh = \mathbb{E}(h|\tau^{-r}\mathcal{E}) = h$, pour tout entier $r \geq 0$

b) On note A l'ensemble sur lequel h atteint son maximum. En utilisant la propriété de barycentre de l'opérateur TP, montrer que A est invariant par la transformation $x \to \frac{1}{\tau x + k}$, pour tout entier $k \geq 1$.

c) Montrer que h est constante (on pourra également considérer l'ensemble où h atteint son minimum).

d) Montrer (13.1), pour toute fonction f lipschitzienne sur $[0,1]$.

e) Montrer que le sytème défini par la transformation de Gauss est mélangeant et que la tribu $\bigcap_n \tau^{-n}\mathcal{E}$ est la tribu triviale (réduite à $(\mathcal{E},\emptyset)$), aux ensembles négligeables près).

Bibliographie

[Bala00] Baladi (V.) : *Positive Transfer Operators and Decay of Correlations*, Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific (2000).

[Bil99] Billingsley (P.) : *Convergence of probability measures*, 2nd ed. (1999), Wiley, New York.

[Bow75] Bowen (R.) : *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math. 470 (1975) Springer Verlag.

[Bre68] Breiman (L.) : *Probability* (1968), Addison-Wesley.

[Brow71] Brown (B. M.) : Martingale central limit theorems (1971), The Annals of Math. Stat. vol. 42, no 1, p. 59-66.

[DeGrSy76] Denker (M.), Grillenberger (C.), Sygmund (K.) : *Ergodic theory on compact spaces*, Lecture Notes in Math. 527 (1976) Springer Verlag.

[Denk89] Denker (M.). - The central limit theorem for dynamical systems, in Dynamical systems and ergodic theory, Banach center publications, volume 23, PWN-Polish Scientific Publishers, Warszawa, 1989.

[DerLin96] Derriennic (Y.), Lin (M.), Sur le théorème limite central de Kipnis et Varadhan pour les chaînes réversibles ou normales, C. R. Acad. Sci. Paris, t.323, Série 1, 1053-1057 (1996).

[Doo53] Doob (J. L.) : *Stochastic processes* (1953), Wiley, New York.

[Douk94] Doukhan (P.) : *Mixing, properties and examples* (1994), Lecture Notes in Statistics 85, Springer-Verlag.

[DunSch58] Dunford (N.), Schwartz (J.) : *Linear Operators* (Vol. I) 1958, Wiley-Interscience, New York.

[Gar70] Garsia (A.) : *Topics in almost everywhere convergence*, (1970) Lectures in adanced mathematics, Markham Publishing Company, Chicago.

[Gor69] Gordin (M. I.) : On the central limit theorem for stationary processes, (Dokl. Akad. Nauk SSSR), Soviet Math. Dokl., 10, no 5, 1174-1176, 1969.

[GoLi61] Gordin (M. I.), Lifshitz (B. A.) : The central limit theorem for stationary markov processes, Soviet Math. Dokl., vol. 19, no 2, 392-394, 1978.

[GuiHar88] Guivarc'h (Y.), J. Hardy (J.) : Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov, Ann. I. H. P., sér. Proba. et Stat., 24, 73-98, 1988.

[HaHe80] Hall (P.), Heyde (C.C.) : *Martingale limit theory and its applications* (1980), Academic Press, New York.

[HenHer01] Hennion (H.), Hervé (L.) : *Limit theorem for Markov Chains and Stochastic Properties of Dynamical Systems by Quasi-compactness*, Springer Lectures Notes, no 1766 (2001).

[Hop61] Hopf (E.) : On the ergodic theorem for positive linear operators, J. Math. Mech., 205, 101-106 (1961).

[Ibrag63] Ibragimov (I. A.) : A central limit theorem for a class of dependant random variables, Theory Prob. Appl., 1963, vol. 8, p. 83-89.

[IbLi71] Ibragimov (I. A.), Y. V. Linnik (Y. V.) : *Independant and stationary sequences of random variables* (1971), Wolters-Noordhoff, Groningen.

[Jam64] Jamison (J.), Asymptotic behavior of successive iterates of continuous functions under Markov operator, J. Math. Anal. Appl., 9, 1964, p. 203-214.

[Kel98] Keller (G.) : *Equilibium States in ergodic theory* (1998), London Mathematical Society, Student Texts 42 ;

[Khi] Khintchin (A. Ya.) : *Continued Fractions*, republication Dover (1997).

[King68] Kingman (J.F.C.) : The ergodic theory of subadditive processes, J. Royal Statist. Soc. (1968), 30, p. 499-510.

[Kre85] Krengel (U.) : *Ergodic theorems* (1985), De Gruyter Studies in Mathematics, volume 6, Berlin.

[Nev64] Neveu (J.) : *Bases Mathématiques du Calcul des Probabilités* (1964), Masson.

[Par81] Parry (W.) : *Topics in ergodic theory* (1981), Cambridge.

[Rau92] Raugi (A.) : Théorie spectrale d'un opérateur de transition sur un espace métrique compact, Ann. Inst. Henri Poincaré, 28, 2, 1992, p. 281-309.

[Rev75] Revuz (D.) : *Markov chains* (1975), North Holland Pub. Comp..

[Rokh67] Rokhlin (V.) : Lectures on the entropy theory, Russian Math. Surveys, (1967) vol. 22, no 5.

[Rud67] Rudin (W.) : *Fourier analysis on groups*, (1967) Interscience pub., J. Wiley, NY.

Index