RÉPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITÉ MOULOUD MAMMERI, TIZI-OUZOU
FACULTÉ DE SCIENCES
DÉPARTEMENT DE MATHÉMATIQUES

THÈSE DE DOCTORAT ES-SCIENCES

SPÉCIALITE : MATHÉMATIQUES
OPTION : PROBABILITÉS ET STATISTIQUE

Présentée par :
Mme LADJIMI Fetima épouse BOUDJEDAIMI

Sujet :

Théories de L'apprentissage, Fonctions Aléatoires, Processus et Chaînes de Markov

Devant le jury d'examen composé de :

M. MORSLI Mohamed ;	Professeur ;	U.M.M.T.O ;	Président
M. BOUDIBA Mohand Arezki ;	M. de Conférences A ;	U.M.M.T.O ;	Rapporteur
M. HAMADOUCHE Djamel ;	Professeur ;	U.M.M.T.O ;	Examinateur
M. KESSI Arezki ;	Professeur ;	U.S.T.H.B ;	Examinateur
M. YOUSFATE Abderrahmane ;	Professeur ;	U.Sidi-Bel-Abbès ;	Examinateur
M. ZOUGAB Nabil ;	M. de Conférences A ;	U.Béjaia ;	Examinateur
M. PEIGNE Marc ;	Professeur ;	U.Tours France ;	Invité

Soutenue le : 15/09/2020

Remerciements

Ma reconnaissance va tout d'abord à mon directeur de thèse Monsieur Mohand Arezki BOUDIBA, Maître de Conférences à l'U.M.M.T.O, pour m'avoir confié ce travail. Je le remercie pour sa disponibilité, ses remarques judicieuses et ses relectures attentives dont il sait me faire profiter, pour ses conseils, ses encouragements et ses critiques, pour sa confiance en moi et sa générosité.

Je tiens tout particulièrement à exprimer ma profonde gratitude et toute ma reconnaissance à Monsieur PEIGNE Marc, Professeur à l'université François Rabelais de Tours. Je le remercie de m'avoir accueillie et offert l'opportunité de travailler avec lui au sein de laboratoire LMPT durant mes stages de perfectionnement. Il m'a guidée avec autant de disponibilité que d'attention et sans qui cette thèse et le travail qu'elle représente n'auraient pas vu le jour. Avec ses grandes qualités scientifiques, il m'a initiée au monde des opérateurs quasi-compacts ; j'ai énormément appris à son contact. Je le remercie pour toute l'énergie et tout le temps qu'il m'a consacré.

Je remercie l'ensemble des membres du laboratoire LMPT, pour les conditions de travail dont j'ai bénéficié. En particulier l'équipe de Probabilités et Statistique.

Je suis très honorée que Monsieur MORSLI Mohamed, Professeur à l'U.M.M.T.O, ait accepté de présider ce jury. Je le remercie pour l'intérêt qu'il porte à ce travail.

Je remercie Monsieur HAMADOUCHE Djamel Professeur à l'U.M.M.T.O, Monsieur KESSI Arezki Professeur à l'U.S.T.H.B, Monsieur YOUSFATE Abderrahmane Professeur à l'université de Sidi Bel-Abbès et Monsieur ZOUGAB Nabil Maître de Conférences à l'université de Béjaia, pour l'intérêt qu'ils accordent à ce travail et l'honneur qu'ils me font en acceptant de faire partie de ce jury.

J'adresse mes remerciements à tous ceux qui ont contribué de près ou de loin à ma formation, en particulier mes enseignants depuis la première année primaire.

Je tiens à remercier tous les collègues et amies, qui ont répondu avec calme et patience aux questions quotidiennes dont je les accablais. Je remercie particulièrement Madame HARMIM Dehbia pour ses conseils et encouragements, ainsi que Madame ZIDI Yamina pour ses remarques en analyse fonctionnelle.

Enfin, mes pensées vont à mes parents, mon époux et toute ma famille pour leur soutien et leurs encouragements. Je leur dédie cette thèse à laquelle ils ont participé plus qu'ils ne le pensent.

A la mémoire de mon frère MOULOUD,

de mes Beaux Parents et de mon amie BESSAD Baya.

Liste des Publications et Communications

Communications

1. B. Bessad, F. Ladjimi, M. A. Boudiba.
   Itération de Fonctions aléatoires et application à la simulation.
   Séminaire Mathématique de Béjaia Séance du Mardi 07 Mai 2013.

2. F. Ladjimi, M. Peigné
   On the Diaconis- Freedman's Chain on $[0, 1]$ *with place dependent transition probabilities.*
   Congrès des Mathématiciens Algériens CMA'2018, Boumerdes, 12-13 mai 2018

Publications

1. F. Ladjimi, M. Peigné
   On the asymptotic behavior of the Diaconis-Freedman chain on $[0, 1]$.
   Statistics and Probability Letters (2019), Vol. 145, 1–11.

Résumé

Le but de la thèse est l'étude des chaînes de Markov itératives i.e obtenues par itération de fonctions aléatoires. En utilisant des techniques d'opérateurs linéaires quasi-compacts, nous montrons l'existence et l'unicité d'une mesure invariante (stationnaire) pour ces chaînes, dans le cas où les fonctions aléatoires itérées sont Lipschitziennes. Nous appliquons cette approche à l'étude de comportement asymptotique de la chaîne de Diaconis-Freedman sur $[0, 1]$. Nous obtenons une condition nécessaire et suffisante d'unicité de la loi invariante. Nous explorons le cas où cette condition n'est pas vérifiée et nous montrons alors que les lois invariantes de la chaîne sont les combinaisons convexes des mesures de Dirac δ_0 et δ_1. Nous indiquons pour terminer quelques idées d'extension de la chaîne de Diaconis-Freedman et nous conjecturons quelques résultats.

Mots-Clefs : Itération de fonctions aléatoires ; Opérateur quasi-compact ; Chaînes de Markov ; Mesure invariante ; Compact absorbant ; Récurrence.

Abstract

The aim of this thesis is the study of iterated Markov chains i.e generated by iteration of random functions. Using quasi-compact linear operator technics, we show the existence and uniqueness of an invariant measure for these chains, in the case where the iterated random functions are Lispchitz. We apply this approach to study the asymptotic behavior of the Diaconis-Freedman chain on $[0, 1]$. We obtain a necessary and sufficient condition for the uniqueness of the stationary probability measure. We explore the case where this condition is not hold and then we show that the invariant probability measures of the chain are the convex combinations of the Dirac measures δ_0 and δ_1. We end up with some ideas for extending the Diaconis-Freedman chain and we conjecture some results about.

Keywords : Iterated random functions ; quasi-compact linear operators ; Markov chains ; invariant measure ; absorbing compact set ; Recurrence.

Table des matières

Introduction

Les processus et chaînes de Markov sous la forme d'itérées de fonctions aléatoires ou de systèmes dynamiques sont intensément étudiés ces dernières années. Rappelons que cela fait suite aux travaux de Doeblin et Fortet (1937), Norman (1972), Fuhrstenberg (1963),... Kifer (1986) assurant en particulier que toute chaîne de Markov peut être mise sous la forme itérative. Sous cette forme et suite à ces développements, les processus et chaînes de Markov ont offert un cadre théorique adapté aux théories de l'apprentissage et ont ouvert la voie à de nombreux champs d'application : fractals (cf. Barnsley et al [3] et Barnsley et al [5]), traitement d'images (cf. Barnsley et Elton [4]), dynamique des populations, météorologie, actuariat,... Cela a permis, en particulier, l'émergence des techniques de simulation exacte (cf. Propp et Wilson ([56], [57])) réalisant ainsi une avancée remarquable dans les méthodes de simulation.

Les données un peu simplifiées sont un ensemble E, une suite i.i.d. de fonctions aléatoires $(F_n)_n$ à valeurs dans E et une variable aléaoire X_0. Nous nous intéressons alors au processus $(X_n)_n$ défini par X_0 et pour $n \geq 1$, $X_n = F_n(X_{n-1})$. Les F_n sont, souvent, sous la forme $F_n = f_{Y_n}$ où f est une fonction de $E \times E$ dans E, avec $f_y(x) = f(x,y)$ et (Y_n) une suite de variables aléatoires i.i.d. à valeurs dans E. Dans les théories de l'apprentissage, en psychologie on considère quelquefois que les Y_n sont les stimuli et les X_n les réponses du sujet.

Les exemples les plus simples de systèmes dynamiques sont les marches aléatoires sur les groupes. Dans le cas général (cas où les f_{Y_n} sont non linéaires) les méthodes d'étude changent car les f_{Y_n} génèrent en général des semi-groupes et de ce fait les méthodes et techniques d'étude des marches aléatoires ne sont pas très adaptées.

Notre problématique s'insère dans ce cadre. De façon plus précise nous nous intéressons à la chaîne de Diaconis-Freedman, en tant qu'exemple de système dynamique où les transitions sont aussi des fonctions spatiales. Il s'agit de caractériser le comportement limite de cette chaîne dans le cadre général des travaux de Letac [42], Diaconis et Freedman [20], Kaijser [36],...

Ce mémoire est structuré en quatre chapitres. Le premier chapitre est composé de trois parties, dans la première partie nous donnons quelques motivations en lien avec la théorie de l'apprentissage pour l'étude des processus itératifs. Nous définissons les systèmes aléatoires à liaisons complètes qui ont servi de modèles généraux pour les théories de l'apprentissage en psychologie. Dans la deuxième partie nous faisons une synthèse des éléments de base sur les processus et chaînes de Markov à partir de la littérature classique sur le sujet. Dans la troisième partie nous faisons un résumé des travaux sur les itérations de fonctions aléatoires : Le théorème de Kifer [38] et sa démonstration, le principe de contraction de Letac [42], le théorème de Diaconis et Freedman [20],...

Dans le deuxième chapitre, nous rappelons quelques résultats fondamentaux et notions de base sur les opérateurs linéaires. Une grande partie de ce chapitre est consacrée aux opérateurs linéaires quasi-compacts. Dans le cas particulier où P est l'opérateur de transition associé à une chaîne de Markov $(X_n)_n$ a espace d'états E compact, alors moyennant quelques hypothèses, la quasi-compacité de P permet une étude de la chaîne $(X_n)_n$ très similaire à celle qui peut être faite dans le cas où E est fini (cf. Hennion [29]). Voila donc l'intérêt de la quasi compacité pour l'étude des chaînes de Markov. Ensuite nous donnons une propriété spectrale importante pour ce type d'opérateur et le théorème de Ionescu Tulcea et Marinescu sur une condition suffisante de quasi-compacité. Dans la suite du chapitre on s'intéresse à la définition donnée par Hennion et Hervé [30]. Les propriétés spectrales et structurelles importantes qui s'ensuivent permettent la généralisation du théorème de Ionescu Tulcea et Marinescu. Nous terminons le chapitre par une application faite par Hervé dans [32], de ces propriétés à l'étude des fonctions continues f solutions de l'équation $Pf = \rho f$ et des mesures $P-$invariantes, où ρ est le rayon spectral d'un opérateur linéaire positif et quasi-compact P donné, en se basant sur la notion de compact absorbant. Il s'agit des résultats utiles pour l'étude de la chaîne de Diaconis-Freedman.

Dans le chapitre 3, nous étudions une chaîne de Markov itérative. Il s'agit de la chaîne de Diaconis- Freedman sur $[0,1]$ décrivant les mouvements d'une particule entre deux points attractifs $A_0 = 0$ et $A_1 = 1$. Si on note par Z_n la position de la particule à l'instant n, $n \in \mathbb{N}$ et si à l'instant $n, Z_n = x, x \in [0,1]$ alors à l'instant $n+1$, elle se dirige vers 0 avec la probabilité $p(x)$ ou vers 1 avec la probabilité $q(x) = 1 - p(x)$ et se déplace uniformément à l'intérieur de l'intervalle choisi.

Au début du chapitre, nous faisons une synthèse des travaux réalisés concernant cette chaîne. En utilisant la théorie spectrale des opérateurs linéaires quasi-compact, nous montrons l'existence et l'unicité d'une mesure invariante pour les chaînes de Markov générées par itération de fonctions aléatoires lipschitziennes indépendantes et identiquement distribuées. Dans la suite nous étendons ce résultat au cas des systèmes d'itération de fonctions aléatoires, avec dépendance spatiale. Nous appliquons cette approche à l'étude de comportement asymptotique de la chaîne de Diaconis-Freedman sur [0,1]. On montre l'unicité de la loi invariante sous la condition de stricte positivité de $p(x)$. Nous terminons ce chapitre, qui contient notre résultat principal, par l'étude de la chaîne dans un cadre plus général, en exploitant la notion de compact absorbant.

Le chapitre 4 est consacré à quelques extensions de la chaîne de Diaconis-Freedman. Nous résumons les extensions abordées dans la littérature actuelle concernant la chaîne dans le cas de deux points attractifs. Nous étudions ensuite la chaîne de Markov modélisant les mouvements d'une particule attirée par trois points alignés dans des conditions particulières. Nous développons le chapitre par une extension au cas bidimensionnel avec quelques exemples et nous terminons par une conjecture.

Nous terminons cet exposé par une conclusion et les perspectives ouvertes par cette étude.

Chapitre 1

Modèles d'apprentissage, processus et chaînes de Markov

1.1 Modèles stochastiques d'apprentissage

1.1.1 Historique et motivation

En psychologie l'apprentissage est l'acquisition de savoir faire et de connaissances. Les modèles d'apprentissage sont classés selon trois principaux courants : le modèle transmissif, le modèle behavioriste et le modèle socio-constructiviste .

La conception transmissive de l'apprentissage est basée sur l'idée que pour apprendre, l'apprenant doit être attentif, écouter, suivre, imiter, répéter et appliquer. Quelques auteurs utilisent l'image de la boite vide qu'il s'agirait de remplir, pour définir ce modèle.

Pour le modèle behavioriste, l'apprentissage est vu comme la mise en relation entre une action provoquée de l'extérieur (stimulus) et une réaction adéquate du sujet qui cause un changement de comportement significatif. Pavlov fut le premier à expérimenter ce modèle par ses travaux sur les comportements réflexes ou le conditionnement. Pour rappel, il s'agit de l'expérience sur un chien qui salive si on active une clochette en même temps qu'on lui apporte à manger et qui par la suite fait l'apprentissage de saliver dès qu'on active le stimulus(clochette). Skinner (1978) a repris cette théorie en l'appliquant à l'apprentissage. Il met en évidence dans la réponse de l'apprenant, l'automatisation du comportement.

Il introduit le concept de maintien de la satisfaction par la récompense et découvre un processus qui "positivise" l'apprentissage :

- stimulus,
- réponse,
- récompense (en cas de bonne réponse),
- renforcement.

Le renforcement est ici une stabilisation de la connaissance apprise par la répétition d'une réponse correcte donnée. Le principe qui sous-tend cette approche est que l'on ne peut pas connaître complètement les processus internes (dans le cerveau). On doit alors s'appuyer sur l'expérimentation et ses données (stimulus et résultat). L'apprentissage apparaît alors comme un mécanisme. On résume souvent le behaviorisme dans ce schéma simple :
Stimulus $\rightarrow$ Formation $\rightarrow$ Rsultat.

Quant au modèle socio-constructiviste issu des recherches de Piaget (1925), contrairement au modèle précédent, il considère que l'on peut étudier ce qui se passe dans la boîte noire (le cerveau). Piaget pense que la connaissance se construit. Ce modèle a été développé aussi par l'école russe de psychologie, en particulier par les travaux de Vygostsky, Leontiev (cf. [39]) et d'autres qui considèrent que les connaissances se construisent par l'activité. Pour eux, on apprend mieux par les interactions de groupe qui permettent les échanges d'expériences.

Dans le courant behavioriste, le besoin d'un modèle quantitatif pour l'analyse des situations expérimentales a donné naissance aux modèles mathématiques d'apprentissage. L'adaptation de ces modèles, pour tenir compte de la variation des stimuli et des réponses, est formalisée dans les modèles stochastiques d'apprentissage. Ces modèles sont apparus dans les travaux de Bush et Mosteller (1955) sur la théorie des modèles linéaires, Este (1950) sur la théorie de l'échantillonnage des stimuli... Pour plus de détails le lecteur peut consulter Norman [46] et ses références pour une bonne bibliographie sur le sujet.

Par la suite certains de ces modèles ont trouvé leur cadre naturel dans des théories mathématiques. C'est le cas du modèle stochastique d'apprentissage dans le cas markovien. Dans ce modèle, les données sont un espace des états(réponses) et un espace des paramètres(stimuli) et une procédure qui permet de relier ces deux espaces. Les systèmes à liaison complètes en sont une première formalisation.

1.1.2 Systèmes aléatoires à liaisons complètes

Soient $(E, \mathcal{E})$ et $(F, \mathcal{F})$ deux espaces mesurables et $p : E \times \mathcal{F} \to [0,1]$ un noyau de transition de $(E, \mathcal{E})$ dans $(F, \mathcal{F})$ (i.e. pour $x \in E$, $p(x, .)$ est une probabilité sur $\mathcal{F}$ et pour tout $B \in \mathcal{F}$, $p(., B)$ est $\mathcal{E}-$ mesurable). Soit $T : E \times F \to E$ une application $\mathcal{E} \times \mathcal{F}$ mesurable. Le système $((E, \mathcal{E}), (F, \mathcal{F}), p, T)$ est appelé système aléatoire (homogène) à liaisons complètes. Le processus $Z = X_0, Y_0, X_1, Y_1, ...$ des variables aléatoires sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ est appelé processus associé au système $((E, \mathcal{E}), (F, \mathcal{F}), P, T)$ si X_n et Y_n prennent leurs valeurs dans $(E, \mathcal{E})$ et $(F, \mathcal{F})$ respectivement,

$$X_{n+1} = T(X_n, Y_n) \tag{1.1}$$

et

$$\forall A \in \mathcal{F} \;\; \mathbb{P}(Y_n \in A / X_n, Y_{n-1}, ...) = p(X_n, A) \tag{1.2}$$

presque sûrement(p.s). Les processus $X = (X_n)_n$ et $Y = (Y_n)_n$ sont appelés processus des états et processus des indices (événements) respectivement, $(E, \mathcal{E})$ est l'espace des états et $(F, \mathcal{F})$ est l'espace des indices (événements). La loi μ_0 de X_0 est la loi initiale du processus associé au système $((E, \mathcal{E}), (F, \mathcal{F}), p, T)$.

Le concept de système aléatoire à liaisons complètes a été considéré dans la littérature comme modèle général dans la plupart des modèles stochastiques d'apprentissages. Dans ce contexte X_n caractérise la réponse du sujet à l'essai n et l'événement Y_n est la cause qu'on considère comme un stimulus qui affect la réponse du sujet.

C'est le cadre général des processus itératifs adapté aux théories de L'apprentissage. Nous nous intéressons à des modèles de chaînes de Markov de ce type, sous des hypothèses plus restrictives et plus naturelles.

1.2 Processus et chaînes de Markov

Nous rassemblons ici quelques outils de base sur les chaînes de Markov.

Définition 1.1. *Soit $(E, \mathcal{E})$ un espace mesurable. Une application $P : E \times \mathcal{E} \longrightarrow [0,1]$ est un noyau de transition sur E si :*

1. *Pour tout $x \in E$, l'application $P(x,.) : \mathcal{E} \longrightarrow [0,1]$ est une probabilité sur $(E, \mathcal{E})$;*
2. *Pour tout $A \in \mathcal{E}$, l'application $P(.,A) : \ E \longrightarrow [0,1]$ est mesurable.*

Définition 1.2. *Soit $(E, \mathcal{E})$ un espace mesurable et μ_0 une mesure de probabilité sur E. Un processus aléatoire $(X_n)_n$ défini sur l'espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans E, est une chaîne de Markov d'espace des états E et de loi initiale μ_0 si :*

1. *μ_0 est la loi de X_0 ;*
2. *Le processus $(X_n)_n$ vérifie la propriété de Markov i.e. $\forall A \in \mathcal{E}, \forall n \geq 0$, on a*

$$\mathbb{P}[X_{n+1} \in A | X_n, \ ..., \ X_1, \ X_0] = \mathbb{P}[X_{n+1} \in A | X_n] \quad \text{p.s.}$$

L'application $P : E \times \mathcal{E} \longrightarrow [0,1]$ définie par

$$\forall A \in \mathcal{E}, \ P(X_n, A) = \mathbb{P}[X_{n+1} \in A | X_n]$$

est le noyau de transition de la chaîne $(X_n)_n$. La chaîne $(X_n)_n$ est dite stationnaire ou homogène si son noyau de transition P ne dépend pas de n.

Remarquons qu'étant donnée une chaîne de Markov on peut lui associer un noyau de transition P. Réciproquement si on a un noyau de transition on peut lui associer une chaîne de Markov (cf. Neveu [46]).

Proposition 1.1. *Soit $(X_n)_n$ une chaîne de Markov homogène d'espace des états E de loi initiale μ_0 et de noyau de transition P sur $(E, \mathcal{E})$. Alors $\forall \ A_0, A_1, \cdots A_n \in \mathcal{E}$,*

$$\mathbb{P}(X_0 \in A_0, ..., X_n \in A_n) = \int_{A_0} \mu_0(\mathrm{d}x_0) \int_{A_1} P(x_0, \mathrm{d}x_1) ... \int_{A_n} P(x_{n-1}, \mathrm{d}x_n). \tag{1.3}$$

Démonstration. Il suffit de vérifier l'équation (1.3) par récurrence sur n. Pour $n = 0$, c'est évident car $\mathbb{P}(X_0 \in A_0) = \mu_0(A_0) = \int_{A_0} \mu_0(\mathrm{d}x_0)$.
Pour $n = 1$,

$$\begin{aligned}
\mathbb{P}(X_0 \in A_0, X_1 \in A_1) &= \int_{\{X_0 \in A_0\}} \mathbb{P}(X_1 \in A_1 | X_0) \mathrm{d}\mathbb{P} \\
&= \int_{A_0} \mathbb{P}(X_1 \in A_1 | X_0 = x_0) \mu_0(\mathrm{d}x_0) \\
&= \int_{A_0} \big[\int_{A_1} \mathbb{P}(X_1 \in \mathrm{d}x_1 | X_0 = x_0) \big] \mu_0(\mathrm{d}x_0) \\
&= \int_{A_0} \mu_0(\mathrm{d}x_0) \int_{A_1} P(x_0, \mathrm{d}x_1),
\end{aligned}$$

La formule générale s'en suit par récurrence. □

Remarque 1.1. Cette proposition montre, en particulier, que le noyau de transition P caractérise complètement la chaîne de Markov $(X_n)_n$.

Définition 1.3. *Soit $(X_n)_n$ une chaîne de Markov sur E, de noyau de transition P. Soit $(\mathcal{F}_n)_n$ une filtration i.e. une suite croissante de sous-tribus de $\mathcal{A}$. On appelle temps d'arrêt ou temps markovien relativement à la filtration $(\mathcal{F}_n)_n$, toute variable aléatoire τ à valeurs dans $\mathbb{N} \cup \{+\infty\}$, tel que*

$$\forall n \in \mathbb{N}, \quad \{\tau \leq n\} \in \mathcal{F}_n.$$

La tribu définie par

$$\mathcal{F}_\tau = \{A \in \mathcal{A},\ \forall n \in \mathbb{N},\ A \cap \{\tau \leq n\} \in \mathcal{F}_n\}$$

est appelée tribu engendrée par τ ou tribu des événements antérieurs à τ.

L'espace produit $(E^\infty, \mathcal{E}_\infty)$ est défini par l'ensemble $E^\infty = \prod_{i=1}^\infty E_i$ avec $E_i = E\ \forall i$, muni de la tribu engendrée par les ensembles cylindriques $\mathcal{E}_\infty$.

Proposition 1.2 (Propriété de Markov forte). *Avec les notations ci-dessus, soit $(X_n)_n$ une chaîne de Markov sur E de noyau de transition P et $\tau < \infty$ un temps d'arrêt relativement à la filtration naturelle $(\mathcal{F}_n)_n$ i.e. $\mathcal{F}_n = \sigma(X_0, \cdots X_n)$. Pour $B \in \mathcal{E}_\infty$, on a alors*

$$\mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B | \mathcal{F}_\tau] = \mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B | X_\tau].$$

Démonstration. Soit $A \in \mathcal{F}_\tau$, on a

$$\begin{aligned}
\mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B,\ A] &= \sum_{n=0}^{\infty} \mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B,\ A,\ \tau = n] \\
&= \sum_{n=0}^{\infty} \mathbb{P}[(X_n, X_{n+1}, ...) \in B,\ A,\ \tau = n]
\end{aligned}$$

Comme $A \cap \{\tau = n\} \in \mathcal{F}_n$ alors

$$\begin{aligned}
\mathbb{P}[(X_n, X_{n+1}, ...) \in B,\ A,\ \tau = n] &= \int_{A\cap\{\tau=n\}} \mathbb{P}[(X_n, X_{n+1}, ...) \in B | X_n, ...X_0) \mathrm{d}\mathbb{P} \\
&= \int_{A\cap\{\tau=n\}} \mathbb{P}[(X_n, X_{n+1}, ...) \in B | X_n] \mathrm{d}\mathbb{P},
\end{aligned}$$

d'après la propriété de Markov.

Soit alors φ la fonction définie par $\varphi(X_n) = \mathbb{P}[(X_n, X_{n+1}, ...) \in B | X_n]$. Nous avons donc

$$\begin{aligned}
\mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B,\ A] &= \sum_{n=0}^{\infty} \int_{A\cap\{\tau=n\}} \mathbb{P}[(X_n, X_{n+1}, ...) \in B | X_n) \mathrm{d}\mathbb{P} \\
&= \sum_{n=0}^{\infty} \int_A 1_{\{\tau=n\}} \varphi(X_\tau) \mathrm{d}\mathbb{P} \\
&= \int_A \varphi(X_\tau) \mathrm{d}\mathbb{P}.
\end{aligned}$$

On a donc par unicité de la probabilité conditionnelle

$$\mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B | \mathcal{F}_\tau] = \varphi(X_\tau) = \mathbb{P}[(X_\tau, X_{\tau+1}, ...) \in B | X_\tau]$$

□

Soit $(E, \mathcal{E})$ un espace mesurable, $(X_n)_n$ une chaîne de Markov d'espace des états E de loi initiale μ_0 et de noyau de transition P sur $(E, \mathcal{E})$. Pour f mesurable et bornée, définissons l'application $f \mapsto Pf$ en posant

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

De même, pour toute mesure bornée μ soit l'opérateur $\mu \mapsto \mu P$, défini pour $B \in \mathcal{E}$, par

$$\mu P(B) = \int_E \mu(\mathrm{d}x) P(x, B)\ .$$

Soient $\mathcal{M}_b(E)$ l'espace des fonctions mesurables et bornées et $\mathbf{M}_b(E)$ l'espace des mesures bornées sur $(E, \mathcal{E})$.

Définissons le crochet de dualité $< \mu, f >$ pour $\mu \in \mathbf{M}_b(E)$ et $f \in \mathcal{M}_b(E)$ par

$$< \mu, f >= \mu(f) = \int f \mathrm{d}\mu.$$

Les deux opérateurs définis ci dessus opérant le premier sur l'espace $\mathcal{M}_b(E)$ le deuxième sur l'espace $\mathbf{M}_b(E)$ sont adjoints l'un de l'autre i.e.

$$\forall \mu \in \mathbf{M}_b(E),\ \forall f \in \mathcal{M}_b(E);\ < \mu, Pf >=< \mu P, f > .$$

Si on définit les itérés $P^n(x,.)$ de $P(x,.)$ en posant $P^0(x,.) = \delta_x$ et pour $n > 0$,

$$P^n(x, B) = \int_E P(x, dy) P^{n-1}(x, B)$$

alors on peut considérer les opérateurs P^n en posant, pour toute fonction mesurable et bornée sur E,

$$P^n f(x) = \int_E f(y) P^n(x, dy)$$

De même, on peut considérer les opérateurs P^n en posant, pour toute mesure bornée μ

$$\mu P^n(B) = \int_E \mu(dx) P^n(x, B) \ .$$

L'étude des opérateurs P et P^n pour la chaîne de Markov $(X_n)_n$ est intéressante car nous avons immédiatement

$$\mathbb{P}[X_{n+m} \in B | X_n] = P^m \mathbb{1}_B(X_n) = P^m(X_n, B),$$

$$\mathbb{E}[f(X_{n+m}) | X_n] = P^m f(X_n)$$

et si $\mathcal{L}(X_n) = \mu_n$ est la loi de X_n, alors pour $m \geq 0$,

$$\mu_{n+m} = \mathcal{L}(X_{n+m}) = \mu_n P^m \ .$$

Définition 1.4. *Une mesure ν sur $(E, \mathcal{E})$ est dite P-invariante ou est une mesure stationnaire pour la chaîne de Markov $(X_n)_n$, si*

$$\nu(B) = \int_E P(x, B)\nu(\mathrm{d}x) \qquad \forall B \in \mathcal{E}.$$

De façon équivalente, ν est P- invariante si

$$\int_E Pf(x)\nu(\mathrm{d}x) = \int_E f(x)\nu(\mathrm{d}x), \quad \forall f \in \mathcal{M}_b(E).$$

où $\mathcal{M}_b(E)$ est l'espace des fonctions réelles bornées et mesurables sur E.

Un ensemble $B \in \mathcal{E}$ est dit stochastiquement fermé (ou invariant ou absorbant) pour P s'il est non vide et $P(x, B) = 1$ pour chaque $x \in B$. Une mesure invariante ν est ergodique si $\nu(B) = 0$ ou $\nu(B) = 1$ pour tout ensemble stochastiquement fermé B.

Une mesure stationnaire pour la chaîne $(X_n)_n$ est un point fixe pour l'application P définie sur l'espace $\mathbf{M}(E)$ des mesures signées sur $(E, \mathcal{E})$ par

$$\mu P(B) = \int_E \mu(\mathrm{d}x)P(x, B), \quad \forall \mu \in \mathbf{M}(E).$$

Une mesure invariante n'est pas toujours une loi.

Un des critères importants pour vérifier si une chaîne de Markov de noyau P admet une mesure invariante est la propriété de Feller que nous définissons, ci-dessous.

Définition 1.5. *Soit E un espace métrique. L'opérateur P a la propriété faible de Feller si, pour toute suite (x_n) de E convergente vers un élément x de E et toute fonction continue et bornée f sur E on a*

$$\int_E f(y)P(x_n, dy) \;\longrightarrow \int_E f(y)P(x, dy) \quad \text{quand} \quad \text{n} \longrightarrow \infty.$$

Dans le cas où cette propriété est vraie pour toute fonction f mesurable et bornée, on dit que l'opérateur P a la propriété de Feller forte.

Théorème 1.1 (Lerma et Lasserre [41]). *Soit E un espace métrique compact et $(X_n)_n$ une chaîne de Markov d'espace des états E de noyau de transition P. Si P a la propriété de Feller faible alors il admet une mesure invariante.*

Théorème 1.2 (Lerma et Lasserre [41]). *Soit E un espace métrique localement compact et séparable. Soit $(X_n)_n$ une chaîne de Markov d'espace des états E et de noyau de transition P. Supposons que P a la propriété de Feller faible. Soit f_0 une fonction continue, strictement positive et tendant vers 0 à l'infini. S'il existe un nombre réel positif b et une fonction mesurable $g : E \to \mathbb{R}^+$ non identiquement nulle tels que*

$$Pg(x) \leq g(x) - 1 + bf_0(x) \quad \forall x \in E,$$

alors il existe une mesure P-invariante.

Définition 1.6. *Soit $(X_n)_n$ une chaîne de Markov à espace d'états dénombrable E et irréductible. Pour $x \in E$ un état, soit τ_x la variable aléatoire à valeurs dans $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ définie par :*

$$\tau_x = \begin{cases} \inf\{n > 0/\ X_n = x\} & si \ \bigcup_n \{\omega / X_n(\omega) = x\} \neq \emptyset \\ \\ +\infty & si \ \bigcup_n \{\omega / X_n(\omega) = x\} = \emptyset \end{cases}$$

τ_x est le premier instant d'entrée dans l'état x, pour la chaîne $(X_n)_n$.
On dit qu'un état $x \in E$ est :

- *récurrent pour la chaîne $(X_n)_n$ si $\mathbb{P}[\tau_x < \infty \,|X_0 = x] = 1$,*
- *transient pour la chaîne $(X_n)_n$ si $\mathbb{P}[\tau_x < \infty \,|X_0 = x] = 0$.*

Une chaîne de Markov irréductible est dite transiente (resp. récurrente) si elle admet un état transient (resp. récurrent).

Dans le cas où E n'est pas dénombrable, nous précisons ces propriétés ci-dessous (cf. Meyn et Tweedie [58]).

Définition 1.7. *Soit $(X_n)_n$ une chaîne de Markov à valeurs dans l'espace mesurable $(E, \mathcal{E})$ quelconque. Pour $A \in \mathcal{E}$, soit η_A la variable aléatoire à valeurs dans $\overline{\mathbb{N}}$ définie par : $\eta_A := \sum_{n=1}^{\infty} \mathbb{1}_{\{X_n \in A\}}$, η_A est le nombre de visite de A.*

L'ensemble A est dit :

- *uniformément transient si $\mathbb{E}_x(\eta_A) = \sum_{n=1}^{\infty} P^n(x, A) < \infty \ \ \forall x \in A$,*

- *récurrent si* $\mathbb{E}_x(\eta_A) = \sum_{n=1}^{\infty} P^n(x, A) = \infty \ \ \forall x \in A$, *où* $P^n(x, A) = \mathbb{P}(X_n \in A | X_0 = x)$,
- *Harris récurrent si* $\mathbb{P}_x(\eta_A = \infty) = 1$,
- *transient s'il peut être recouvert par une famille dénombrable d'ensembles uniformément transients.*

Définition 1.8. *Soit* μ *une mesure* $\sigma-$*finie sur* E. *Une chaîne de Markov* $(X_n)_n$ *à espace d'états* E *et de noyau de transition* P *est dite* $\mu-$ *irréductible si, pour tout* $A \in \mathcal{E}$ *tel que* $\mu(A) > 0$, *on a* $\sum_{k=1}^{\infty} P^k(x, A) > 0$ *pour tout* $x \in E$.

Théorème 1.3 (**Meyn et Tweedie [58]**). *Pour toute probabilité de transition* P *d'une chaîne de Markov,* μ*- irréductible, deux cas seulement sont possibles*

a) *Ou bien* E *est transient c'est à dire il existe une famille dénombrable d'ensembles* A_j *qui forment une partition de* E, *telle que tout ensemble* $A \in \mathcal{E}$ *où* $A \subset A_j$ *pour certain* j *est uniformément transient. La chaîne est alors dite transiente.*

b) *Ou bien tout ensemble* $A \in \mathcal{E}$, *tel que* $\mu(A) > 0$ *est récurrent ; la chaîne est alors dite récurrente.*

1.3 Processus itératifs

Une chaîne de Markov est dite itérative si elle est construite par itération de fonctions aléatoires. Plus précisément, soient $(E, \mathcal{E})$ et $(F, \mathcal{F})$ deux espaces mesurables et $(T_n)_n$ une suite de fonctions aléatoires à valeurs dans E indépendantes et identiquement distribuées. Pour X_0 une variable aléatoire à valeurs dans E et indépendante des T_n, le processus $(X_n)_{n\geq 0}$ défini par

$$X_0 \text{ et } \ \forall n \geq 0, \quad X_{n+1} := T_{n+1}(X_n).$$

est une chaîne de Markov itérative, appelée aussi *chaîne associée au système d'itération de fonctions* $(T_n)_{n\geq 1}$.

Dans la littérature, souvent on considère l'application

$$\begin{aligned} T \ : \ E \times F \ &\to E \\ (x, y) \ &\mapsto \ T(x, y) = T_y(x), \end{aligned}$$

et les T_n sont alors spécialisées sous la forme $T_n = T_{Y_n}$ où $(Y_n)_n$ est une suite de variables aléatoires indépendantes et identiquement distribuées à valeurs dans F.

La proposition suivante montre que $(X_n)_{n\geq 0}$ est bien une chaîne de Markov.

Proposition 1.3. *Avec les notations ci-dessus, si μ est la loi des Y_n, alors $(X_n)_n$ est une chaîne de Markov de noyau de transition P défini pour $x \in E$ et $A \in \mathcal{E}$, par*

$$P(x, A) = \mathbb{P}\{T_{Y_1}(x) \in A\} = \mu\{y \in F, \quad T_y(x) \in A\}$$

Démonstration.
Montrons que (X_n) vérifie la propriété de Markov *i.e.*

$$\forall A \in \mathcal{E},\ \forall n,\ \mathbb{P}[X_{n+1} \in A | X_n,\ X_{n-1},\ \dots\ ,\ X_0] = \mathbb{P}[X_{n+1} \in A | X_n]$$

Calculons $\mathbb{P}[X_{n+1} \in A | X_n,\ \dots\ ,\ X_0],\ \forall n$
Par définition de la probabilité conditionnelle $\mathbb{P}[X_{n+1} \in A | X_n,\ X_{n-1}, \dots,\ X_0]$ est l'unique variable aléatoire $\sigma[X_n,\ X_{n-1} \dots,\ X_0]$ mesurable qui vérifie $\forall\ A,\ B_n,\ B_{n-1},\ \dots\ , B_0 \in \mathcal{E}$.

$$\begin{aligned}
\mathbb{P}[X_{n+1} \in A,\ X_n \in B_n,\ \dots\ ,\ X_0 \in B_0] &= \int_{\{X_n \in B_n,\ \dots\ ,\ X_0 \in B_0\}} \mathbb{P}[X_{n+1} \in A | X_n,\ X_{n-1},\ \dots\ ,\ X_0] \mathrm{d}\mathbb{P} \\
&= \int_{\{X_n \in B_n,\ \dots\ ,\ X_0 \in B_0\}} \mathbb{E}[\mathbb{1}_A(X_{n+1}) | X_n,\ X_{n-1},\ \dots\ ,\ X_0] \mathrm{d}\mathbb{P} \\
&= \int_{\{X_n \in B_n,\ \dots\ ,\ X_0 \in B_0\}} \mathbb{1}_A(X_{n+1}) \mathrm{d}\mathbb{P} \\
&= \int_{\{X_n \in B_n,\ \dots\ ,\ X_0 \in B_0\}} \mathbb{P}[X_{n+1} \in A | X_n] \mathrm{d}\mathbb{P}
\end{aligned}$$

Donc si $\varphi(X_n,\ X_{n-1},\ \dots\ ,\ X_0)$ est une fonction mesurable dans $\mathbb{R}$ tel que

$$\varphi(X_n,\ X_{n-1},\ \dots\ ,\ X_0) = \mathbb{P}[X_{n+1} \in A | X_n,\ X_{n-1}, \dots,\ X_0]$$

Alors

$$\varphi(X_n,\ X_{n-1},\ \dots\ ,\ X_0) = \mathbb{P}[X_{n+1} \in A | X_n] \text{ p.s.}$$

Par unicité de la probabilité conditionnelle, on a :

$$\mathbb{P}[X_{n+1} \in A | X_n,\ X_{n-1}, \dots,\ X_0] = \mathbb{P}[X_{n+1} \in A | X_n].$$

Calculons le noyau de transition de la chaîne $(X_n)_n$.

Par définition de la probabilité conditionnelle $\forall A \in \mathcal{E},\ \mathbb{P}[X_{n+1} \in A \,|\, X_n]$ est l'unique variable aléatoire $\sigma(X_n)$-mesurable qui vérifie :

$$\forall B \in \mathcal{E},\ \mathbb{P}[X_{n+1} \in A,\ X_n \in B] = \int_{\{X_n \in B\}} \mathbb{P}[X_{n+1} \in A \,|\, X_n] \mathrm{d}\mathbb{P}$$

Cette variable aléatoire étant $\sigma(X_n)$-mesurable, il existe une fonction $\varphi : E \to [0,\, 1]$ mesurable telle que

$$\mathbb{P}[X_{n+1} \in A \,|\, X_n] = \varphi(X_n).$$

Donc, avec cette notation, pour A et $B \in \mathcal{E}$:

$$\begin{aligned}
\mathbb{P}[X_{n+1} \in A,\ X_n \in B] &= \int_{\{X_n \in B\}} \varphi(X_n) \mathrm{d}\mathbb{P} \\
&= \int_{\{X_n \in B\}} \mathbb{E}[\mathbb{1}_A(X_{n+1}) \,|\, X_n] \mathrm{d}\mathbb{P} \\
&= \int_{\{X_n \in B\}} \mathbb{1}_A(X_{n+1}) \mathrm{d}\mathbb{P} \\
&= \int_E \int_B \mathbb{1}_{\{T_y(x) \in A\}}(x) \mathbb{P}[X_n \in \mathrm{d}x,\ Y_{n+1} \in \mathrm{d}y]
\end{aligned}$$

Par indépendance de Y_{n+1} et X_n, nous avons :

$$\int_E \int_B \mathbb{1}_{\{T_y^{-1}(A)\}}(x) \mathbb{P}[X_n \in \mathrm{d}x,\ Y_{n+1} \in \mathrm{d}y] = \int_E \int_B \mathbb{1}_{\{T_y^{-1}(A)\}}(x) \mathbb{P}[X_n \in dx] \cdot \mathbb{P}[Y_{n+1} \in \mathrm{d}y]$$

En appliquant le théorème de Fubini, nous aurons :

$$\begin{aligned}
\int_E \int_B \mathbb{1}_{\{T_y(x) \in A\}}(x) \mathbb{P}[X_n \in \mathrm{d}x] \cdot \mathbb{P}[Y_{n+1} \in dy] &= \int_E \int_B \mathbb{1}_{\{T_y(x) \in A\}} \mathbb{P}[X_n \in \mathrm{d}x] \mu(\mathrm{d}y) \\
&= \int_B \mathbb{P}[X_n \in \mathrm{d}x] \int_E \mathbb{1}_{\{T_y(x) \in A\}}(x) \mu(dy) \\
&= \int_B \mathbb{P}[X_n \in \mathrm{d}x] \mu\{y \in F,\ T_y(x) \in A\}
\end{aligned}$$

Par unicité de φ

$$P(x,\ A) = \mathbb{P}[X_{n+1} \in A \,|\, X_n = x] = \mu\{y \in F,\ T_y(x) \in A\}, \quad x \in E,\ A \in \mathcal{E}$$

□

Ainsi nous venons de montrer que tout système d'itération de fonctions aléatoires i.i.d définit un processus de Markov. La question inverse à savoir est ce que tout processus de Markov, peut être construit par itération de fonctions aléatoires est tout a fait naturelle. Le théorème suivant répond a cette question par l'affirmative.

Théorème 1.4 (Kifer [38]). *Soit E un sous ensemble Borélien d'un espace Polonais i.e. un espace métrique complet et séparable. Alors, pour toute famille de mesures de probabilité $\{P(x,.),\ x \in E\}$ tel que l'application $x \mapsto P(x,B)$ est mesurable pour tout Borélien B de E, il existe une mesure de probabilité μ sur l'espace des applications mesurables T de E dans E vérifiant :*

$$\mu\{T,\ T(x) \in B\} = P(x,B)$$

pour tout $x \in E$ et $B \in \mathcal{E}$, où $\mathcal{E}$ est la tribu des Boréliens de E.

Démonstration. Nous reprenons ici la preuve de Kifer [38] et nous renvoyons le lecteur à [7], [2] et [9] pour d'autres variantes du théorème et leurs démonstrations.

La démonstration se fait en deux étapes.

i) On suppose dans un premier temps que $E = [0,1] = \mathbb{I}$. Soit Φ l'application de $\mathbb{I} \times \mathbb{I} \to \mathbb{I}$ définie par :

$$\Phi(x,y) = \inf\{\gamma;\ P(x,[0,\gamma]) \geq y\}.$$

L'application $x \mapsto \Phi(x,y)$ est mesurable pour chaque $y \in \mathbb{I}$. En effet,

$$\begin{aligned} \{x, \Phi(x,y) > a\} &= \{x, \inf\{\gamma;\ P(x,[0,\gamma]) \geq y\} > a\} \\ &= \{x;\ P(x,[0,a]) < y\}. \end{aligned}$$

l'ensemble $\{x;\ P(x,[0,a]) < y\}$ est mesurable puisque l'application $x \mapsto P(x,B)$ est mesurable pour tout Borélien $B \in \mathcal{E}$.

Soit λ la mesure de Lebesgue sur $\mathbb{I}$, comme

$$\begin{aligned} \lambda\{y \in \mathbb{I};\ \Phi(x,y) > a\} &= \lambda\{y \in \mathbb{I};\ \inf\{\gamma;\ P(x,[0,\gamma]) \geq y\} > a\} \\ &= \lambda\{y \in \mathbb{I};\ P(x,[0,a]) < y\} = 1 - P(x,[0,a]) \end{aligned}$$

alors $\lambda\{y \in \mathbb{I};\ \Phi(x, y) \in [0, a]\} = P(x, [0, a])$.

Par suite pour tout Borélien Δ de $\mathbb{I}$ on a

$$\lambda\{y \in \mathbb{I};\ \Phi(x, y) \in \Delta\} = P(x, \Delta).$$

Ceci prouve le théorème dans un cas particulier.

ii) Dans le cas général où E est un sous ensemble Borélien d'un espace Polonais, on utilise le résultat suivant : E est mesurablement isométrique à un Borélien de $\mathbb{I} = [0, 1]$, c'est- à -dire qu'il existe une application mesurable injective $\varphi : E \to \mathbb{I}$ tel que $\Gamma = \varphi(E)$ est un sous ensemble Borélien de $\mathbb{I}$ et $\varphi^{-1} : \Gamma \to E$ est mesurable. Pour la démonstration de ce résultat nous renvoyons le lecteur à [7] et [9]. Pour tout $x \in E$ définissons une mesure de probabilité $\widetilde{P}(x, .)$ sur $\mathbb{I}$ par

$$\widetilde{P}(x, \Delta) = P(x, \varphi^{-1}(\Delta \cap \Gamma))$$

où Δ est un Borélien de $\mathbb{I}$.

Pour tout $x \in E$ et tout $y \in I$ posons

$$\widetilde{\Phi}(x, y) = \inf\{\gamma; \widetilde{P}(x, [0, \gamma]) \geq y\}.$$

Soit $\psi : \mathbb{I} \to E$ une application tel que $\psi = \varphi^{-1}$ sur Γ et $\psi = x_0$ sur $\mathbb{I} \setminus \Gamma$ où x_0 un point dans E, alors l'application $T_y := \psi \circ \widetilde{\Phi}(x, .)$ définie de E dans E est mesurable pour chaque $y \in \mathbb{I}$. Ainsi nous obtenons une application g de $\mathbb{I}$ dans l'espace $\mathcal{M}$ des applications mesurables de E dans E définie par $g(y) = T_y$. L'application g induit une structure de mesurabilité sur $\mathcal{M}$, de cette façon : un sous ensemble $A \subset \mathcal{M}$ est mesurable si $g^{-1}(A)$ est un sous ensemble borélien de $\mathbb{I}$. Pour tout $A \in \mathcal{M}$ tel que $g^{-1}(A)$ est un sous ensemble borélien de $\mathbb{I}$, on pose

$$\mu(A) = \lambda(g^{-1}(A)).$$

Comme précédemment on montre que :

$$\lambda\{y \in \mathbb{I};\ \widetilde{\Phi}(x, y) \in \Delta\} = \widetilde{P}(x, \Delta).$$

De plus Pour tout $B \in \mathcal{E}$ on a :

$$\begin{aligned}
\mu\{T,\ T(x) \in B\} &= \lambda\{y \in \mathbb{I};\ \ T_y(x) \in B\} = \lambda\{y \in \mathbb{I};\ \widetilde{\phi}(x,y) \in \psi^{-1}(B)\} \\
&= \widetilde{P}(x, \psi^{-1}(B)) = P(x, \varphi^{-1}(\Gamma \cap \psi^{-1}(B))) \\
&= P(x, \varphi^{-1}(\varphi(B))) = P(x, B).
\end{aligned}$$

Ce qui prouve le théorème dans le cas général. □

La construction de tout processus de Markov comme composition de fonctions aléatoires continues est donnée par le théorème suivant.

Théorème 1.5 (Kifer [38]). *Soit E un espace métrique connexe et localement compact. Supposons que l'application $x \to P(x, .)$ de E dans l'espace $\mathcal{P}(E)$ des mesures de probabilité sur E est continue en respectant la topologie faible sur $\mathcal{P}(E)$. Si pour tout $x \in E$ le support de $P(x, .)$ est E tout entier alors il existe une mesure de probabilité μ sur l'espace $C(E, E)$ des applications continues de E dans E vérifiant :*

$$\mu\{T,\ T(x) \in B\} = P(x, B)$$

pour tout $x \in E$ et $B \in \mathcal{E}$.

Les chaînes de Markov itératives sont à la base des méthodes de simulations exactes introduites par Propp et Wilson (cf. [56] et [57]). Cela a été une révolution dans les méthodes de simulation. Le théorème suivant éclaire un peu la situation. Étant donné une mesure ν continue sur $\mathbb{R}$, de fonction de répartition Ψ le théorème suivant construit une chaîne de Markov dont ν est la loi stationnaire.

Théorème 1.6 (Stenflo [62]). *Une loi de probabilité ν continue définie sur $\mathbb{R}$, de fonction de répartition Ψ est une mesure stationnaire d'une chaîne de Markov associée à un système d'itération de fonctions aléatoires (IFS) généré par les applications monotones $f_i(x) := \Psi^{-1} \circ u_i \circ \Psi(x)$, pour tout x tel que $\Psi(x) > 0$ et pondérées par les probabilités $p_i = \dfrac{1}{k}$ où $u_i(u) = \dfrac{u}{k} + \dfrac{i-1}{k}$, $0 \leq u \leq 1$, $i = 1, 2, \cdots, k$, pour tout $k \geq 2$.*

Exemple 1. Soit ν la mesure de probabilité de fonction de densité, par rapport à la mesure de Lebesgue, la fonction triangulaire d avec

$$d(x) = \begin{cases} x & 0 \leq x \leq 1 \\ 2 - x & 1 \leq x \leq 2. \end{cases}$$

Le théorème assure que la mesure ν est la loi stationnaire de la chaîne de Markov $(X_n)_n$ obtenue par itération des fonctions

$$f_1(x) = \begin{cases} \frac{x}{\sqrt{2}} & 0 \leq x \leq 1 \\ \sqrt{2x - \frac{x^2}{2} - 1} & 1 \leq x \leq 2, \end{cases}$$

et

$$f_2(x) = \begin{cases} 2 - \sqrt{1 - \frac{x^2}{2}} & 0 \leq x \leq 1 \\ 2 - \sqrt{2 - 2x + \frac{x^2}{2}} & 1 \leq x \leq 2 \end{cases}$$

choisies uniformément.

Les figures suivantes illustrent bien la convergence de $(X_n)_n$ vers la loi triangulaire.

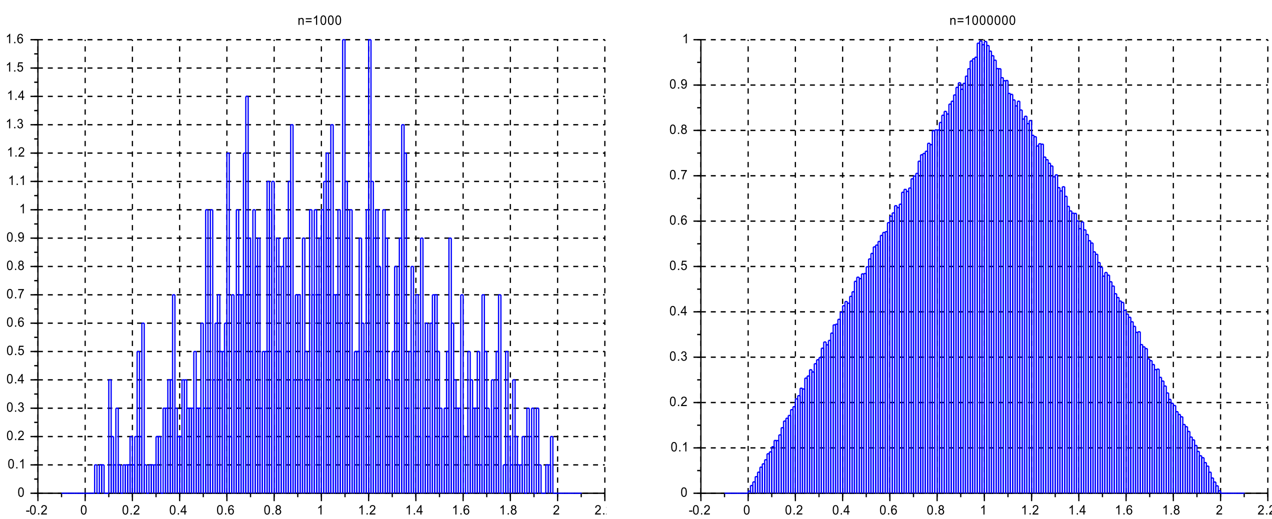

Figure 1.1 : Histogrammes des échantillons de tailles $n = 1000$ et $n = 100000$ de la loi de X_n.

1.4 Mesures invariantes et convergence des processus itératifs

Dans cette section nous nous intéressons aux conditions sur les T_n pour lesquelles la chaîne itérative $(X_n)_{n\geq 0}$ admet une limite éventuelle ou admet une mesure stationnaire. C'est à dire on s'intéresse au comportement limite de la chaîne $(X_n)_{n\geq 0}$. De nombreux travaux ont été consacrés à la recherche des mesures stationnaires pour de tels processus. Nous présentons ici les principaux résultats.

Proposition 1.4. *Soit $(Y_n)_n$ une suite de variables aléatoires réelles i.i.d. S'il existe deux fonctions réelles f et g de la variable réelle x, f mesurable et g dérivable et de dérivée $g' < 1$ sur $\mathbb{R}$ tel que $T_y(x) = f(y) + g(x)$ pour x et $y \in \mathbb{R}$, alors le processus $(X_n(x))_n$ défini par*

$$X_0 = x, \; X_{n+1} = T_{Y_{n+1}}(X_n)$$

converge en loi.

Démonstration.

Définissons le processus $(H_n(x))_n$, pour $x \in \mathbb{R}$, en posant

$$H_0 = x \text{ et pour } n > 0, \; H_n(x) = T_{Y_1} \circ T_{Y_2} \circ \cdots \circ T_{Y_n}(x)$$

Remarquons que le processus $(H_n(x))_n$ n'est pas une chaîne de Markov en général. Si g est dérivable et de dérivée $g' < 1$, alors $(H_n(x))_n$ est une suite de Cauchy p.s. Par suite $(H_n(x))_n$ converge p.s. Il s'en suit qu'il converge en loi. Comme les processus $(H_n(x))_n$ et $(X_n(x))_n$ ont la même loi, la proposition, est donc établie. $\square$

Dans le cas où les T_n sont continues et E localement compact, nous avons par exemple le théorème suivant basé sur le principe de contraction.

Théorème 1.7 (Letac [42]). *Soient $(E, \mathcal{E})$ un espace mesurable avec E localement compact et $\mathcal{B}$ sa tribu borélienne, $(F, \mathcal{F}, \mu)$ un espace de probabilité et $T : E \times F \to E$ une application mesurable par rapport à la tribu $\mathcal{E} \otimes \mathcal{F}$ tel que pour tout y fixé dans F, $x \mapsto T_y(x) = T(x, y)$ est continue. Soit alors $(Y_n)_n$ une suite de v.a. i.i.d à valeurs dans F de loi μ, et X_0 une variable aléatoire à valeurs dans E indépendante de $(Y_n)_n$. Soit $(X_n)_n$ la chaîne définie par X_0 et pour $n \geq 0$ par*

$$X_{n+1} = T(X_n, Y_{n+1}) \ ,$$

et $(H_n)_n$ la suite de variables aléatoires définie par $H_n = T_{Y_1} \circ T_{Y_2} \circ \cdots \circ T_{Y_n}(X_0)$. Notons H_n par $H_n(x)$(resp X_n par $X_n(x)$) lorsque $X_0 = x$. Si $H = \lim_n H_n(x)$ existe p.s. et ne dépend pas de x alors si ν est la loi de H, elle est unique et $(X_n)_n$ admet ν comme mesure stationnaire.

La démonstration du théorème repose sur la proposition suivante.

Proposition 1.5 (Letac [42]). *Avec les notations ci-dessus, soit ν_n la loi de $X_n(x)$ i.e. $\nu_n = \mathcal{L}(X_n(x))$. Si la suite $(\nu_n)_n$ converge étroitement vers une mesure ν indépendante de x, alors ν est une loi P- invariante.*

Démonstration de la Proposition. Nous reprenons et détaillons ici les idées de la preuve de Letac [42]. Soit $C_b(E)$ l'ensemble des fonctions à valeurs réelles continues et bornées sur E.

On a (ν_n) converge étroitement vers ν c'est à dire :

$$\forall g \in C_b(E) \quad \int_E g(x)\nu_n(dx) \to \int_E g(x)\nu(dx).$$

Soit $g \in C_b(E)$. Montrons que pour tout $n \geq 1$ on a

$$\int_E g(x_1)\nu_n(dx_1) = \int_{E\times F} g(T_y(x_1))\nu_{n-1}(dx_1)\mu(dy) \tag{1.4}$$

En effet :

$$\nu_n(dx_1) = P[X_n(x) \in dx_1] = P[T_{Y_n}(X_{n-1}(x)) \in dx_1] = \int_F P[X_{n-1}(x) \in T_y^{-1}(dx_1), Y_n \in dy].$$

Par indépendance de Y_n et X_{n-1} on aura

$$\nu_n(dx_1) = \int_F P[X_{n-1}(x) \in T_y^{-1}(dx_1)]\mu(dy) = \int_F \nu_{n-1}(T_y^{-1}(dx_1))\mu(dy),$$

donc

$$\int_E g(x_1)\nu_n(dx_1) = \int_F \int_E g(T_y(x_1))\nu_{n-1}(dx_1)\mu(dy) = \int_{E\times F} g(T_y(x_1))\nu_{n-1}(dx_1)\mu(dy).$$

Montrons maintenant que l'application :

$$x_1 \to \int_F g(T_y(x_1))\mu(dy)$$

est continue sur E.

Par hypothèse, la fonction $x \to T_y(x)$ est continue. Donc, si $(x_n)_n$ est une suite convergente vers x_1 alors

$$T_y(x_n) \to T_y(x_1).$$

D'autre part, g est continue. Donc

$$\lim_{n\to\infty} g(T_y(x_n)) = g(\lim_{n\to\infty} T_y(x_n)) = g(T_y(x_1)).$$

Comme g est bornée, par convergence dominée, on obtient :

$$\lim_{n\to\infty} \int_F g(T_y(x_n))\mu(dy) = \int_F \lim_{n\to\infty} g(T_y(x_n))\mu(dy) = \int_F g(T_y(x_1))\mu(dy).$$

Ainsi la fonction

$$x_1 \to \int_F g(T_y(x_1))\mu(dy) \quad \text{est continue.}$$

En passant à la limite dans l'équation (1.4), on aura :

$$\int_E g(x_1)\nu(dx_1) = \int_{E\times F} g(T_y(x_1))\nu(dx_1)\mu(dy). \tag{1.5}$$

Car d'une part, on a :

$$\lim_{n\to\infty} \int_E g(x_1)\nu_n(dx_1) = \int_E g(x_1)\nu(dx_1)$$

puisque $g \in C_b(E)$ et (ν_n) converge étroitement vers ν.

Et d'autre part, on a :

$$\lim_{n\to\infty} \int_{E\times F} g(T_y(x_1))\nu_{n-1}(dx_1)\mu(dy) = \int_E \big[\int_F g(T_y(x_1))\mu(dy)\big]\nu(dx_1)$$

(car $x_1 \to \int_F g(T_y(x_1))\mu(dy)$ est dans $C_b(E)$ et (ν_n) converge étroitement vers ν). Comme E est localement compact, l'équation (1.5) reste valable pour $g = \mathbb{1}_B, \quad B \in \mathcal{E}$. Ainsi en remplaçant g par $\mathbb{1}_B$ dans la relation(1.5), on aura

$$\int_E \mathbb{1}_B(x_1)\nu(dx_1) = \int_{E\times F} \mathbb{1}_B(T_y(x_1))\nu(dx_1)\mu(dy)$$

donc

$$\forall B \in \mathcal{E}, \quad \nu(B) = \int_E P(x_1, B)\nu(dx_1).$$

Cette dernière égalité montre que ν est P- invariante.

Il reste a montrer que ν est unique. Soit ν_0 une autre loi stationnaire pour la chaîne (X_n). Alors, si $\nu_0 = \mathcal{L}(X_0)$ on obtient

$$\nu_0 = \mathcal{L}(X_0) = \mathcal{L}(X_n(X_0)) \xrightarrow[n\to\infty]{} \nu \text{ et donc } \nu = \nu_0. \quad \square$$

Démonstration du théorème. Nous avons,

$$\lim_{n\to\infty} H_n(x) = H \Longrightarrow \lim_{n\to\infty} \mathcal{L}(H_n(x)) = \mathcal{L}(H).$$

Comme

$$\mathcal{L}(H_n(x)) = \mathcal{L}(X_n(x))$$

et d'après la proposition 1.5 on déduit que $\nu = \mathcal{L}(H)$ est P- invariante. $\square$

Exemple 1.1. Considérons les deux applications affines f_1 et f_2 définies de $[0,1]$ dans $[0,1]$ par $f_1(x) = \frac{1}{2}x$ et $f_2(x) = \frac{1}{2}x + \frac{1}{2}$. Soit X_0 est une variable aléatoire à valeurs dans $[0,1]$. On considère la chaîne de Markov itérative $(X_n)_n$ obtenue à partir de X_0 en composant les deux applications f_1 et f_2 pondérées par les probabilités p et $1-p$ respectivement.

La chaîne $(X_n)_n$ est donc définie par X_0 et pour $n \geq 0$

$$X_{n+1} = T_{n+1}(X_n)$$

tel que $\forall n \in \mathbb{N},\ \mathbb{P}\{T_{n+1} = f_1\} = p$ et $\mathbb{P}\{T_{n+1} = f_2\} = 1-p$.

Ainsi si $(Y_n)_n$ est une suite de variables aléatoires i.i.d de loi de Bernoulli de paramètre p sur $\{0, \frac{1}{2}\}$ et indépendantes de X_0, alors $(X_n)_n$ est donnée par

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

où $T_{Y_n}(x) = \frac{1}{2}x + Y_n$. i.e. $X_{n+1} = \dfrac{1}{2}X_n + Y_{n+1}$.

En calculant $H_n(x) = T_{Y_1} \circ T_{Y_2} \circ \cdots \ \circ T_{Y_n}(x)$ pour $x \in [0,1]$, on obtient

$$H_n(x) = \frac{x}{2^n} + \sum_{k=1}^{k=n} \frac{Y_k}{2^{k-1}}.$$

Comme $|Y_k| \leq \frac{1}{2}$, on a

$$\lim_{n \to +\infty} H_n(x) = \sum_{1}^{\infty} \frac{Y_k}{2^{k-1}} \quad \text{p.s.}$$

La limite de $H_n(x)$ existe donc presque sûrement et ne dépend pas de x. D'après le théorème précedent, on conclut que $(X_n)_n$ admet une unique loi stationnaire ν.

Il a été démontré dans Bhattacharya et Majumdar [9] que ν est une loi continue. Lorsque $p = \frac{1}{2}$, elle est la loi uniforme sur $[0,1]$. Ci-dessous quelques simulations pour illustrer ce résultat.

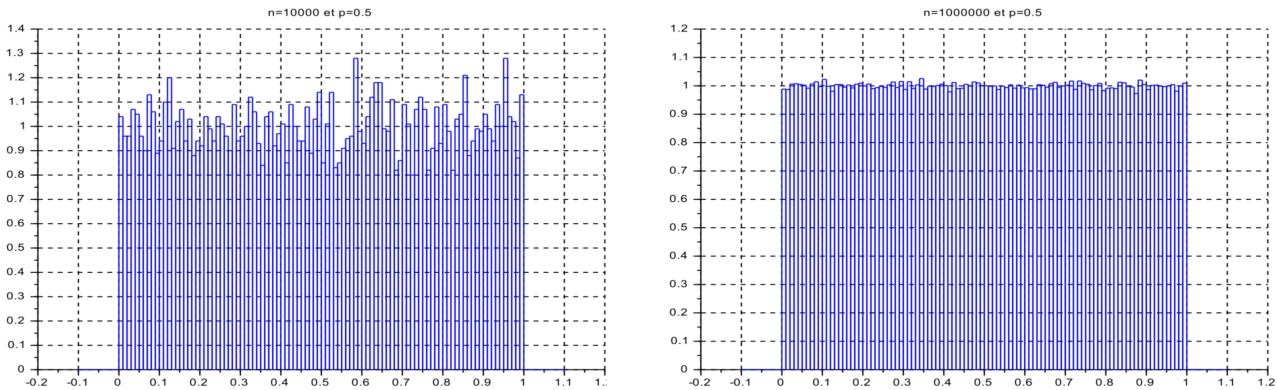

Figure 1.2 : Histogrammes des échantillons de tailles $n = 10^4$ et $n = 10^6$ de la loi de X_n, pour $p = 1/2$.

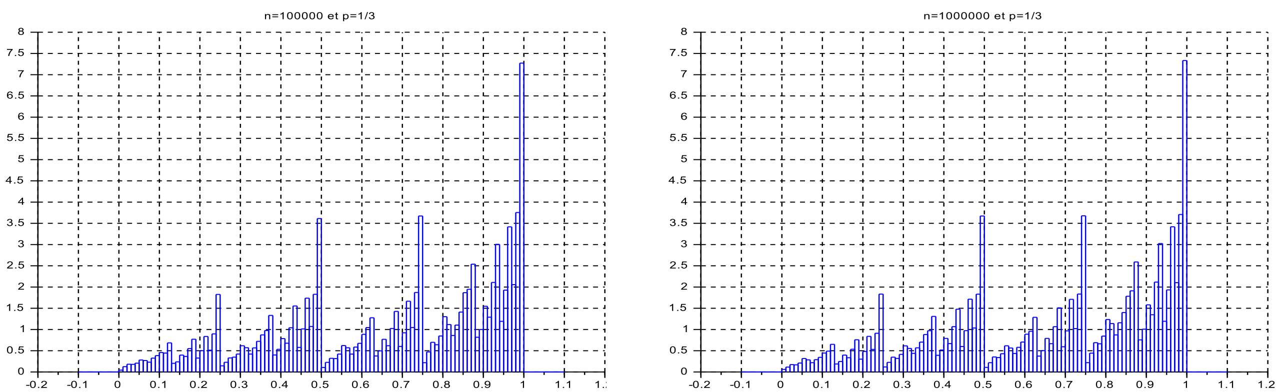

Figure 1.3 : Histogrammes des échantillons de tailles $n = 10^4$ et $n = 10^6$ de la loi de X_n, pour $p = 1/3$.

Nous signalons que Benaïm et El Karaoui [7] ont fait une étude plus générale de ce type de chaînes. Ils ont étudié la chaîne $(X_n)_n$ obtenue en composant aléatoirement les fonctions $f_1(x) = ax - 1$ et $f_2(x) = ax + 1$. i.e. $X_{n+1} = aX_n \pm 1$ où $a \in \mathbb{R}$ et $|a| < 1$. $\square$

Diaconis et Freedman [20] généralisent le théorème 1.4 aux cas où les T_n sont des fonctions lipschitziennes et contractantes en moyenne et montre la convergence géométrique de la chaîne au sens de la métrique de Prokhorov. Avant d'énoncer leur théorème nous donnons quelques définitions.

Définition 1.9 (Distance de Prokhorov)**.** *La distance de Prokhorov de deux mesures de probabilité μ et ν sur la tribu borélienne $\mathcal{E}$ d'un espace métrique (E, d) est définie par :*

$$d_\pi(\mu, \nu) := \inf\{\epsilon > 0/\ \forall A \in E, \mu(A) \leq \nu(A^\epsilon) + \epsilon \ \text{ et } \ \nu(\mathrm{A}) \leq \mu(\mathrm{A}^\epsilon) + \epsilon\},$$

où A^ϵ désigne l'ensemble des points de E dont la distance à A est strictement inférieure à ϵ.

Définition 1.10. *Une variable aléatoire U est dite a queue algébrique s'il existe deux constantes positives α et β telles que pour $u > 0$ assez grand*

$$\mathbb{P}\{U > u\} \leq \frac{\beta}{u^\alpha}.$$

Soit (E, d) un espace métrique complet et séparable et $(T_y)_{y \in F}$ une famille de fonctions de E dans E lipschitziennes de rapport K_y i.e.

$$\forall x, x' \in E,\ d(T_y(x), T_y(x')) \leq K_y d(x, x').$$

Soit $(Y_n)_n$ une suite de variables aléatoires i.i.d. de loi μ, à valeurs dans F. Considérons la chaîne de Markov $(X_n)_n$ définie par X_0 une variable aléatoire indépendante des Y_n et à valeurs dans E et pour $n > 0$

$$X_n = T_{Y_n}(X_{n-1}).$$

Soit $(H_n)_n$ la suite de variables aléatoires définie par $H_n = T_{Y_1} \cdot T_{Y_2} \cdot \cdots \cdot T_{Y_n}(X_0)$. On note $P_n(x, dy) := \mathbb{P}(X_n \in\ dy | X_0 = x)$.

Théorème 1.8 (**Diaconis et Freedman [20]**)**.** *Soit (E, d) un espace métrique complet et séparable. Soit $\{T_y :\ y \in F\}$ une famille de fonctions de E dans E lipschitziennes de rapport K_y, et μ une loi de probabilité sur F, tels que*

$$\int K_y \mu(\mathrm{d}y) < \infty,\ \int d(T_y(x_0), x_0)\mu(\mathrm{d}y) < \infty \text{ pour tout } x_0 \in E \text{ et } \int \log K_y \mu(\mathrm{d}y) < \infty.$$

Alors,

1. *la chaîne de Markov $(X_n)_n$ admet une probabilité stationnaire ν unique.*

2. *Il existe des constantes* $0 < A_x < \infty$ *et* $r \in]0,1[$ *telles que*

$$\forall n \in \mathbb{N},\ \forall x \in E, \quad d_\pi(P_n(x,.),\nu) \leq A_x r^n.$$

3. *La constante* r *ne dépend pas de* n *ou de* x *; la constante* A_x *ne dépend pas de* n *et* $A_x < a + bd(x, x_0)$ *avec* $0 < a, b < \infty$.

Proposition 1.6 (Diaconis et Freedman [20]). *Sous les conditions du théorème 1.8, la suite* (H_n) *converge p.s. à vitesse exponentielle vers une variable aléatoire de loi* ν. *La loi* ν *est la loi stationnaire de la chaîne* $(X_n)_n$.

Soit maintenant $\mathbb{L}\mathrm{ip}(E,E)$ l'espace de Banach des fonctions lipschitziennes et continues de E dans E. Soit μ une mesure de probabilité sur $\mathbb{L}\mathrm{ip}(E,E)$. On considère la chaîne $(X_n)_n$ définie de la façon suivante. Si $X_0 = x$ on choisit T_1 avec la loi μ et on calcule $X_1 = T_1(x)$ de même on choisit T_2 avec la loi μ et on calcule $X_2 = T_2(X_1)$ et ainsi de suite....

Proposition 1.7 (Diaconis et Freedman [20]). *Supposons que les conditions suivantes sont satisfaites :*

1) *la variable* $T \to K_T := \sup\limits_{\substack{x,y\in E\\ x\neq y}} \dfrac{d(T(x),T(y))}{d(x,y)}$ *est a queue algébrique relativement à* μ *i.e. il existe* α *et* β *tel que pour* u *assez grand*

$$\mu\{T,\ K_T \geq u\} \leq \frac{\beta}{u^\alpha}.$$

2) *Fixons un point* $x_0 \in E$ *et supposons que la variable aléatoire* $f \to \zeta(T) = d(T(x_0), x_0)$ *est aussi a queue algébrique.*

3) $\displaystyle\int_{\mathbb{L}\mathrm{ip}(E,E)} \log(K_T)\mu(dT) < 0$, *en admettant que cette intégrale peut prendre* $-\infty$

Alors

1. *La chaîne de Markov* $(X_n)_n$ *admet une probabilité stationnaire* ν *unique ;*

2. *Le processus* (H_n) *associé converge p.s. à vitesse géométrique et sa limite ne dépend pas de* x.

Wu et Shao dans [64] ont reprit le travail de Diaconis et Freedman en remplaçant les conditions du théorème 1.8 par des conditions plus générales et montrent la convergence de la suite itérative au sens de la contraction en moyenne géométrique.

Soit (E, d) un espace métrique complet et séparable, de tribu Borélienne $\mathcal{E}$ et $(Y_n)_n$ une suite de variables aléatoires i.i.d de loi μ, à valeurs dans un espace mesurable $(F, \mathcal{F})$. Soit X_0 une variable aléatoire à valeurs dans E et indépendante des Y_n, considérons la chaîne de Markov $(X_n)_{n\geq 0}$ définie par

$$X_0 \text{ et } \forall n \geq 0, \quad X_{n+1} = T(X_n, Y_n) = T_{Y_{n+1}}(X_n),$$

où $T : E \times F \to E$ est une application mesurable par rapport à la tribu $\mathcal{E} \times \mathcal{F}$. Soit $X_0^{'}$ une variable aléatoire indépendante de X_0 et $H_n(x) = T_{Y_1} \circ T_{Y_2} \circ ... \circ T_{Y_n}(x)$.

Définition 1.11. *On dit que X_n est contractante en moyenne géométrique, si il existe $\alpha > 0$, $C = C(\alpha) > 0$ et $r = r(\alpha) \in [0, 1]$ tels que, pour tout $n \in \mathbb{N}$,*

$$\mathbb{E}\big[(d(X_n(X_0^{'}), X_n(X_0)))^{\alpha}\big] \leq Cr^n. \tag{1.6}$$

Condition 1. Il existe $\theta_0 \in E$ et $\alpha > 0$ tels que

$$I(\alpha, \theta_0) := \mathbb{E}\big[(d(\theta_n, T_y(\theta_0)))^{\alpha}\big] = \int_F d(\theta_n, T_y(\theta_0))^{\alpha}\mu(dy) < \infty$$

Condition 2. Il existe $x_0 \in E$, $\alpha > 0$, $r(\alpha) \in [0, 1]$ et $C(\alpha) > 0$ tels que

$$\mathbb{E}\big[(d(X_n(x), X_n(x_0)))^{\alpha}\big] \leq C(\alpha)r(\alpha)^n d(x, x_0)^{\alpha}, \quad \text{pour tous } \ x \in E \ \text{ et } \ n \in \mathbb{N}.$$

Théorème 1.9 (Wu et Shao [64]). *On suppose que les conditions 1 et 2 sont satisfaites. Alors il existe une variable aléatoire H_∞ tels que, pour tout $x \in E$, $H_n(x)$ converge presque sûrement vers H_∞. La limite H_∞ est $\sigma(Y_1, Y_2, ...)-$mesurable et indépendante de x. Par ailleurs, pour tout $n \in \mathbb{N}$*

$$\mathbb{E}\big[(d(H_n(x), H_\infty))^{\alpha}\big] \leq Cr(\alpha)^n,$$

$C > 0$ dépend uniquement de x, x_0, θ_0 et α, et $0 < r(\alpha) < 1$. De plus l'inégalité (1.6) est vérifiée.

Alsmeyer et Fuh dans [1] et Peigné et Woess ([52], [53]) montrent la convergence de la chaîne itérative sous d'autres conditions sur les T_n.

Chapitre 2

Quelques outils sur la théorie des opérateurs linéaires.

Dans ce chapitre, nous rappelons quelques outils sur la théorie des opérateurs linéaires nécessaires pour l'étude des itérées de transformations aléatoires. L'exposé repose sur les articles [28], [29] et [59] et également sur les ouvrages [18], [30] et [47]. Dans toute la suite, $(\mathbf{B}, \|.\|)$ est un espace de Banach sur le corps $\mathbb{C}$ des nombres complexes et P un opérateur linéaire borné de $\mathbf{B}$ dans $\mathbf{B}$. L'opérateur identique sera noté I. La norme de l'opérateur P est notée aussi $\|P\|$. L'espace $\mathbf{B}$ n'est pas nécessairement un espace de fonctions mais nous utilisons la notation f pour un de ses éléments. La restriction de P à un sous-espace G est désignée par $P_{|G}$.

2.1 Généralités

Un opérateur linéaire continu (borné) Π de $\mathbf{B}$ dans $\mathbf{B}$ est une projection dans $\mathbf{B}$ si Π vérifie : $\Pi \circ \Pi = \Pi$.

Si $\Pi : \mathbf{B} \longrightarrow \mathbf{B}$ est une projection alors $I - \Pi$ est aussi une projection dans $\mathbf{B}$ et on a : $\ker \Pi = Im(I - \Pi)$ et $Im\Pi = \ker(I - \Pi)$.

Un sous-espace fermé F de $\mathbf{B}$ est dit complémenté (dans $\mathbf{B}$) s'il existe une projection $\Pi : \mathbf{B} \to \mathbf{B}$ tel que $Im\Pi = F$. C'est le cas si et seulement si, il existe un sous-espace fermé H de $\mathbf{B}$ tel que

$$\mathbf{B} = F \oplus H \ i.e. \ \ \mathbf{B} = F + H \ et \ F \cap H = \{0\}$$

Le sous-espace H (respectivement le sous-espace F) est alors appelé le supplémentaire de F (respectivement de H) dans $\mathbf{B}$.

Un sous espace G de $\mathbf{B}$ est dit P stable (où P invariant) si $P(G) \subset G$.

Définition 2.1. *Un opérateur linéaire borné $P : \mathbf{B} \to \mathbf{B}$ est dit compact si pour tout ensemble bornée $A \subset \mathbf{B}$, son image $P(A)$ est relativement compacte (son adhérence est compacte.)*

Définition 2.2. *Un opérateur borné $P : \mathbf{B} \to \mathbf{B}$ est dit faiblement compact si toute suite bornée $(f_n)_n$ de $\mathbf{B}$ possède une sous-suite $(f_{n_k})_k$ telle que la suite des images (Pf_{n_k}) converge faiblement dans $\mathbf{B}$.*

Définition 2.3 (Lerma et Lasserre [41]). *Soit $(\mathbf{V}, \|.\|_{\mathbf{V}})$ un espace vectoriel normé partiellement ordonné par le cone positif $\mathbf{V}^+ \subset \mathbf{V}$(i.e. pour tous vecteurs f et g dans $\mathbf{V}$, $f \geq g$ si et seulement si $f - g \in \mathbf{V}^+$). Un opérateur linéaire P de $\mathbf{V}$ dans $\mathbf{V}$ est dit :*

- *Positif si $Pf \in \mathbf{V}^+$ pour tout $f \in \mathbf{V}^+$,*
- *contraction si $\|Pf\|_{\mathbf{V}} \leq \|f\|_{\mathbf{V}}$ pour tout $f \in \mathbf{V}$,*
- *préserve la norme si $\|Pf\|_{\mathbf{V}} = \|f\|_{\mathbf{V}}$ pour tout $f \in \mathbf{V}^+$,*
- *Markovien si P est positif et il préserve la norme.*

2.2 Théorie spectrale des opérateurs linéaires

Définitions. *Le spectre de l'opérateur linéaire P est le sous ensemble $\sigma(P)$ de $\mathbb{C}$ défini par*

$$\sigma(P) = \{\lambda \in \mathbb{C}/ \ \lambda I - P \ \text{ est non inversible}\}.$$

Un élément de $\sigma(P)$ est une valeur spectrale de P. Le complémentaire de $\sigma(P)$ est appelé l'ensemble résolvant de P et est noté $\mathcal{R}(P)$. Si $\lambda \in \mathcal{R}(P)$ alors $R(\lambda, P) = (\lambda I - P)^{-1}$ est appelé résolvante de P au point λ.

Le rayon spectral de P est le nombre $\rho(P)$ défini par

$$\rho(P) = \sup\{|\lambda| : \lambda \in \sigma(P)\}$$

et on sait que $\lim_n \|P^n\|^{1/n} = \rho(P)$ *(voir par exemple [18]).*

On dit que λ *est une valeur propre de* P *si* $\lambda I - P$ *n'est pas injective.*

Le sous- espace $\ker(\lambda I - P)$ *est l'espace propre associé à* λ*. Une valeur propre* λ *est dite simple si le sous espace* $\ker(\lambda I - P)$ *est de dimension un.*

Le sous- espace $\bigcup_{k\geq 1} \ker(\lambda I - P)^k$ *est appelé sous espace caractéristique ; la valeur propre* λ *est dite d'indice p, si p est le plus petit entier tel que*

$$\ker(\lambda I - P)^p = \bigcup_{k\geq 1} \ker(\lambda I - P)^k.$$

On appelle spectre ponctuel de P*, et on note* $\sigma_p(P)$*, l'ensemble des valeurs propres de* P*.*
On appelle multiplicité géométrique de la valeur propre λ *d'indice p, la dimension de* $\ker(\lambda I - P)$ *et multiplicité algébrique la dimension de* $\ker(\lambda I - P)^p$*.*

Le rayon spectral essentiel de P *est le plus petit* $\rho > 0$ *tel que le spectre de* P *en dehors du disque de rayon* ρ *ne contienne que des valeurs propres isolées de multiplicité algébrique finie.*

Dans le cas des opérateurs compacts nous avons le théorème suivant.

Théorème 2.1 (Yosida [65]). *Soit* P *un opérateur linéaire compact de* $\mathbf{B}$*. Si* $\lambda_0 \neq 0$ *n'est pas une valeur propre de* P *alors* λ_0 *est dans l'ensemble résolvant de* P*.*

Si P est un opérateur linéaire compact de $\mathbf{B}$, alors $\sigma(P)$ est soit fini soit infini dénombrable ayant 0 comme unique point d'accumulation.

Nous verrons dans la section ci-dessous, que les opérateurs quasi-compacts possèdent des propriétés spectrales aussi importantes.

Le lemme suivant joue un rôle important dans la suite.

Lemme 2.1 (Hennion et Hervé [30]). *Soit* P_0 *un opérateur linéaire borné de* $\mathbf{B}$ *et* $\alpha > \rho(P_0)$*. Il existe deux constantes* $\eta > 0$ *et* $C \in \mathbb{R}$ *telle que, si* $\|P - P_0\| < \eta$ *alors pour tout* $n \in \mathbb{N}$*,* $\|P^n\| \leq C\alpha^n$*. En particulier* $\limsup_{P\to P_0} \rho(P) \leq \rho(P_0)$*.*

Démonstration. Choisissons n_0 tel que $\alpha > \|P_0^{n_0}\|^{1/n_0} \geq \inf_n \|P_0^n\|^{1/n} = \rho(P_0)$. Fixons $\eta > 0$ tel que $\|P - P_0\| < \eta \Rightarrow \|P^{n_0}\| < \alpha^{n_0}$.

Soit $n \in \mathbb{N},\ n = kn_0 + l,\ 0 \leq l < n_0$, on a

$$\|P^n\| \leq \|P^{kn_0}\|\|P\|^l \leq \|P^{n_0}\|^k\|P\|^l \leq \alpha^{kn_0}\|P\|^l \leq \alpha^n\left(\frac{\|P\|}{\alpha}\right)^l \leq \alpha^n\left(\frac{\|P_0\|+\eta}{\alpha}\right)^l \leq C\alpha^n,$$

avec $C = \max\left\{\left(\frac{\|P_0\|+\eta}{\alpha}\right)^l : l = 0, ..., n_0 - 1\right\}$.
Pour $\|P - P_0\| < \eta$, on a $\rho(P) \leq \alpha$, d'où $\limsup_{P \to P_0} \rho(P) \leq \alpha$, la deuxième assertion est obtenue en faisant varier α. □

2.3 Théorie spectrale des opérateurs quasi-compacts

La quasi-compacité est un outil très efficace pour l'étude, dans un cadre aléatoire, des itérées de transformations. Cette notion a été introduite en 1937 par Kryloff et Bogolioboff (cf. [65] ou [47] et leurs références) sous l'appellation quasi-complètement continu dans le but d'étudier les propriétés ergodiques d'un opérateur markovien. Neuveu signale dans [46] l'équivalence entre la quasi-compacité et l'ergodicité forte d'un opérateur markovien. Ionescu Tulcea et Marinescu [35] ont donné une condition suffisante de quasi-compacité. Jointe au fait que cette propriété est stable par perturbation, cela fournit un argument essentiel dans la preuve des théorèmes limites pour les fonctions de certaines chaînes de Markov (cf. Guivarc'h et Raugi [26] et Guivarc'h et Hardy [25].) La quasi-compacité est aussi utile en Analyse (cf. Hervé [33]). Hennion dans [29] et Hervé dans [32] montrent comment la quasi-compacité permet l'étude des fonctions harmoniques pour un noyau markovien sur un espace compact ayant une propriété de quasi-compacité. Nous donnons ci dessous les définitions et propriétés essentielles de la quasi-compacité.

2.3.1 Définitions et propriétés

Définition 2.4. *Un opérateur linéaire borné $P : \mathbf{B} \to \mathbf{B}$ est dit quasi-compact s'il existe une suite $(Q_n)_{n\geq 1}$ d'endomorphismes compacts Q_n sur $\mathbf{B}$ telle que les itérés P^n de P vérifient* $\lim_{n\to\infty} \|P^n - Q_n\| = 0$.

Cette définition est équivalente à la suivante (cf. [16] et [46]).

Définition 2.5. *L'endomorphisme P de $\mathbf{B}$ est quasi compact s'il existe un entier n_0 et un opérateur compact K tels que $\|P^{n_0} - K\| < 1$.*

Ainsi on peut dire qu'un opérateur P de $\mathbf{B}$ dans $\mathbf{B}$ est quasi-compact si à partir d'une certaine puissance, il est assez proche d'un opérateur compact.

Exemple 2.1. Les opérateurs compacts et les opérateurs à puissance compacte sont quasi-compacts.

Soit μ une mesure positive et $(E, \mathcal{E}, \mu)$ un espace mesuré. Un opérateur faiblement compact de $\mathbf{B} = \mathbb{L}^1(E, \mathcal{E}, \mu)$ dans lui même est un opérateur quasi- compact puisque son carré est compact (cf. Dunford et Schwarz [18], page 510).

Proposition 2.1 (**Brunel et Revuz [16]**)**.** *Si l'opérateur borné $P : \mathbf{B} \to \mathbf{B}$ est quasi compact alors l'ensemble des valeurs propres de P n'a aucun point d'accumulation de module ≥ 1 (c'est à dire toutes les valeurs propres λ de P telles que $|\lambda| \geq 1$ sont des points isolés du spectre de P).*

Démonstration. Par hypothèse il existe un entier n_0 et un opérateur compact K tels que $\|P^{n_0} - K\| < \frac{1}{8}$. Il suffit de démontrer ce résultat pour l'opérateur P^{n_0} puisque on a $\sigma(P^{n_0}) = (\sigma(P))^{n_0}$, où $(\sigma(P))^{n_0}$ désigne l'ensemble $\{\lambda^{n_0}, \lambda \in \sigma(P)\}$. Supposons qu'il soit faux. Il existe alors une suite (f_i) dans $\mathbf{B}$ et une suite de scalaires (λ_i) tels que : $(\lambda_i I - P^{n_0})(f_i) = 0, \quad \lambda_i \neq \lambda_j \quad \text{pour} \quad \mathrm{i} \neq \mathrm{j}, \ \lim_{\mathrm{i} \to \infty} \lambda_\mathrm{i} = \lambda \quad \text{avec} \quad |\lambda| \geq 1$. Pour tout i, les éléments $f_1, ...; f_i$ sont linéairement indépendants et nous noterons E_i le sous-espace qu'ils engendrent. E_{i-1} est un sous-espace propre de E_i. D'après le théorème de F. Riesz (cf. [65], page 84), il existe, pour chaque i, $g_i \in E_i$ tel que :

$$\|g_i\| = 1, \ \|g_i - f\| > \frac{1}{2} \ \text{ pour } \ \text{tout } \ f \in E_{i-1}.$$

On a aussi pour tout i,

$$g_i - P^{n_0}(\frac{g_i}{\lambda_i}) \in E_{i-1} \ \text{ et } \ P^{n_0}(\frac{g_j}{\lambda_j}) \in E_{i-1} \ \ si \ \ j < i,$$

et par suite, en écrivant

$$P^{n_0}(\frac{g_i}{\lambda_i})-P^{n_0}(\frac{g_j}{\lambda_j})= g_i - \Big((g_i - P^{n_0}(\frac{g_i}{\lambda_i})) + P^{n_0}(\frac{g_j}{\lambda_j})\Big),$$

on obtient

$$i > j \Rightarrow \|P^{n_0}(\frac{g_i}{\lambda_i})-P^{n_0}(\frac{g_j}{\lambda_j})\| > \frac{1}{2}.$$

Cela entraîne, pour $i > j$,

$$\Big\|K(\frac{g_i}{\lambda_i})-K(\frac{g_j}{\lambda_j})\Big\|+\frac{1}{8}\Big\|\frac{g_i}{\lambda_i} - \frac{g_j}{\lambda_j}\Big\|> \frac{1}{2}.$$

Comme $\|g_i\| = 1$ et $\lambda_i \to \lambda$, l'ensemble des nombres $\|K(\frac{g_i}{\lambda_i})-K(\frac{g_j}{\lambda_j})\|$ pour $i \neq j$ est minoré à partir d'un certain rang par un nombre strictement positif ce qui est en contradiction avec la compacité de l'opérateur K. D'où la proposition. □

Remarque 2.1. Si P est quasi-compact, alors la partie du spectre de P qui n'est pas constituée de valeurs propres de module ≥ 1 est contenue dans un disque centré en 0 de rayon strictement inférieur à 1.

Théorème 2.2 (Brunel et Revuz [16]). *Un opérateur P est quasi-compact si et seulement si il est égal à la somme d'un opérateur de rang fini et d'un opérateur de rayon spectral inférieur à* 1.

Démonstration. Voir Brunel et Revuz [16]. □

Exemple 2.2 (Condition de Doeblin et quasi-compacité.). Nous reprenons ici l'exemple donné dans [30]. Soient $(E, \mathcal{E})$ un espace mesurable et $\mathbf{M}$ l'espace de Banach des fonctions d'ensembles $\sigma-$finies à variation totale bornée sur $(E, \mathcal{E})$. Soit P une probabilité de transition sur $(E, \mathcal{E})$ et on note aussi P l'opérateur de transition agissant sur $\mathbf{M}$ associé à la probabilité de transition P. C'est à dire, P est l'opérateur qui, à toute mesure $\mu \in \mathbf{M}$, associé la mesure $\mu P \in \mathbf{M}$ définie pour $B \in \mathcal{E}$ par :

$$\mu P(B) = \int_E \mu(\mathrm{d}x)P(x, B).$$

On dit que P vérifie la condition de Doeblin s'il existe une mesure de probabilité μ sur $(E, \mathcal{E})$, $n_0 \in \mathbb{N}$ et $\eta > 0$ tels que, si $A \in \mathcal{E}$ vérifie $\mu(A) \leq \eta$, alors

$$\sup_{x \in E} P^{n_0}(x, A) \leq 1 - \eta.$$

Proposition 2.2 (Doob [22]). *Si P est une probabilité de transition vérifiant la condition de Doeblin alors P admet une mesure de probabilité invariante ν. De plus, il existe $0 \leq r < 1$ et $C \geq 0$ tel que, pour tout $n \geq 1$, $\sup_{x \in E} \|P^n(x,.) - \nu(.)\|_{TV} \leq Cr^n$, où $\|\beta\|_{TV}$ est la variation totale de la mesure β.*

La proposition, exprime que si P vérifié la condition de Doeblin, alors l'opérateur P est uniformément ergodique(cf. Lerma et Lasserre[41], définition 9.2.1.).

Si $\mathbf{B}$ est l'espace des fonctions complexes boréliennes et bornées sur $(E, \mathcal{E})$, muni de la norme $\|f\|_\infty = \sup\{|f(x)| : x \in E\}$, alors on a $\mathbf{B} = vect\{\mathbf{1}\} \oplus H$ avec $\mathbf{1}$ est la fonction identiquement égale à 1, $H = \{f : f \in \mathbf{B}, \;\; \nu(f) = 0\}$ et $\rho(P_{|H}) \leq r < 1$.

On a aussi le résultat suivant.

Théorème 2.3 (Fortet [27]). *Si P satisfait à la condition de Doeblin alors, l'opérateur de transition P est quasi- compact.* $\square$

La vérification du théorème 2.2 peut parfois être très délicate. Le théorème suivant est d'une grande utilité pour vérifier la quasi-compacité d'une large classe d'opérateurs.

Théorème 2.4 (Ionescu Tulcea et Marinescu [35]). *Considérons deux espaces de Banach $(\mathbf{B}, \|.\|)$ et $(\mathbf{L}, |.|)$ avec $\mathbf{B} \subset \mathbf{L}$; soit un opérateur $P : \mathbf{B} \to \mathbf{B}$*

Supposons que :

(i) *Si $f_n \in \mathbf{B}$, $f \in \mathbf{L}$, $\lim_n |f_n - f| = 0$, et $\|f_n\| \leq C$ pour tout n, alors $f \in \mathbf{B}$ et $\|f\| \leq C$.*

(ii) $\sup_{n \geq 0} \|P^n\| < \infty$.

(iii) *Il existe deux constantes positives R et $r(0 < r < 1)$, tels que pour tout $f \in \mathbf{B}$:*

$$\|Pf\| \leq r\|f\| + R|f|.$$

(iv) *Si B est un sous-ensemble borné de $(\mathbf{B}, \|.\|)$, alors la fermeture de $P(B)$ est compacte dans $(\mathbf{L}, |.|)$.*

Alors, si toutes ces hypothèses sont vérifiées, P est un opérateur quasi-compact sur $\mathbf{B}$. Il n'y a donc qu'un nombre fini des valeurs propres de module 1 de P, $\lambda_1, \lambda_2, ..., \lambda_p$, et on a

$$P^n = \sum_{i=1}^{i=p} \frac{1}{\lambda_i^n}\Pi_i + S^n \quad n \in \mathbb{N}^\star$$

où Π_i $(i = 1, 2, ..., p)$ sont des endomorphismes de $\mathbf{B}$ de rang fini et bornés et S est un endomorphisme borné de $\mathbf{B}$, tel que

$$\Pi_i^2 = \Pi_i, \quad \Pi_i\Pi_j = 0 \ (i \neq j,) \quad \Pi_i S = S\Pi_i = 0,$$

$\|S\| \leq \dfrac{M}{(1+h)^n}(n \in \mathbb{N}^\star)$ *M et h étant deux constantes positives.*

La condition (iii) est appelée inégalité de Doeblin et Fortet, le théorème 2.4 reste valable si on la remplace par la condition suivante.

Il existe des constantes positives $n_0 > 0$, R et r $(0 < r < 1)$, tels que :

$$\|P^{n_0}f\| \leq r\|f\| + R|f| \text{ pour tout } f \in \mathbf{B}.$$

Démonstration. Pour la démonstration du théorème voir l'article de Ionescu Tulcea et Marinescu [35]. On peut aussi consulter le livre de Norman [47] page 45, pour une autre présentation de ce théorème et sa démonstration. □

Hennion dans [29] donne cette autre définition pour un opérateur quasi-compact.

Définition 2.6. *On dit qu'un opérateur linéaire borné P de $\mathbf{B}$ dans $\mathbf{B}$ est quasi-compact, si l'on a une décomposition de $\mathbf{B}$ en sous espaces fermés P-stables,*

$$\mathbf{B} = F \oplus H,$$

où

F est de dimension finie et $P_{|F}$ n'a que des valeurs propres de module $\rho(P)$ tandis que $\rho(P_{|H}) < \rho(P)$.

Remarque 2.2. Considérons Π_F et Π_H les projections sur F et H respectivement (Π_F et Π_H sont bornées puisque F et H sont fermés). Soit $r \in \mathbb{R}$ tel que $\rho(P_{|H}) < r < \rho(P)$, alors puisque $r > \rho(P_{|H})$ d'après le lemme 2.1 il existe $C \in \mathbb{R}$ tel que pour tout $n \in \mathbb{N}^\star$, $\|P^n - (P_{|F})^n \Pi_F\| \leq Cr^n$. Posons $Q_n = (P_{|F})^n \Pi_F$, Q_n est un opérateur compact car $P_{|F}$ est de rang fini. Comme $(\rho(P))^n = \rho(P^n)$, on déduit que P est quasi-compact au sens de la définition 2.4. De plus l'étude du comportement asymptotique de la suite $(P^n)_n$ se ramène à celle de la suite $((P_{|F})^n \Pi_F)_n$ des itérés d'un endomorphisme d'un espace de dimension finie.

Le sous- espace F introduit dans la définition 2.6 se décrit canoniquement.

Proposition 2.3 (Hennion [29]). *Si P es quasi-compact, alors P n'a qu'un nombre fini de valeurs spectrales de module $\rho(P)$, ce sont les valeurs propres notées $\lambda_1, ..., \lambda_n$ et*

$$F = \bigoplus_{i=1}^{i=n} \bigsqcup_{l \geq 1} \ker(\lambda_i I - P)^l.$$

Démonstration. Voir l'article de Hennion [29], page 2 □

La quasi-compacité peut- être aussi caractérisée par la proposition suivante.

Proposition 2.4 (Hennion [29]). *Si P est un opérateur quasi-compact, alors $P = \Pi + \mathrm{R}$, où Π et R sont des endomorphismes de $\mathbf{B}$ tels que $\Pi\mathrm{R} = \mathrm{R}\Pi = 0$, $\rho(\mathrm{R}) < \rho(P)$; Π est de rang fini, les valeurs propres non nulles de Π coïncident avec les valeurs propres de module $\rho(P)$ de P et les sous-espaces caractéristiques sont identiques.*

La plus part des propriétés spectrales d'un opérateur P envisagées dans les travaux de Hennion ([28, 29]) et Hennion et Hervé [30] sont décrites à l'aide du rayon spectral essentiel que nous définissons ci-dessous.

Définition et Proposition. *Le rayon spectral essentiel de P, $\rho_e(P)$, est la borne inférieure de $\rho(P)$ et des réels $r', r' \geq 0$, pour lesquels il existe des sous-espaces $F_{r'}$ et $H_{r'}$ de* $\mathbf{B}$ *tels que* $\mathbf{B} = F_{r'} \oplus H_{r'},\ P(F_{r'}) \subset F_{r'},\ P(H_{r'}) \subset H_{r'};\ 1 < dim F_{r'} < +\infty$ *et* $P_{|F_{r'}}$ *n'a que des valeurs propres de modules $\geq r'$; H est fermé et $\rho(P_{|H_{r'}}) < r'$.*
Lorsque $\rho_e(P) < \rho(P)$, l'opérateur P est quasi-compact.

La proposition 2.3 se généralise comme suit.

Proposition 2.5 (Hennion et Hervé [30]). *Si P est quasi-compact et si $\rho_e(P) < r' \leq \rho(P)$, alors P n'a qu'un nombre fini de valeurs spectrales de module $\geq r'$, ce sont les valeurs propres notées $\lambda_1, ..., \lambda_m$ et*

$$F_{r'} = \bigoplus_{i=1}^{i=m} \bigsqcup_{l \geq 1} \ker(\lambda_i I - P)^l \quad \text{et} \quad H_{r'} = \{f :\ \limsup_n \|P^n f\| < r'\}.$$

Pour $k \in \mathbb{N}^\star$, on a $\rho_e(P^k) = (\rho_e(P))^k$.

Le théorème 2.6 est une généralisation de théorème de Ionescu Tulcea et Marinescu [35] obtenu par application de la formule de Nussbaum (cf. Hennion [29], page 19) pour le calcul du rayon spectral essentiel. Avant d'énoncer ce théorème rappelons que l'élément de base de la formule de Nussbaum est la fonction d'ensemble γ mesurant le défaut de compacité des parties de $\mathbf{B}$. Pour $r > 0,\ f \in \mathbf{B}$, on pose $\mathrm{B}(f, r) = \{g : \|g - f\| < r\}$ et plus particulièrement, $\mathrm{B}_1 = \mathrm{B}(0, 1)$.

Définition 2.7. *Soit $A \subset \mathbf{B}$, $\gamma(A)$ est la borne inférieure des réels $r, r > 0$, tels qu'il existe un recouvrement fini de A par des boules de $\mathbf{B}$ de rayon r.*

La fonction d'ensemble γ induit une fonction d'opérateur, encore notée γ. $\gamma(P)$ est la borne inférieure de l'ensemble des réels $k,\ k > 0$, tels que, pour tout $A \subset \mathbf{B}$,

$$\gamma(P(A)) < k\gamma(A).$$

Proposition 2.6 (Hennion [29]). $\gamma(P) = \gamma(P(B_1)) < \|P\|$, *la suite $(\gamma(P^n))_n$ est sous-multiplicative et, en posant $\rho_\gamma(P) = \lim_n(\gamma(P^n))^{1/n}$, on a $\rho_\gamma(P) < \rho(P)$.*

Le résultat de Nussbaum est donné par :

Théorème 2.5 (Formule de Nussbaum). $\rho_e(P) = \rho_\gamma(P)$.

Rappelons qu'une partie A de $\mathbf{B}$ est dite précompacte dans $(\mathbf{B}, |.|)$ si, pour tout $\epsilon > 0$, il existe un recouvrement fini de A par des boules $D(f, \epsilon) = \{g : |g - f| < \epsilon\}, f \in \mathbf{B}$.

Théorème 2.6 (Hennion [29]). *Soit* $(\mathbf{B}, \|.\|)$ *un espace de Banach,* $|.|$ *une semi-norme sur* $\mathbf{B}$ *et* P *un opérateur borné sur* $\mathbf{B}$ *telles que*

i) $P(\{f : \ f \in \mathbf{B}, \|f\| \leq 1\})$ *est précompact dans* $(\mathbf{B}, |.|)$,

ii) *il existe une constante* M *telle que, pour tout* $f \in \mathbf{B}, |Pf| \leq M|f|$,

iii) *Il existe* $k \in \mathbb{N}^\star$ *et des réels positif* r *et* R *tels que,*

$r < \rho(P)$ *et, pour tout* $f \in \mathbf{B}, \|P^k f\| \leq R|f| + r^k \|f\|$,

alors P *est quasi-compact et* $\rho_e(P) \leq r$.

Exemple 2.3. Soit u une fonction lipschitzienne sur $E = [0, 1]$, à valeurs positives ou nulles, vérifiant $u(x) + u(\frac{x+1}{2}) = 1$, pour tout $x \in E$. Nous considérons l'opérateur de transition P_u associé à u qui, à f mesurable sur E, fait correspondre $P_u f$ définie par :

$$P_u f(x) = u(\frac{x}{2}) f(\frac{x}{2}) + u(\frac{x+1}{2}) f(\frac{x+1}{2}).$$

Ce type d'opérateur a été introduit et étudié par Conze et Raugi [17].

Soit $\mathbb{L}\mathrm{ip}(E)$ l'espace des fonctions lipschitziennes continues de E dans $\mathbb{C}$ muni de la norme $\|f\|_{\mathbb{L}\mathrm{ip}}$ définie par : $\|f\|_{\mathbb{L}\mathrm{ip}} = |f|_\infty + m(f)$ Où $m(f) = \sup\limits_{\substack{x,y \in E \\ x \neq y}} \dfrac{|f(x) - f(y)|}{|x - y|}$ avec $|f|_\infty$ la norme du supremum. L'opérateur P_u est quasi-compact sur $\mathbb{L}\mathrm{ip}(E)$, puisque les conditions du théorème 2.6 sont vérifiés pour $|.| = |f|_\infty$.

En effet, L'opérateur P_u est positif puisque $\forall f \geq 0, P_u f \geq 0$, de plus on a $P_u \mathbf{1} = \mathbf{1}$ où $\mathbf{1}$ est la fonction identiquement égale à 1, donc pour $f \in \mathbb{L}\mathrm{ip}(E)$, on a $|P_u f|_\infty \leq |f|_\infty$ d'où la condition (ii). Et

$$\begin{aligned} m(P_u f) &\leq \sup_x \big(|f(\frac{x}{2})| + |f(\frac{x}{2})|\big) \frac{m(u)}{2} + \sup_x \big(|u(\frac{x}{2})| + |u(\frac{x}{2})|\big) \frac{f(u)}{2} \\ &\leq |f|_\infty m(u) + \frac{1}{2} m(f), \end{aligned}$$

d'où

$$\|P_u f\|_{\mathbb{L}\mathrm{ip}} \leq (1 + m(u))|f|_\infty + \frac{1}{2}\|f\|_{\mathbb{L}\mathrm{ip}}.$$

Comme $P_u\mathbf{1} = \mathbf{1}$ alors 1 est une valeur propre de P_u donc $\rho(P_u) \geq 1$ et la condition (iii) s'ensuit. Comme l'ensemble $P_u(\{f : \ f \in \mathbb{L}\mathrm{ip}(E), \|f\|_{\mathbb{L}\mathrm{ip}} \leq 1\})$ est équicontinue et borné dans $(\mathbb{L}\mathrm{ip}(E), |.|_\infty)$, alors d'après Ascoli, elle est relativement compact donc précompact dans $(\mathbb{L}\mathrm{ip}(E), |.|_\infty)$.

Puisque 1 est une valeur propre de P_u et $|P_u|_\infty = 1$ alors $\rho(P_u) = 1$.

Remarque 2.3. Le choix de l'espace fonctionnel sur lequel on fait opérer un opérateur linéaire P est très important pour assurer sa quasi-compacité. En effet dans le cas de l'exemple précédent, si $u = \frac{1}{2}$, il a été mentionné dans Hennion [31] que si $\lambda \in \mathbb{C}, |\lambda| \leq 1$ la fonction $f_\lambda(x) = \sum_{n\geq 1} \lambda^{n-1}\cos(2^n \pi x)$ est une fonction propre associée à la valeur propre λ. D'où $P_{\frac{1}{2}}$ n'est pas quasi compact en tant qu'opérateur agissant sur $C_b(E)$ l'espace de toutes les fonctions continues et bornées sur E.

En pratique d'après les travaux de Hennion [28] pour démontrer qu'un opérateur P est quasi-compact, l'un des outils les plus employés consiste à démontrer tout d'abord une inégalité appelée inégalité de Doeblin Fortet donnée dans la définition suivante.

Définition 2.8. *On dit que P a la propriété DF(r), s'il existe une norme $|.|$ sur $\mathbf{B}$ telle que*

(i) *P est compact de $(\mathbf{B}, \|.\|)$ dans $(\mathbf{B}, |.|)$,*

(ii) *pour tout n, $n \in \mathbb{N}$, il existe des réels positifs R_n, r_n, tels que $\liminf_n (r_n)^{1/n} = r < \rho(P)$ et, pour tout $f \in B, \|P^n f\| < R_n|f| + r_n\|f\|$.*

Nous avons la propriété suivante fort utile suivant le théorème 2.5.

Corollaire 2.1 (Hennion [28]). *Si P a la propriété DF(r), alors P est quasi-compact et $\rho_e(P) \leq r$.*

Le résultat de Hennion [28] est une amélioration du théorème de Ionescu Tulcea et Marinescu [35] qui rajoutaient à la propriété DF(r) l'hypothèse $\sup_n \|P^n\| < +\infty$, et concluaient à la quasi- compacité sans donner une majoration de rayon spectral essentiel de P.

2.3.2 Application aux opérateurs positifs

Dans cette section nous rappelons les éléments essentiels étudiés dans Hervé [32] pour déterminer les fonctions propres associé à la valeur propre ρ où ρ est le rayon spectral d'un opérateur positif donné, en se basant sur la notion de compact absorbant. Soit E un espace métrique compact, $\{\eta(x,.), x \in E\}$ une famille de mesures positives sur E telles que, $\sup\limits_{x\in E} \eta(x,E) < \infty$. Soit P l'opérateur positif défini par $Pf(x) = \int_E f(y)\eta(x,dy)$, que nous supposons continu sur $C(E)$ et quasi compact sur l'espace $\mathcal{H}_\alpha(E),\ 0 < \alpha \leq 1$, des fonctions α-Hölder continues de E dans $\mathbb{C}$, défini par :

$$\mathcal{H}_\alpha(E) := \{f \in C(E) \mid m_\alpha(f) < +\infty\}$$

où $m_\alpha(f) := \sup\limits_{\substack{x,y\in E\\ x\neq y}} \dfrac{|f(x) - f(y)|}{d(x,y)^\alpha}$.

On munira $\mathcal{H}_\alpha(E)$ de la norme $\|f\|_\alpha := |f|_\infty + m_\alpha(f)$. Soit K un compact absorbant i.e. $\eta(x; K^c) = 0$ pour tout $x \in K$. Nous dirons qu'il est absorbant minimal s'il n'existe pas de compact absorbants non vide strictement inclus dans K. Pour tout $f \in C(K)$ et tout $x \in K$, nous posons

$$P_K f(x) = \int_K f(x)\eta(x,dy).$$

L'opérateur P_K est borné et positif sur $C(K)$, vérifiant en outre pour tout $f \in C(E)$,

$$P_K(f_{|K}) = (Pf)_{|K}.$$

Il est clair que chaque P_K est quasi compact sur $\mathcal{H}_\alpha(K)$.

Soit ρ (resp. ρ_K) le rayon spectral de P (resp. de P_K). Des propriétés spectrales des opérateurs quasi-compact, on déduit que ρ est une valeur propre de P sur $\mathcal{H}_\alpha(E)$. Soit p_ρ l'indice de ρ pour P. Soit n_c le nombre maximal de compact absorbant disjoints. Si n_c est fini, nous notons $K_1, ..., K_{n_c}$ les n_c compacts absorbant minimaux disjoints.

Théorème 2.7 (Hervé [32]). *Les trois propriétés suivantes sont équivalentes :*

1. *Il existe une fonction $\gamma \in \mathcal{H}_\alpha(E)$ strictement positive sur E telle que $P\gamma = \rho\gamma$.*
2. *On a $p_\rho = 1$, et $\rho_K = \rho$ pour tout compact K absorbant.*
3. *On a $p_\rho = 1, n_c < \infty$, et $\rho_{K_l} = \rho$ pour chaque $l = 1, ..., n_c$.*

Pour la description de sous espace propre associé à la valeur propre ρ on a besoin du théorème suivant.

Théorème 2.8 (Hervé [32]). *Sous l'une quelconque des conditions 1, 2 et 3 du Théorème 2.7, l'espace des fonctions continues sur E solution de l'équation $P\gamma = \rho\gamma$ admet une base $\{\gamma_1, ..., \gamma_{n_c}\}$ formée de fonctions de $\mathcal{H}_\alpha(E)$ positives ou nulles sur E, vérifiant $\gamma_{i_{|K_j}} = 0$ pour tous $i, j \in \{1, ..., n_c\}$ tels que $i \neq j$. En particulier toute fonction continue f telle que $Pf = \rho f$ appartient à $\mathcal{H}_\alpha(E)$.*

Le théorème suivant est utile pour l'étude de l'ensemble des mesures $P-$invariantes.

Théorème 2.9 (Hervé [32]). *On suppose que $\rho = 1$. Sous l'une quelconque des conditions 1, 2 et 3 du théorème 2.7, l'espace des mesures $P-$invariantes admet une base $\{\nu_1, ..., \nu_{n_c}\}$ formée de mesures positives telles que le support de chaque ν_i soit inclus dans K_i.*

Chapitre 3

Chaîne de Diaconis -Freedman.

3.1 Introduction

Nous nous intéressons ici à la chaîne de Markov $(Z_n)_{n\geq 0}$ sur $[0,1]$ introduite par Diaconis et Freedman [20]. Elle est définie en considérant que si à l'instant n elle est en x i.e. $Z_n = x$, à l'instant $n+1$ elle choisit l'un des intervalles $[0,x]$ ou $[x,1]$ avec la probabilité $\frac{1}{2}$ et se déplace en un point y distribué selon la loi uniforme dans l'intervalle ainsi choisi.

Pour $x \in]0,1[$, la probabilité de transition de la chaîne $(Z_n)_{n\geq 0}$ possède alors, selon Diaconis et Freedman [20], une densité $k(x,.)$ par rapport à la mesure de Lebesgue, donnée par

$$\forall y \in]0,1[\qquad k(x,y) = \frac{1}{2} \times \frac{1}{x} \mathbb{1}_{]0,x[}(y) + \frac{1}{2} \times \frac{1}{1-x} \mathbb{1}_{]x,1[}(y).$$

Partant de 0 (resp. de 1), on reste en effet en 0 (resp. en 1) avec la probabilité p (resp. q) ou bien l'on se déplace en un point y distribué selon la loi uniforme sur $]0,1[$.

Diaconis et Freedman dans [20] montrent qu'elle possède une unique mesure de probabilité invariante ν sur $]0,1[$, dont la densité $f_{\frac{1}{2}}$ par rapport à la mesure de Lebesgue est celle de la célèbre loi de l'Arcsinus

$$\forall x \in]0,1[\quad f_{\frac{1}{2}}(x) = \frac{1}{\pi\sqrt{x(1-x)}} \mathbb{1}_{]0,1[}(x).$$

En effet, supposons qu'il existe une mesure stationnaire de densité $f_{\frac{1}{2}}$, La densité $f_{\frac{1}{2}}$ doit alors vérifier l'équation :

$$f_{\frac{1}{2}}(y) = \int_0^1 k(x,y) f_{\frac{1}{2}}(x) \mathrm{d}x = \frac{1}{2} \int_y^1 \frac{f_{\frac{1}{2}}(x)}{x} \mathrm{d}x + \frac{1}{2} \int_0^y \frac{f_{\frac{1}{2}}(x)}{1-x} \mathrm{d}x.$$

Si de plus $f_{\frac{1}{2}}$ est dérivable sur $]0,1[$ on aura,

$$f'_{\frac{1}{2}}y) = -\frac{1}{2}\frac{f_{\frac{1}{2}}(y)}{y} + \frac{1}{2}\frac{f_{\frac{1}{2}}(y)}{1-y},$$

ce qui équivalent à

$$\frac{f'_{\frac{1}{2}}(y)}{f_{\frac{1}{2}}(y)} = \frac{1}{2}(-\frac{1}{y} + \frac{1}{1-y}),$$

d'où

$$f_{\frac{1}{2}}(y) = \frac{1}{\pi\sqrt{y(1-y)}}\mathbb{1}_{]0,1[}(y).$$

Il en est de même lorsque les intervalles $]0,x[$ et $]x,1[$ sont choisis avec les probabilités respectives $p \in]0,1[$ et $q = 1-p$, la mesure invariante dans ce cas admet la densité f_p de la loi Beta $\mathcal{B}(q,p)$ de paramètres q et p donnée par

$$\forall x \in]0,1[\quad f_p(x) = \frac{1}{\Gamma(p)\Gamma(q)}x^{q-1}(1-x)^{p-1}\mathbb{1}_{]0,1[}(x).$$

En effet, f_p vérifie bien l'égalité

$$\begin{aligned} f_p(y) = \int_0^1 k(x,y)f_p(x)\mathrm{d}x &= p\int_y^1 \frac{f_p(x)}{x}\mathrm{d}x + q\int_0^y \frac{f_p(x)}{1-x}\mathrm{d}x \\ &= \frac{1}{\Gamma(p)\Gamma(q)}\big[p\int_x^1 \frac{1}{x^{p+1}(1-x)^q}\mathrm{d}x + q\int_0^x \frac{1}{x^p(1-x)^{q+1}}\mathrm{d}x\big] \end{aligned}$$

car :

$$\begin{aligned} \int_y^1 \frac{p}{x^{p+1}(1-x)^q}\mathrm{d}x + \int_0^y \frac{q}{x^p(1-x)^{q+1}}\mathrm{d}x &= \int_0^{1-y} \frac{p}{x^q(1-x)^{p+1}}\mathrm{d}x + \int_0^y \frac{(1-p)}{x^p(1-x)^{2-p}}\mathrm{d}x \\ &= \big(\frac{1-y}{y}\big)^{1-q} + \big(\frac{y}{1-y}\big)^{1-p} \\ &= \frac{1}{y^p(1-y)^q} \end{aligned}$$

De la description ci-dessus, il ressort que le noyau de transition Q de $(Z_n)_{n\geq 0}$ est donné, pour toute fonction borélienne bornée $\varphi : [0,1] \to \mathbb{C}$ par la formule suivante.

$$Q\varphi(0) = p\varphi(0) + q\int_0^1 \varphi(y)\mathrm{d}y, \qquad Q\varphi(1) = p\int_0^1 \varphi(y)\mathrm{d}y + q\varphi(1)$$

et

$$\forall x \in]0,1[\qquad Q\varphi(x) = \frac{p(x)}{x}\int_0^x \varphi(y)\mathrm{d}y + \frac{q(x)}{1-x}\int_x^1 \varphi(y)\mathrm{d}y.$$

Ceci se réécrit de façon plus concise comme suit : pour tout $x \in [0,1]$

$$Q\varphi(x) = p(x)\int_0^1 \varphi(tx)\mathrm{d}t + q(x)\int_0^1 \varphi(tx+1-t)\mathrm{d}t. \tag{3.1}$$

Cette dernière expression montre que la chaîne $(Z_n)_{n\geq 0}$ s'inscrit bien dans le cadre des itérations de fonctions aléatoires continues. Pour $0 \leq t \leq 1$ notons H_t l'homothétie $x \mapsto tx$ et A_t la transformation affine $x \mapsto tx + 1 - t$, introduisons alors la mesure μ sur l'espace $C([0,1],[0,1])$ des fonctions continues de $[0,1]$ dans $[0,1]$, donnée par

$$\mu(\mathrm{d}T) = p\int_0^1 \delta_{H_t}(\mathrm{d}T)\mathrm{d}t + q\int_0^1 \delta_{A_t}(\mathrm{d}T)\mathrm{d}t$$

où δ_T désigne la masse de Dirac en T. L'égalité (3.1) s'écrit alors

$$Q\varphi(x) = \int_{C([0,1],[0,1])} \varphi(T(x))\mu(\mathrm{d}T).$$

De cette expression, il apparaît clairement que si $(T_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes définies sur un espace de probabilité $(\Omega, \mathcal{T}, \mathbb{P})$ de loi μ sur $C([0,1],[0,1])$, on a $Z_n = T_n \cdots T_1 \cdot Z_0$ i.e. Z_n est le produit de composition des applications $T_1, \cdots, T_n$ appliqué à Z_0 où Z_0 est une variable aléatoire fixée, à valeurs dans $[0.1]$ et indépendante des $T_i, i = 1, \ldots, n$. La chaîne $(Z_n)_{n\geq 0}$ est associée à un système de fonctions itérées. Son étude se fait par celle des propriétés contractantes des applications H_t et $A_t,\ 0 \leq t \leq 1$. Nous nous réfèrerons à $(Z_n)_{n\geq 0}$ par *La chaîne de Diaconis-Freedman.*

Cette chaîne a été considérée par Stoyanov et Pirinsky dans [55], comme modèle de l'histoire triste et bien connue de « l'âne de Buridan » attribué au scientifique et philosophe Buridan(1225- 1358). Un âne dont la faim était égale à sa soif est placé à mi-chemin entre un picotin d'avoine et un seau d'eau. L'âne était libre de décider que faire, commencer à boire ou à manger. Étant très hésitant, il n'arrivait pas à prendre une décision. L'âne ne bougea même pas et mourut immobile dans sa position initiale. Évidemment pour survivre, l'âne devait se diriger dans un temps raisonnable vers le picotin d'avoine ou le seau d'eau.

Une fois son choix de direction fait, soit de façon déterministe (D) ou aléatoire (R), son déplacement à venir se fera aussi de façon aléatoire ou déterministe. Ainsi nous distinguons quatre cas : (DD), (DR), (RD) et (RR). Nous donnons ici quelques détails concernant les différents modèles qui décrivent cette histoire étudiée dans [55] .

On peut décrire les mouvements de l'âne en termes d'une particule A, se déplaçant entre deux points attractifs $A_0 = 0$ et $A_1 = 1$ du segment $[0,\ 1]$, à chaque unité de temps. Soit $(X_n)_n$ la suite des positions successives de la particule. Remarquons que son mouvement se fait en deux temps : D'abord le choix de la direction, ensuite la façon de se déplacer.

Cas (DD). Dans ce cas supposons que la particule A se déplace vers les points attractifs 0 et 1 alternativement (i.e. vers 0 puis 1, 0 puis 1 et ainsi de suite) de façon que la distance entre sa position actuelle et le point attractif se contracte d'un rapport $\lambda \in]0,1[$ i.e. si la particule se trouve en $M_0 = x \in [0,1]$ à un instant n, à l'instant $n+1$ elle se dirige vers $A_0 = 0$ et si M_1 est sa position à cet instant alors on a $A_0M_1 = \lambda M_0A_0$, et à l'instant $n+2$ elle se déplacera vers un point M_2 dans la direction de $A_1 = 1$ de telle sorte que $M_2A_1 = \lambda M_1A_1$. Nous avons la proposition suivante du méme type que celle obtenue dans [55].

Proposition 3.1. *Avec la description ci dessus la suite $(X_n)_n$ est divergente et les deux sous suites $(X_{2n})_n$ et $(X_{2n+1})_n$ sont convergentes et on a*

$$\lim_{n\to+\infty} X_{2n} = \frac{1}{1+\lambda} \quad \text{et} \qquad \lim_{\mathrm{n}\to+\infty} \mathrm{X_{2n+1}} = \frac{\lambda}{1+\lambda}.$$

Démonstration. Supposons qu'à l'instant 0 la particule est en x i.e. $X_0 = x$ et donc $X_1 = \lambda X_0 = \lambda x$, $X_2 = 1-\lambda(1-X_1) = 1-\lambda+\lambda^2X_0, X_3 = \lambda X_2 = (\lambda-\lambda^2+\lambda^3X_0)$ ainsi par récurrence sur n on obtient

$$\forall n \in \mathbb{N},\ X_{2n+2} = 1-\lambda+\lambda^2 X_{2n} \ \text{ et } \ \forall \mathrm{n} \in \mathbb{N}^\star,\ \mathrm{X_{2n+1}} = \lambda-\lambda^2+\lambda^2\mathrm{X_{2n-1}}.$$

Puisque $\lambda \in]0,1[$ alors les applications $f_\lambda : x \mapsto 1-\lambda+\lambda^2 x$ et $g_\lambda : x \mapsto \lambda-\lambda^2+\lambda^2 x$ sont contractantes sur $[0,1]$, par suite les suites $(X_{2n+2})_n$ et $(X_{2n+1})_{n\in\mathbb{N}}$ converges vers les points fixes des applications f_λ et g_λ respectivement. D'où la proposition. $\square$

Cas (DR). Dans ce cas l'étape 1 du mouvement de la particule est déterministe et l'étape 2 est aléatoire.

Précisément, à partir de sa position initiale $x \in [0,1]$ la particule se déplace alternativement dans la direction de 0, de 1, de 0 de 1 etc. vers un point y choisi uniformément dans l'intervalle correspondant en changeant de direction à chaque instant. La suite $(X_{2n})_n$ (respectivement $(X_{2n+1})_n$) est une chaîne de Markov d'opérateur de transition $Q_1 = P_1P_0$ (respectivement $Q_2 = P_0P_1$), où P_0 et P_1 sont deux opérateurs définis pour toute fonction borélienne bornée $\varphi : [0,1] \to \mathbb{C}$ et tout $x \in [0,1]$ par :

$$P_0\varphi(x) = \frac{1}{x}\int_0^x \varphi(y)\mathrm{d}y = \int_0^1 \varphi(tx)\mathrm{d}t \text{ et } P_1\varphi(x) = \frac{1}{1-x}\int_x^1 \varphi(y)\mathrm{d}y = \int_0^1 \varphi(x+t(1-x))\mathrm{d}t.$$

Avec ces hypothèses, nous donnons ci-après, une version d'un résultat de Stoyanov et Pirinsky [55].

Proposition 3.2. *Avec la description ci dessus, la chaîne $(X_n)_n$ est divergente. Cependant les deux chaînes $(X_{2n})_n$ et $(X_{2n+1})_n$ convergent en loi . Nous avons, quand $n \to \infty$,*

$$X_{2n} \to \theta_1 \text{ avec } \theta_1 \rightsquigarrow \mathcal{B}(1,2) \text{ (i.e. } \theta_1 \text{ suit la loi } \mathcal{B}(1,2)) \text{ et } X_{2n+1} \to \theta_2 \text{ avec } \theta_2 \rightsquigarrow \mathcal{B}(2,1).$$

Nous vérifions que la loi Beta $\mathcal{B}(1,2)$ (respect. $\mathcal{B}(2,1)$) est une loi de probabilité invariante sur $[0,1]$ pour la chaîne $(X_{2n})_n$ (respect. $(X_{2n+1})_n$). En effet, on a

$$\begin{aligned} Q_1\varphi(x) &= P_1P_0\varphi(x) = P_1[P_0\varphi(x)] = \frac{1}{1-x}\int_x^1 P_0\varphi(y)\mathrm{d}y \\ &= \frac{1}{1-x}\int_0^1 \varphi(z)\Big[\int_{x\vee y}^1 \frac{1}{y}\Big]\mathrm{d}z \\ &= \frac{-\ln(x)}{1-x}\int_0^x \varphi(z)\mathrm{d}z - \frac{1}{1-x}\int_x^1 \ln(z)\varphi(z)\mathrm{d}z \end{aligned}$$

Considérons l'opérateur Q_1^* sur $\mathbb{L}^1([0,1])$ l'adjoint de Q_1 relativement à la mesure de Lebesgue sur $[0,1]$, défini pour toute fonction $\varphi \in \mathbb{L}^1[0,1], \psi \in \mathbb{L}^\infty[0,1]$, par :

$$\int_0^1 \varphi(x)Q_1\psi(x)\mathrm{d}x = \int_0^1 Q_1^*\varphi(x)\psi(x)\mathrm{d}x.$$

Un calcul direct montre que l'on peut définir $Q_1^*\varphi$, pour tout $x \in [0,1]$, par
$$Q_1^*\varphi(x) = \int_x^1 \frac{-\varphi(t)\ln(t)}{1-t}\mathrm{d}t - \int_0^x \frac{\varphi(t)\ln(x)}{1-t}\mathrm{d}t.$$
La fonction $x \mapsto 2(1-x)\mathbb{1}_{[0,1]}(x)$ est bien une solution de l'equation $Q^*\varphi = \varphi$. Ainsi la mesure de densité $x \mapsto 2(1-x)\mathbb{1}_{[0,1]}(x)$ est une mesure Q_1- invariante.

Cas (RD). Dans ce cas, si la particule est on $x \in [0,1]$ à l'instant n, elle se dirige à l'instant $n+1$ vers 0 (respectivement 1), avec la probabilité p (resp. q), et elle se déplace de tel sorte que la distance x (resp. $1-x$) se réduit avec un rapport $\lambda \in]0,1[$. La suite $(X_n)_n$ est une chaîne de Markov d'opérateur de transition Q défini pour toute fonction borélienne bornée $\varphi : [0,1] \to \mathbb{C}$ et tout $x \in [0,1]$ par :
$$Q\varphi(x) = p\varphi(\lambda x) + q\varphi(1-\lambda(1-x)).$$

Remarquons que dans le cas où $\lambda = 1/2$, $(X_n)_{n\geq 0}$ est la chaîne discutée dans le chapitre 1 (voir exemple 1.1). Nous avons montré qu'elle admet une unique loi invariante. Nous vérifions que dans le cas particulier où $\lambda = 1/2$ et $p = q = 1/2$ cette mesure invariante est la loi uniforme. Considérons l'opérateur Q^* sur $\mathbb{L}^1([0,1])$, l'adjoint de Q relativement à la mesure de Lebesgue sur $[0,1]$, défini pour toute fonction $\varphi \in \mathbb{L}^1[0,1], \psi \in \mathbb{L}^\infty[0,1]$, par l'égalité suivante.
$$\int_0^1 \varphi(x)Q\psi(x)\mathrm{d}x = \int_0^1 Q^*\varphi(x)\psi(x)\mathrm{d}x.$$
Un calcul direct montre que l'on peut définir $Q^*\varphi$ pour tout $x \in [0,1]$, par :
$$Q^*\varphi(x) = \frac{p}{\lambda}\varphi\left(\frac{x}{\lambda}\right)\mathbb{1}_{[0,\lambda]}(x) + \frac{q}{1-\lambda}\varphi\left(\frac{x-\lambda}{1-\lambda}\right)\mathbb{1}_{[\lambda,1]}(x).$$
Dans le cas où $\lambda = 1/2$ et $p = q = 1/2$, la fonction $x \mapsto f(x) = \mathbb{1}_{[0,1]}(x)$ est bien une solution de l'equation $Q^*\varphi = \varphi$. Ainsi la mesure de densité f par rapport à la mesure de Lebesgue, est une mesure invariante pour la chaîne $(X_n)_n$.

Cas (RR). Dans la cas où les deux étapes de mouvement de la particule sont aléatoires, la chaîne $(X_n)_{n\geq 0}$ obtenue est la chaîne de Diaconis-Freedman introduite au début de cette introduction, qui fera l'objet de notre étude.

Nous renvoyons aussi le lecteur à [11], [49] et [55] pour une généralisation de la chaîne de Diaconis-Freedman au cas ou elle admet toujours une loi stationnaire de type Beta.

Diaconis et Freedman évoquent aussi dans leur travail le cas où les poids p et q sont autorisés à varier en fonction de la position x ; dans ce cas la chaîne rentre dans le cadre des systèmes d'itérations de transformations aléatoires contractantes avec des probabilités de transitions dépendantes de la position spatiale de la chaîne. Une vaste littérature existe sur le sujet on citera par exemple les travaux de Kaijser [36] au début des années 80, Barnsley et al [5], [4] et [3] quelques années plus tard et des développements récents par Kapika et Seleczka [37]. Une des méthodes les plus efficaces pour l'étude de tels systèmes repose sur un argument de couplage voir Kapika et Seleczka [37] . L'approche par la théorie spectrale nécessite des hypothèses un peu plus restrictives mais s'avère très fructueuse pour décrire le comportement de la chaîne en présence de plusieurs mesures invariantes. Ce phénomène est très riche à étudier, il échappe au cadre des travaux évoqués ci-dessus.

Dans le cas de la chaîne de Diaconis-Freedman et lorsque $p(x) = x$, les variables $Z_1, Z_2, ...$ sont indépendantes et uniformément distribuées sur $[0,1]$ (cf. [20]). Le cas où $p(x) = 1-x$ est intéressant : la suite $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-p.s. vers une variable aléatoire Z_∞ à valeur dans $\{0,1\}$(cf. [20]). Dans ce qui suit, nous verrons que ces deux cas s'inscrivent dans une étude plus systématique de la situation générale abordée par les deux auteurs. Nous adoptons l'approche décrite dans [50] et [30] par la théorie des opérateurs quasi-compact ; elle s'appuie sur la description du spectre périphérique de l'opérateur de transition Q, en utilisant un contrôle précis de l'action du semi-groupe généré par les transformations (aléatoires) du système et la description de ses ensembles invariants.

Nous terminons cette introduction par une présentation d'un résultat partiel sur l'un des modèles de l'histoire de l'âne de Buridan proposé dans [55]. Soit $\lambda \in]0,1[$ un nombre réel fixe. Si l'âne est en $x \in [0,1]$ à l'instant n, il se dirige à l'instant $n+1$ vers 0 (respectivement 1), avec la probabilité $p(x)$ (resp. $q(x)$), et il se déplace de tel sorte que la distance x (resp. $1-x$) se réduit avec une proportion λ. En d'autre terme, la suite $(Z_n)_{n\geq 0}$ des positions successives de l'âne est une chaîne de Markov sur $[0,1]$ d'opérateur de transition Q_λ donné, pour toute fonction borélienne bornée $\varphi : [0,1] \to \mathbb{C}$ et tout $x \in [0,1]$ par :

$$Q_\lambda \varphi(x) = p(x)\varphi(\lambda x) + q(x)\varphi(1 - \lambda(1-x)).$$

Le cas où $\lambda = 1/2$ a été considéré par Hervé dans [32]. Il y explicite en détail le spectre de la restriction de $Q_{1/2}$ à un espace de Banach convenablement choisi en fonction de la régularité des fonctions p et q. L'auteur décrit soigneusement les sous ensembles invariants de $[0,1]$ sous l'action de semi-groupe des contractions engendré par les deux applications $x \mapsto \frac{x}{2}$ et $x \mapsto \frac{1+x}{2}$. Cela présente un grand intérêt, en particulier pour les applications en Analyse Mathématique (cf. Hervé [33]).

3.2 Approche par la théorie des opérateurs pour les systèmes d'itération de fonctions aléatoires

Dans cette section nous présentons quelques éléments sur l'approche par la théorie des opérateurs pour étudier les systèmes d'itération de fonctions contractantes. Nous présentons dans un contexte assez général les outils et méthodes basées sur la théorie des opérateurs. Les deux propositions suivantes concernent l'unicité de la mesure invariante pour de tels systèmes, mais ce ne sont pas de nouveaux résultats. Néanmoins cette approche est très utile pour l'étude de comportement de la chaîne de Diaconis-Freedman que nous verrons dans la section suivante. Par souci de clarté, nous détaillons leurs démonstrations. Dans la sous rubrique 3.2.1, nous considérons le cas où les fonctions qui définissent notre système d'itération de fonctions aléatoires forment une suite de variables aléatoires indépendantes et identiquement distribuées, dans la sous rubrique suivante nous étendons cette approche, au cas des systèmes d'itérations de transformations aléatoires lipschitziennes avec des probabilités de transitions qui dépendent du point où se trouve la chaîne. Une première application de ces propositions à la chaîne de Diaconis-Freedman sera donnée dans le paragraphe 3.3.1. Néanmoins, l'approche utilisée pour les prouver est flexible et fonctionne bien, pour certains systèmes spécifiques d'itérations de fonctions aléatoires comme la chaîne de Diaconis-Freedman, même s'il existe plusieurs mesures de probabilité invariantes.

Soit (E,d) un espace métrique compact et on note $C(E,E)$ l'espace des fonctions continues de E dans E muni de la norme de la convergence uniforme $|\cdot|_\infty$.

Soit $(T_n)_{n\geq 1}$ une suite de fonctions aléatoires continues de E dans E de loi μ. Le cas où les fonctions T_n sont lipschitziennes de E dans E est très utile en particulier pour l'utilisation d'un argument spectral, basé sur les propriétés de contraction du semi-groupe T_μ engendré par le support de μ.

3.2.1 Cas de fonctions aléatoires lipschitziennes indépendantes et identiquement distribuées

Soit $\mathbb{L}\mathrm{ip}(E,E)$ l'espace de Banach des fonctions lipschitziennes continues de E dans E i. e. l'espace des fonction $f : E \to E$ tel que

$$[f] = \sup_{\substack{x,y\in E\\ x\neq y}} \frac{d(f(x),f(y))}{d(x,y)} < \infty,$$

et on munira $\mathbb{L}\mathrm{ip}(E,E)$ de la norme $\|\cdot\| = |\cdot|_\infty + [\cdot]$.

On note $\mathcal{B}$ la tribu des boréliens de $\mathbb{L}\mathrm{ip}(E,E)$ et l'on fixe une mesure de probabilité μ sur $(\mathbb{L}\mathrm{ip}(E,E),\mathcal{B})$. Soit $(T_n)_{n\geq 1}$ une suite de fonctions aléatoires indépendantes et de même loi μ définie sur un espace de probabilité $(\Omega,\mathcal{F},\mathbb{P})$ et à valeurs dans $(\mathbb{L}\mathrm{ip}(E,E)$. On considère la chaîne de Markov sur E définie par

$$X_0 \ \text{ et } \ \forall \mathrm{n} \geq 0 \quad \mathrm{X_{n+1}} := \mathrm{T_{n+1}}(\mathrm{X_n}) = \mathrm{T_{n+1}} \cdot \mathrm{X_n}$$

où X_0 est une variable aléatoire fixée à valeurs dans E et indépendantes des T_n. La suite $(X_n)_{n\geq 0}$ est une chaîne de Markov d'espace des états E et de noyau de transition P défini pour toute fonction borélienne bornée $\varphi : E \to \mathbb{C}$ et tout $x \in E$, par

$$P\varphi(x) = \int_{\mathbb{L}\mathrm{ip}(E,E)} \varphi(T(x))\mu(\mathrm{d}T).$$

La chaîne $(X_n)_{n\geq 0}$ a la propriété de Feller i.e. l'opérateur P agit sur $C(E)$ l'espace des fonctions continues de E dans $\mathbb{C}$. Les transformations T_n étant lipschitziennes sur E, l'opérateur

P agit aussi sur l'espace des fonctions lipschitziennes de E dans $\mathbb{C}$ et plus généralement sur l'espace $\mathcal{H}_\alpha(E), 0 < \alpha \leq 1$, des fonctions α-Hölder continues de E dans $\mathbb{C}$, défini par :

$$\mathcal{H}_\alpha(E) := \{f \in C(E) \mid \|f\|_\alpha := |f|_\infty + m_\alpha(f) < +\infty\}$$

où $m_\alpha(f) := \sup\limits_{\substack{x,y\in E \\ x\neq y}} \dfrac{|f(x) - f(y)|}{d(x,y)^\alpha}$.

Rappelons que $\mathcal{H}_\alpha(E)$ muni de la norme $\|\cdot\|_\alpha$ est un espace de Banach et que l'injection canonique de $\mathcal{H}_\alpha(E)$ dans l'espace $C(E)$ muni de la norme de la convergence uniforme $|\cdot|_\infty$ est compacte. Le comportement asymptotique de la chaîne $(X_n)_{n\geq 0}$ est lié au spectre de P sur ces espaces ; sous l'hypothèse de "contraction en moyenne" des T_n, la restriction de l'opérateur P à $\mathcal{H}_\alpha(E)$ vérifie certaines propriétés spectrales. C'est le point capital pour contrôler le spectre de P. Nous avons la proposition suivante.

Proposition 3.3. *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires à valeurs dans E associée à un système d'itération de fonctions de loi μ sur $\mathbb{L}\mathrm{ip}(E,E)$. Supposons qu'il existe $\alpha \in]0,1]$ tel que*

$$r := \sup_{\substack{x,y\in E \\ x\neq y}} \int_{\mathbb{L}\mathrm{ip}(E,E)} \Big(\frac{d(T(x),T(y))}{d(x,y)}\Big)^\alpha \mu(\mathrm{d}T) < 1. \tag{3.2}$$

Alors il existe sur E une unique mesure de probabilité P-invariante ν. De plus, il existe des constantes $\kappa > 0$ et $\rho \in]0,1[$ telles que, pour toute fonction $\varphi \in \mathcal{H}_\alpha(E)$, on ait

$$\forall x \in E \quad |P^n\varphi(x) - \nu(\varphi)| \leq \kappa\rho^n. \tag{3.3}$$

Démonstration. Par souci de clarté avec en vue la généralisation qui va suivre, nous détaillons ici la démonstration ; nous renvoyons le lecteur à [20] lorsque E n'est pas compact. L'opérateur P est markovien sur $C(E)$, son rayon spectral $\rho_\infty(P)$ sur cet espace est donc égal à 1. Malheureusement, sans hypothèse supplémentaire, le contrôle du spectre de P sur $C(E)$ est très délicat (voir par exemple [28]) ce qui explique pourquoi on s'intéresse plutôt à la restriction de P au sous-espace $\mathcal{H}_\alpha(E)$ de $C(E)$.

L'opérateur P agit aussi sur $\mathcal{H}_\alpha(E)$ puisque pour toute fonction $\varphi \in \mathcal{H}_\alpha(E)$ on a

$$m_\alpha(P\varphi) \leq r\ m_\alpha(\varphi), \tag{3.4}$$

d'où en particulier

$$\forall \varphi \in \|P\varphi\|_\alpha \leq r\|\varphi\|_\alpha + |\varphi|_\infty. \tag{3.5}$$

Cette inégalité, nous permet d'appliquer le théorème de Ionescu-Tulcea et Marinescu sur les opérateurs quasi-compacts. D'après les travaux de Hennion (voir chapitre 3, proposition 2.5) on déduit que le rayon spectral essentiel de P sur $\mathcal{H}_\alpha(E)$ est inférieur ou égal à r. Autrement dit, les valeurs spectrales de P de module $> r$ sont en nombre fini et ce sont des valeurs propres isolées de multiplicité finie dans le spectre de P.

Pour démontrer la proposition 3.3, il nous suffit de contrôler le spectre périphérique de P sur $\mathcal{H}_\alpha(E)$. Soit λ une valeur propre de P de module 1, de fonction propre associée f. L'égalité $P^n f = \lambda^n f$ satisfaite pour tout $n \geq 1$ entraîne

$$m_\alpha(f) = m_\alpha(\lambda^n f) = m_\alpha(P^n f) \leq r^n m_\alpha(f)$$

d'où $m_\alpha(f) = 0$, puisque $0 \leq r < 1$. La fonction f est donc constante sur E et l'on a $\lambda = 1$. La restriction de P à $\mathcal{H}_\alpha(E)$ admet donc 1 comme unique valeur propre de module 1 et cette valeur propre est simple. L'opérateur P sur $\mathcal{H}_\alpha(E)$ se décompose donc sous la forme

$$P = \Pi + \mathrm{R} \tag{3.6}$$

où

1. Π désigne le projecteur de $\mathcal{H}_\alpha(E)$ sur l'espace propre $\mathbb{C} \cdot \mathbf{1}$ associé à la valeur propre 1,
2. R est un opérateur de rayon spectral $\rho < 1$,
3. $\Pi\mathrm{R} = \mathrm{R}\Pi = 0$.

En particulier, pour toute $\varphi \in \mathcal{H}_\alpha(E)$, la suite $(P^n\varphi)$ converge vers $\Pi(\varphi)\mathbf{1}$. Ainsi il existe une unique mesure de probabilité P-invariante ν sur E et le projecteur Π peut s' écrire sous la forme $\Pi : \varphi \mapsto \nu(\varphi)\mathbf{1}$. L'inégalité (3.3) se déduit alors de l'égalité (3.6) et des propriétés spectrales de P sur $\mathcal{H}_\alpha(E)$. $\square$

Application à la chaîne de Diaconis-Freedman pour p constant dans $[0,1]$.

Soit $(Z_n)_{n\geq 0}$ la chaîne sur $[0,1]$ décrite dans l'introduction de ce chapitre. Soit $\mathbb{L}\mathrm{ip}([0,1],[0,1])$ l'espace de Banach des fonctions lipschitziennes de $[0,1]$ dans $[0,1]$.

L'inégalité (3.2) est satisfaite avec $r = \frac{1}{1+\alpha}$ puisque dans ce cas on a $[H_t] = [A_t] = t$ pour tout $0 \leq t \leq 1$. D'où

$$\begin{aligned}\sup_{\substack{x,y\in[0,1]\\ x\neq y}} \int_{\mathbb{L}\mathrm{ip}([0,1],[0,1])} \Big(\frac{d(T(x),T(y))}{d(x,y)}\Big)^{\alpha} \mu(\mathrm{d}T) &\leq p\int_0^1 [H_t]^{\alpha}\mathrm{d}t + q\int_0^1 [A_t]^{\alpha}\mathrm{d}t \\ &= \int_0^1 t^{\alpha}\mathrm{d}t = \frac{1}{1+\alpha}.\end{aligned}$$

Ainsi, la chaîne $(Z_n)_{n\geq 0}$ admet une unique mesure de probabilité invariante sur $[0,1]$, qui est celle associé à la loi Beta $\mathcal{B}(q,p)$.

□

3.2.2 Cas de fonctions aléatoires lipschitziennes avec dépendance spatiale des probabilités de transition.

On remplace ici la mesure μ par une collection $(\mu_x)_{x\in E}$ de mesures de probabilité sur $\mathbb{L}\mathrm{ip}(E,E)$ et l'on considère la chaîne de Markov $(X_n)_{n\geq 0}$ sur E dont le noyau de transition P est donné pour toute fonction borélienne bornée $\varphi : E \to \mathbb{C}$ et tout $x \in E$, par

$$P\varphi(x) = \int_{\mathbb{L}\mathrm{ip}(E,E)} \varphi(T(x))\mu_x(\mathrm{d}T).$$

Avant de proposer un énoncé général, nous introduisons les définitions suivantes.

Définition 3.1. *Une suite $(\xi_n)_{n\geq 0}$ d'éléments de $\mathbb{L}\mathrm{ip}(E,E)$ est dite contractante s'il existe $x_0 \in E$ tel que*

$$\forall x \in E \qquad \lim_{n\to+\infty} \xi_n(x) = x_0.$$

Définition 3.2. *La variation totale d'une mesure signée ν sur l'espace $\mathbb{L}\mathrm{ip}(E,E)$ est la fonction d'ensemble notée $\mid \nu \mid$ donnée par $\mid \nu \mid = \sup_{\mathcal{P}} \sum_{i=1}^{i=k} \mid \nu(B_i) \mid$, où le supremum est pris pour toutes les partitions $\mathcal{P} = (B_1, B2, \cdots, B_k)$ de $\mathbb{L}\mathrm{ip}(E,E)$ en un nombre fini de sous-ensembles mesurables disjoints.*

La variation totale $|\nu|$ *est aussi égale à* $\sup_{\phi} |\int \phi(T)\nu(\mathrm{d}T|$, *le supremum est pris pour toutes les fonctions mesurables* ϕ *définies de* $\mathbb{L}\mathrm{ip}(E,E)$ *dans* $\mathbb{R}$ *tel que* $|\phi| \leq 1$.

On a la proposition suivante.

Proposition 3.4. *Supposons qu'il existe* $0 < \alpha \leq 1$ *tel que*

H1. $r := \sup_{\substack{x,y\in E\\ x\neq y}} \int_{\mathbb{L}\mathrm{ip}(E,E)} \Big(\frac{d(T(x),T(y))}{d(x,y)}\Big)^{\alpha} \mu_x(\mathrm{d}T) < 1;$

H2. $R_\alpha := \sup_{\substack{x,y\in E\\ x\neq y}} \frac{|\mu_x - \mu_y|}{d(x,y)^\alpha} < +\infty.$

et que de plus

H3. *Il existe* $\delta > 0$ *et une mesure de probabilité* μ *sur* E *telle que*

$$\forall x \in E \qquad \mu_x \geq \delta\mu; \tag{3.7}$$

et dont le support S_μ *engendre un semi-groupe* T_μ *qui possède une suite contractante.*

Il existe alors sur E *une unique mesure de probabilité* P*-invariante* ν. *De plus, il existe des constantes* $\kappa > 0$ *et* $\rho \in]0,1[$ *telles que, pour toute fonction* $\varphi \in \mathcal{H}_\alpha(E)$, *on ait*

$$\forall x \in E \quad |P^n\varphi(x) - \nu(\varphi)| \leq \kappa\rho^n. \tag{3.8}$$

Pour démontrer cette proposition, nous utilisons les deux lemmes suivants, valides sous les hypothèses **H1**, **H2** et **H3**.

Lemme 3.1. *Soit* $h \in \mathcal{H}_\alpha(E)$ *telle que* $Ph = h$. *Pour tout* $x \in E$, *la suite* $(h(X_n))_{n\geq 0}$ *est une martingale bornée sur l'espace* $(\Omega, \mathcal{F}, \mathbb{P}_x)$, *où* $\mathbb{P}_x$ *est la probabilité conditionnelle* $\mathbb{P}(.|X_n = x)$. *Elle converge* $\mathbb{P}_x$*-presque sûrement et dans* $\mathbb{L}^1(\Omega, \mathbb{P}_x)$ *vers une variable aléatoire* H^∞ *et l'on a*

$$\forall n \geq 0 \quad h(x) = \mathbb{E}_x(h(X_n)) = \mathbb{E}_x(H^\infty). \tag{3.9}$$

De plus, pour tout $\xi \in T_\mu$, *on a*

$$H^\infty = \lim_{n\to+\infty} h(\xi \cdot X_n) \quad \mathbb{P}_x - \text{p.s.} \tag{3.10}$$

Démonstration du lemme 3.1. La première assertion et l'égalité (3.9) sont immédiates puisque h est P-harmonique et bornée.

En effet, pour tout $n \in \mathbb{N}$, si $\mathcal{F}_n$ est la filtration naturelle, nous avons :

$$\begin{aligned} \mathbb{E}_x(h(X_{n+1})|\mathcal{F}_n) &= \mathbb{E}_x(h(X_{n+1})|X_n) = Ph(X_n) \ \ \text{(propriété de Markov)} \\ &= h(X_n) \ \ \text{(invariance de } h \text{ par } P), \end{aligned}$$

la suite $((h(X_n))_{n\geq 0}$ est une martingale relativement à $(\mathcal{F})_{n\geq 0}$ bornée par $|h|_\infty$. D'après le théorème de convergence des martingales, la suite $(h(X_n))_{n\geq 0}$ converge $\mathbb{P}_x$-presque sûrement et elle est fermée par sa limite c'est à dire $h(X_n) = \mathbb{E}_x(H^\infty/\mathcal{F}_n) \quad \mathbb{P}_x - \text{p.s}$ donc $\mathbb{E}_x(h(X_n)) = \mathbb{E}_x(H^\infty)$.

D'autre part nous avons $\mathbb{E}_x(h(X_n)) = P^n h(x) = h(x)$ d'où l'égalité (3.9)

Pour établir l'égalité (3.10), on pose $u_{n,q}(x) := \mathbb{E}_x\Big(\big|h(X_{n+q}) - h(X_n)\big|^2\Big)$ pour tous entiers positifs n et q. Pour tout $N \geq 1$, en utilisant les propriétés de martingales nous avons

$$\begin{aligned} \sum_{n=1}^N u_{n,q}(x) &= \sum_{n=1}^N \mathbb{E}_x\Big(\big(h(X_{n+q}) - h(X_n)\big)\big(\overline{h(X_{n+q}) - h(X_n)}\big)\Big) \\ &= \sum_{n=1}^N \Big[\mathbb{E}_x\Big(\big|h(X_{n+q})\big|^2\Big) - \mathbb{E}_x\Big(h(X_{n+q})\overline{h(X_n)}\Big) \\ &- \mathbb{E}_x\Big(h(X_n)\overline{h(X_{n+q})}\Big) + \mathbb{E}_x\Big(\big|h(X_n)\big|^2\Big)\Big] \\ &= \sum_{n=1}^N \mathbb{E}_x\Big(\big|h(X_{n+q})\big|^2\Big) - \sum_{n=1}^N \mathbb{E}_x\Big(\big|h(X_n)\big|^2\Big) \leq 2q\big|h\big|_\infty. \end{aligned}$$

Puisque on a

$$\begin{aligned} \mathbb{E}_x\Big(h(X_{n+q})\overline{h(X_n)}\Big) &= \mathbb{E}_x\Big(\mathbb{E}_x\big(h(X_{n+q})\overline{h(X_n)}|\mathcal{F}_n\big)\Big) \\ &= \mathbb{E}_x\Big(\overline{h(X_n)}\mathbb{E}_x\big(h(X_{n+q})|\mathcal{F}_n\big)\Big) \ \ \text{car } \overline{\mathrm{h(X_n)}} \text{ est } \mathcal{F}_\mathrm{n} - \text{mesurable} \\ &= \mathbb{E}_x\Big(\overline{h(X_n)}h(X_n)\Big) \ \ \text{car } \overline{\mathrm{h(X_n)}} \text{ est une martingale} \\ &= \mathbb{E}_x\Big(\big|h(X_n)\big|^2\Big). \end{aligned}$$

De même on a $\mathbb{E}_x\Big(h(X_n)\overline{h(X_{n+q})}\Big) = \mathbb{E}_x\Big(\big|h(X_n)\big|^2\Big)$.

Par conséquent $\sum_{n\geq 1} u_{n,q}(x) < +\infty$ et

$$\sum_{n\geq 1} \mathbb{E}_x\Big(\int_{\mathbb{L}\mathrm{ip}(E,E)^q} \big|h(T_q\cdots T_1\cdot X_n) - h(X_n)\big|^2 \mu_{X_n}(\mathrm{d}T_1)\cdots\mu_{T_{q-1}\cdots T_1\cdot X_n}(\mathrm{d}T_q)\Big) < +\infty \quad (3.11)$$

d'où, en utilisant l'hypothèse **H3**

$$\sum_{n\geq 1} \mathbb{E}_x\Big(\int_{\mathbb{L}\mathrm{ip}(E,E)^q} \big|h(T_q\cdots T_1\cdot X_n) - h(X_n)\big|^2 \mu(\mathrm{d}T_1)\cdots\mu(\mathrm{d}T_q)\Big) < +\infty$$

soit

$$\int_{\mathbb{L}\mathrm{ip}(E,E)^q} \mathbb{E}_x\Big(\sum_{n\geq 1} \big|h(T_q\cdots T_1\cdot X_n) - h(X_n)\big|^2\Big)\mu(\mathrm{d}T_1)\cdots\mu(\mathrm{d}T_q) < +\infty.$$

La suite $\big(h(T_q\cdots T_1\cdot X_n) - h(X_n)\big)_{n\geq 1}$ converge donc vers 0 pour $\mathbb{P}_x$-presque tout $\omega\in\Omega$, tout $q\geq 1$ et $\mu^{\otimes q}$-presque tout $T_1, T_2, \cdots, T_q$. On conclut par un argument de densité (au sens de la norme $\|\cdot\|_\alpha$). $\square$

De même, on peut démontrer le lemme suivant.

Lemme 3.2. *Soit $\phi\in\mathcal{H}_\alpha(E)$ telle que $P\phi = \lambda\phi$ où λ est un nombre complexe de module 1. Pour tout $x\in E$, la suite $(\lambda^{-n}\phi(X_n))_{n\geq 0}$ est une martingale bornée sur l'espace $(\Omega, \mathcal{F}, \mathbb{P}_x)$; elle converge $\mathbb{P}_x$-presque sûrement et dans $\mathbb{L}^1(\Omega, \mathbb{P}_x)$ vers une variable aléatoire Φ^∞ et l'on a*

$$\forall n\geq 0 \quad \phi(x) = \mathbb{E}_x(\lambda^{-n}\phi(X_n)) = \mathbb{E}_x(\Phi^\infty) \quad (3.12)$$

De plus, pour tout $q\geq 1$ et toutes transformations $T_1,\cdots,T_q$ du support S_μ de μ, on a

$$\Phi^\infty = \lim_{n\to+\infty} \lambda^{-(n+q)}\phi(T_q\cdots T_1\cdot X_n) \quad \mathbb{P}_x-\text{p.s.} \quad (3.13)$$

Démonstration de la proposition 3.4. L'opérateur P est markovien, il agit donc sur $C(E)$, son rayon spectral sur cet espace égal à 1. L'opérateur P agit aussi sur $\mathcal{H}_\alpha(E)$ puisque pour toute fonction $\varphi\in\mathcal{H}_\alpha(E)$ et tous $x, y\in E$, on a

$$|P\varphi(x) - P\varphi(y)| \leq \int_{\mathbb{L}\mathrm{ip}(E,E)} |\varphi(T(x)) - \varphi(T(y))|\mu_x(\mathrm{d}T) + |\varphi|_\infty \int_{\mathbb{L}\mathrm{ip}(E,E)} |\mu_x - \mu_y|(\mathrm{d}T)$$

Il vient

$$m_\alpha(P\varphi) \leq r m_\alpha(\varphi) + R_\alpha|\varphi|_\infty \quad (3.14)$$

si bien que la restriction de P à l'espace $\mathcal{H}_\alpha(E)$ vérifie l'inégalité dite de *Doeblin-Fortet*

$$\|P\varphi\| \leq r\|\varphi\|_\alpha + (1+R_\alpha)|\varphi|_\infty. \tag{3.15}$$

C'est un opérateur quasi-compact sur $\mathcal{H}_\alpha(E)$, de rayon spectral 1. Il suffit donc de contrôler le spectre périphérique de P (i.e. l'ensemble des valeurs propres λ de P de module 1). L'argument diffère alors de celui de la démonstration de la proposition 3.3, la propriété de contraction (3.4) n'étant pas satisfaite ici et étant remplacée par (3.14).

Dans un premier temps montrons que 1 est une valeur propre simple de P, autrement dit montrons que les fonctions P-harmoniques sur $\mathcal{H}_\alpha(E)$ sont constantes. Soit $h \in \mathcal{H}_\alpha(E)$ telle que $Ph = h$ et fixons $x \in E$. D'après le lemme 3.1, il existe un ensemble $\Omega_x \subset \Omega$ de $\mathbb{P}_x$-mesure 1 tel que, pour tout $\omega \in \Omega_x$ et toute transformation $\xi \in T_\mu$, les suites $(h(X_n(\omega)))_{n\geq 0}$ et $(h(\xi \cdot X_n(\omega)))_{n\geq 0}$ convergent vers $H^\infty(\omega)$.

Considérons alors une suite contractante $(\xi_k)_{k\geq 0}$ de T_μ, notons x_0 son point limite. La fonction h étant continue, pour tout $\omega \in \Omega_x$, toute valeur d'adhérence x_ω de $(X_n(\omega))_{n\geq 0}$ et tout $k \geq 0$, on a

$$H^\infty(\omega) = h(x_\omega) = h(\xi_k(x_\omega)).$$

D'où $H^\infty(\omega) = h(x_0)$, en faisant tendre k vers $+\infty$. L'égalité (3.9) entraîne alors $h(x) = h(x_0)$; ainsi, les constantes sont les seules fonctions propres de P associées à la valeur propre 1.

Expliquons maintenant comment le lemme 3.2 permet de contrôler le spectre périphérique de P. Fixons tout d'abord une sous-suite $(n_l)_{l\geq 0}$ d'entiers telle que $\lim\limits_{l\to+\infty} \lambda^{-n_l} = 1$ et une suite contractante $(\xi_k)_{k\geq 0}$ de T_μ, de point limite x_0. L'ensemble $\{T_q \cdots T_1 \mid q \geq 1, T_1, \ldots, T_q \in S_\mu\}$ étant dense dans T_μ, on peut sans perdre en généralité, supposer que chaque fonction ξ_k se décompose sous forme d'un produit $T_{q_k} \cdots T_1$ avec $T_i \in S_\mu, 1 \leq i \leq q_k$. D'après le lemme 3.2, il existe $\Omega_x \subset \Omega$ de $\mathbb{P}_x$-mesure 1 tel que, pour tout $\omega \in \Omega_x$ et tout $k \geq 0$, les suites $\Big(\lambda^{-n_l}\phi(X_{n_l}(\omega))\Big)_{l\geq 0}$ et $\Big(\lambda^{-(n_l+q_k)}\phi(\xi_k \cdot X_{n_l}(\omega))\Big)_{l\geq 0}$ convergent vers la même limite $\Phi^\infty(\omega)$. Considérons alors des suites d'entiers $(\varphi(l))_{l\geq 0}$ (dépendante de ω) et $(\psi(k))_{k\geq 0}$ (indépendante de ω) telles que $(\lambda^{-n_{\varphi(l)}} X_{n_{\varphi(l)}}(\omega))_{l\geq 0}$ et

$(\lambda^{-q_{\psi(k)}})_{k\geq 0}$ convergent respectivement vers $x_\omega \in E$ et $e^{i\beta}, \beta \in \mathbb{R}$; la suite $(\xi_k)_{k\geq 0}$ étant contractante de point limite x_0, l'égalité (3.13) entraîne

$$\Phi^\infty(\omega) = \phi(x_\omega) = e^{i\beta}\phi(x_0).$$

D'après l'égalité (3.12), on a $\phi(x) = \mathbb{E}_x(\Phi^\infty) = e^{i\beta}\phi(x_0)$, ce qui prouve que ϕ est constante sur E et $\lambda = 1$. $\square$

3.3 La chaîne de Diaconis- Freedman sur $[0, 1]$

On étudie dans cette partie la chaîne $(Z_n)_{n\geq 0}$ dite de Diaconis-Freedman, décrite dans l'introduction de ce chapitre, dans le cas où les poids p et q peuvent varier en fonction de x. L'opérateur de transition Q de la chaîne $(Z_n)_{n\geq 0}$ est défini pour toute fonction borélienne bornée $\varphi : [0, 1] \to \mathbb{C}$, par :

$$Q\varphi(x) \;=\; p(x)\int_0^1 \varphi(tx)\mathrm{d}t + q(x)\int_0^1 \varphi(tx+1-t)\mathrm{d}t. \tag{3.16}$$

Pour tout $x \in [0, 1]$, on note μ_x la mesure de probabilité sur $\mathbb{L}\mathrm{ip}([0, 1], [0, 1])$ définie par

$$\mu_x(\mathrm{d}T) = p(x)\int_0^1 \delta_{H_t}(\mathrm{d}T)\mathrm{d}t + q(x)\int_0^1 \delta_{A_t}(\mathrm{d}T)\mathrm{d}t, \tag{3.17}$$

où pour $t \in [0.1]$, H_t est l'homothétie $x \mapsto tx$ et A_t est la transformation affine $x \mapsto tx+1-t$. L'opérateur de transition Q de la chaîne $(Z_n)_{n\geq 0}$ est alors donné par

$$Q\varphi(x) = \int_{\mathbb{L}\mathrm{ip}([0,1],[0,1])} \varphi(T(x))\mu_x(\mathrm{d}T).$$

La chaîne $(Z_n)_{n\geq 0}$ rentre donc bien dans le formalisme développé dans la section précédente.

3.3.1 La chaîne de Diaconis-Freedman sur $[0, 1]$: cas particulier des fonctions poids Hölder continues et strictement positives

Soit $\mathcal{H}_\alpha[0, 1]$ l'espace des fonctions α-Hölder continues de $[0, 1]$ dans $\mathbb{C}$. Supposons dans ce cas que la fonction p appartient à $\mathcal{H}_\alpha[0, 1]$ et vérifie

$$\forall x \in [0, 1] \qquad p(x) > 0 \tag{3.18}$$

(ou, de façon symétrique $\forall x \in [0,1] \quad p(x) < 1$).

Sous l'hypothèse (3.18), la chaîne $(Z_n)_{n\geq 0}$ admet une unique mesure de probabilité invariante sur $[0,1]$ car les hypothèses **H1**, **H2** et **H3** de la proposition 3.4 sont satisfaites (avec la famille de mesures $(\mu_x)_{x\in[0,1]}$ donnée par la formule (3.17)) ; en effet

1. Hypothèse **H1** : pour tous $x, y \in [0,1], x \neq y$, on a

$$\begin{aligned}\sup_{\substack{x,y\in[0,1]\\ x\neq y}} \int_{\mathbb{L}\mathrm{ip}([0,1],[0,1])} \Big(\frac{d(T(x),T(y))}{d(x,y)}\Big)^{\alpha} \mu(\mathrm{d}T) &\leq p(x)\int_0^1 [H_t]^{\alpha}\mathrm{d}t + q(x)\int_0^1 [A_t]^{\alpha}\mathrm{d}t \\ &= \int_0^1 t^{\alpha}\mathrm{d}t = \frac{1}{1+\alpha}.\end{aligned}$$

2. Hypothèse **H2** : pour tous $x, y \in [0,1], x \neq y$, on a

$$\frac{|\mu_x - \mu_y|}{|x-y|^{\alpha}} \leq \frac{|p(x)-p(y)|}{|x-y|^{\alpha}}\Big|\int_0^1 \delta_{H_t} dt\Big| + \frac{|q(x)-q(y)|}{|x-y|^{\alpha}}\Big|\int_0^1 \delta_{A_t} dt\Big| \leq 2m_{\alpha}(p).$$

3. *Hypothèse **H3*** : pour tout $x \in [0,1]$, on a

$$\mu_x \geq \delta \int_0^1 \delta_{H_t} dt$$

   avec $\delta := \inf_{x\in[0,1]} p(x)$, le nombre réel δ est positif car p est une fonction continue sur $[0,1]$ et $p(x) > 0$ pour tout $x \in [0,1]$.

   La fonction $H_0 : x \mapsto 0$ appartient au support de $\mu := \int_0^1 \delta_{H_t} dt$ et fournit de façon évidente une suite contractante au semi-groupe T_{μ}.

Conclusion

• Lorsque $p(0) = 1$, l'unique mesure de probabilité invariante pour $(Z_n)_{n\geq 0}$ est la masse de Dirac en 0.

• Lorsque $p(0) < 1$, nous verrons dans la suite que l'unique mesure de probabilité invariante possède une densité par rapport à la mesure de Lebesgue.

3.3.2 La Chaîne de Diaconis-Freedman sur $[0,1]$: cas général des fonctions poids Hölder continues

Dans ce paragraphe nous nous intéressons à la chaîne $(Z_n)_{n\geq 0}$ dans le cas où les fonctions poids p et q sont dans $\mathcal{H}_{\alpha}[0,1]$.

Donnons tout d'abord quelques exemples explicites.

1. Lorsque $p(x) = x$, la chaîne $(Z_n)_{n\geq 0}$ est une suite de variables aléatoires indépendantes et de loi uniforme sur $[0, 1]$.

2. Lorsque $p(x) = 1 - x$, on vérifie que 0 et 1 sont des points absorbants pour $(Z_n)_{n\geq 0}$; les masses de Dirac δ_0 et δ_1 sont donc Q-invariantes. Nous verrons dans la suite que ce sont les seules mesures ergodiques sur $[0, 1]$ et montrerons que la suite $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-presque sûrement vers une variable aléatoire Z_∞ à valeurs dans $\{0, 1\}$.

L'observation sur des exemples variés montre que la condition de stricte positivité de p ou de q, nécessaire pour pouvoir appliquer la proposition 3.4, est clairement trop forte pour assurer l'unicité d'une mesure invariante attractive pour $(Z_n)_{n\geq 0}$. Reprenant l'approche développée par Diaconis et Freedman, on peut expliciter une mesure invariante avec une densité par rapport à la mesure de Lebesgue, si $p(0) < 1$ et $q(1) < 1$. Nous avons alors l'énoncé suivant. Celui-ci n'est pas une conséquence directe de la proposition 3.4 mais sa démonstration s'en inspire fortement.

Théorème 3.1. *Soit $(Z_n)_{n\geq 0}$ la chaîne de Diaconis-Freedman avec les fonctions poids p et q sont dans $\mathcal{H}_\alpha[0, 1]$. Alors*

1. *Si $p(0) < 1$ et $q(1) < 1$, il existe sur $[0, 1]$ une unique mesure de probabilité Q-invariante ν_p. Celle-ci admet la densité f_p par rapport à la mesure de Lebesgue, donnée par*
$$\forall x \in [0, 1] \quad f_p(x) = C \exp\Big(\int_x^{\frac{1}{2}} \frac{p(y)}{y} dy + \int_{\frac{1}{2}}^x \frac{q(y)}{1-y} dy\Big)$$
où C est un constante de normalisation. De plus, il existe $\kappa > 0$ et $0 < \rho < 1$ tels que
$$\forall \varphi \in \mathcal{H}_\alpha[0, 1],\ \forall x \in [0, 1] \quad |Q^n\varphi(x) - \nu_p(\varphi)| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

2. *Si $p(0) = 1$ et $q(1) < 1$, alors la masse de Dirac δ_0 est l'unique mesure de probabilité Q-invariante sur $[0, 1]$ et il existe $\kappa > 0$ et $0 < \rho < 1$ tels que*
$$\forall \varphi \in \mathcal{H}_\alpha[0, 1],\ \forall x \in [0, 1] \quad |Q^n\varphi(x) - \varphi(0)| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

(énoncé analogue lorsque $p(0) < 1$ et $q(1) = 1$).

3. *Si $p(0) = 1$ et $q(1) = 1$ alors les seules mesures de probabilités Q-invariante sur $[0,1]$ sont les combinaisons convexes de δ_0 et δ_1. De plus la chaîne $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-presque sûrement vers une variable aléatoire Z_∞ à valeurs dans $\{0,1\}$ et la loi de Z_∞ est donnée par*

$$\mathbb{P}_x(Z_\infty = 0) = 1 - h(x) \quad \textit{et} \quad \mathbb{P}_x(Z_\infty = 1) = h(x),$$

où h est l'unique fonction vérifiant l'équation $Qh = h$ et telle que $h(0) = 0, h(1) = 1$ (notons que pour tout $x \in [0,1]$ on a $0 \leq h(x) \leq 1$). Enfin, il existe $\kappa > 0$ et $0 < \rho < 1$ tels que

$$\forall \varphi \in \mathcal{H}_\alpha[0,1],\ \forall x \in [0,1] \quad |Q^n\varphi(x) - (1-h(x))\varphi(0) - h(x)\varphi(1)| \leq \kappa\rho^n \|\varphi\|_\alpha.$$

Démonstration Deux remarques essentielles sont à préciser :

1. L'opérateur Q agit clairement sur $\mathcal{H}_\alpha[0,1]$ et vérifie l'inégalité (3.15), il est donc quasi-compact sur $\mathcal{H}_\alpha[0,1]$ et il suffit à présent de contrôler son spectre périphérique.

2. L'opérateur Q est markovien, il agit donc sur $\mathbb{L}^\infty[0,\ 1]$ et l'on peut alors considérer l'opérateur Q^* sur $\mathbb{L}^1[0,1]$ défini, pour toutes fonctions $\varphi \in \mathbb{L}^1[0,1]$ et $\psi \in \mathbb{L}^\infty[0,1]$, par la relation de dualité suivante.

$$\int_0^1 \varphi(x) Q\psi(x)\mathrm{d}x = \int_0^1 Q^*\varphi(x)\psi(x)\mathrm{d}x.$$

Un calcul direct montre que

$$\forall \varphi \in \mathbb{L}^1[0,1],\ \forall x \in [0,1] \qquad Q^*\varphi(x) = \int_0^x \frac{q(t)}{1-t}\varphi(t)dt + \int_x^1 \frac{p(t)}{t}\varphi(t)dt; \tag{3.19}$$

L'opérateur Q^* n'est autre que l'adjoint de Q relativement à la mesure de Lebesgue sur $[0,1]$. Remarquons que si $f \in \mathbb{L}^1[0,1]$ est une fonction positive et vérifie l'équation de point fixe $Q^*f = f$, alors la mesure de densité f relativement à la mesure de Lebesgue sur $[0,1]$ est Q-invariante.

A fin de déterminer une telle solution on suppose que f est dérivable sur $]0,1[$.
Comme $Q1 = 1$, alors $\displaystyle\int_0^1 Q^*f(x)\mathrm{d}x = \int_0^1 f(x)\mathrm{d}x$.
L'équation du point fixe entraîne $(Q^*f)' = f'$. Si f est une solution de l' équation $(Q^*f)' = f'$, alors il existe $C \in \mathbb{R}$ tel que $Q^*f = f + C$ et l'égalité
$\displaystyle\int_0^1 Q^*f(x)\mathrm{d}x = \int_0^1 f(x)\mathrm{d}x$, implique $C = 0$ et ainsi $Q^*f = f$.
Cherchons donc une solution de l'équation $(Q^*f)' = f'$. A partir de (3.19) on a nécessairement

$$\forall x \in]0,1[\quad f'(x) = \Big(\frac{q(x)}{1-x} - \frac{p(x)}{x}\Big) f(x)$$

d'où $f(x) = C\exp\Big(\displaystyle\int_x^{1/2} \frac{p(t)}{t}dt + \int_{1/2}^x \frac{q(t)}{1-t}dt\Big)$ où C est une constante positive donnée. Cette fonction est intégrable par rapport à la mesure de Lebesgue sur $[0,1]$ si et seulement si $p(0) < 1$ et $q(1) < 1$, dans ce cas on pose

$$f : x \mapsto \frac{1}{C_p}\exp\Big(\int_x^{1/2} \frac{p(t)}{t}\mathrm{d}t + \int_{1/2}^x \frac{q(t)}{1-t}\mathrm{d}t\Big) \tag{3.20}$$

avec $C_p := \displaystyle\int_0^1 \exp\Big(\int_x^{1/2} \frac{p(t)}{t}\mathrm{d}t + \int_{1/2}^x \frac{q(t)}{1-t}\mathrm{d}t\Big)\mathrm{d}x$.
Ainsi, la mesure de probabilité ν_p de fonction de densité f_p par rapport à la mesure de Lebesgue sur $[0,1]$ est Q-invariante.

Lorsque $p(0) = 1$, la masse de Dirac δ_0 est Q-invariante ; de façon symétrique, lorsque $q(1) = 1$, c'est la masse δ_1 qui est Q-invariante.

Revenons à la démonstration du théorème, elle se fera en trois étapes.
Étape 1 : Quasi-compacité de l'opérateur Q sur $\mathcal{H}_\alpha[0,1]$

Par hypothèse les poids p et q sont α-Hölder continues sur $[0,1]$ de plus $p + q = 1$ ce qui implique que l'opérateur Q est positif et borné sur $\mathcal{H}_\alpha[0,1]$ et de rayon spectral égal à 1. L'expression (3.16) entraîne de façon immédiate

$$\forall \varphi \in \mathcal{H}_\alpha[0,1] \qquad \|Q\varphi\|_\alpha \leq \frac{1}{\alpha+1}\|\varphi\|_\alpha + (1 + 2m_\alpha(p))|\varphi|_\infty$$

d'où l'on déduit que Q est quasi-compact sur $\mathcal{H}_\alpha[0,1]$.

Étape 2 : Contrôle de la valeur propre $\lambda = 1$ dans $\mathcal{H}_\alpha[0,1]$

La description du sous-espace caractéristique de Q associé à la valeur propre 1 s'inscrit dans un cadre général détaillée dans [32], par la notion de *compact absorbant.*

Un sous-ensemble compact K de $[0,1]$ est dit *Q-absorbant* lorsque pour tout $x \in K$ on a $Q1_{[0,1]\setminus K}(x) = 0$, il est *minimal* lorsqu'il ne contient pas de sous-ensemble compact propre absorbant. Remarquons que pour tout $x \in [0,1]$, la condition $p(x) > 0$ assure que

$$Q(x, I) > 0 \quad \text{pour tout intervalle ferm } I \subset [0, x] \text{ non rduit un point.} \tag{3.21}$$

De même la condition $q(x) > 0$ entraîne que

$$Q(x, I) > 0 \quad \text{pour tout intervalle ferm } I \subset [x, 1] \text{ non rduit un point.} \tag{3.22}$$

Par ailleurs, lorsque $p(x) = 0$ (resp. $q(x) = 0$), on a $Q1_{[x,1]} = 1$ (resp. $Q1_{[0,x]} = 1$). Quatre cas se présentent donc :

1. $q(0) > 0$ *et* $p(1) > 0$

   Dans ce cas, l'intervalle $[0,1]$ est l'unique compact Q-absorbant, il est donc minimal; En effet, tout d'abord $[0,1]$ est bien un compact absorbant (toujours); dans le cas présent il faut montrer que c'est le seul, autrement dit que tout autre compact $K \subsetneq [0,1]$ n'est pas absorbant, ce qui signifie qu'il existe $x_0 \in K$ tel que $Q1_{[0,1]\setminus K}(x_0) \neq 0$.

   Supposons $0 \notin K$, comme K est compact,il existe $\epsilon > 0$ tel que $K \subset [\epsilon, 1]$.

   (a) s'il existe $x_0 \in K$ tel que $p(x_0) > 0$, on aura

   $$Q(x_0, [0,1] \setminus K) \geq Q(x_0, [0, \epsilon[) > 0,$$

   ce qui est la conclusion espérée.

   (b) Sinon $p(x) = 0$, soit $q(x) = 1$, pour tout $x \in K$, comme $p(1) > 0$ on a $q(1) < 1$ si bien que $1 \notin K$ et comme précédemment il existe $\epsilon' > 0$ tel que $K \subset [0, 1-\epsilon']$; mais dans ce cas, quel que soit le choix de $x_0 \in K$, puisque $q(x_0) = 1$, on aura

   $$Q(x_0, [0,1] \setminus K) \geq Q(x_0,]1-\epsilon', 1]) > 0.$$

Si $0 \in K$ et $1 \notin K$, alors comme K est compact, il existe $\epsilon_1 > 0$ tel que $K \subset [0, 1-\epsilon_1]$.

(a) s'il existe $x_0 \in K$ tel que $q(x_0) > 0$, on aura

$$Q(x_0, [0,1] \setminus K) \geq Q(x_0,]1-\epsilon_1, 1]) > 0,$$

ce qui est la conclusion espérée.

(b) sinon $q(x) = 0$, soit $p(x) = 1$, pour tout $x \in K$; comme $p(0) > 0$ on a $q(0) < 1$ si bien que $0 \notin K$ et comme précédemment il existe $\epsilon_1' > 0$ tel que $K \subset [\epsilon_1', 1]$; mais dans ce cas, quel que soit le choix de $x_0 \in K$, puisque $p(x_0) = 1$, on aura

$$Q(x_0, [0,1] \setminus K) \geq Q(x_0, [0, \epsilon_1']) \geq \frac{\epsilon_1'}{x_0} > 0.$$

Lorsque 0 et 1 appartiennent à K ; comme $K \subsetneq [0,1]$, il existe $x_1 \in]0,1[$ tel que $x_1 \notin K$; par compacité de K, il existe $\epsilon_2 > 0$ tel que $]x_1 - \epsilon_2, x_1 + \epsilon_2[\subset [0,1] \setminus K$. On peut prendre $x_0 = 0$ (ou bien 1) dans ce cas : en effet,

$$Q(0, [0,1] \setminus K) \geq Q(0,]x_1 - \epsilon_2, x_1 + \epsilon_2[) > 0.$$

2. $\underline{q(0) = 0 \text{ } et \text{ } p(1) > 0}$

Dans ce cas l'unique compact Q-absorbant minimal est $\{0\}$. Autrement dit $\{0\}$ est un compact invariant (ce qui découle de la condition $q(0) = 0$) et tout autre compact invariant(absorbant) $K \subset [0,1]$ contient 0.

En effet, d'après la propriété (3.21) ci-dessus, si $x \in K$ vérifie $p(x) > 0$, alors K contient l'intervalle $[0, x]$ et donc en particulier 0. Il nous suffit donc de vérifier qu'un tel point x existe dans K. Dans le cas contraire, la fonction q serait identiquement égale à 1 sur K et l'on aurait alors $[x, 1] \subset K$ pour tout $x \in K$, toujours d'après (3.22), le point 1 appartiendrait donc à K et vérifierait en particulier l'égalité $q(1) = 1$. Contradiction avec la condition $p(1) > 0$

3. $\underline{q(0) > 0 \text{ } et \text{ } p(1) = 0}$

Dans ce cas, l'unique compact Q-absorbant minimal est $\{1\}$. La démonstration se fait de façon analogue que dans le cas précédent, en échangeant le rôle de 0 et de 1.

4. $\underline{q(0) = 0 \ et \ p(1) = 0}$

   Dans ce cas, les ensembles $\{0\}$ et $\{1\}$ sont les seuls compacts absorbants minimaux.

De l'égalité $Q1 = 1$ et du théorème 2.7 on déduit que, dans $\mathcal{H}_\alpha[0,1]$, l'indice de la valeur propre 1 est égal à 1 ; en d'autres termes, le sous-espace caractéristique de Q associé à la valeur propre 1 n'est autre que le sous-espace propre correspondant $\mathrm{Ker}(Q - Id)$. On peut donc appliquer le théorème 2.8 du chapitre 2 pour conclure. Nous avons alors :

1. Lorsque $q(0) > 0$ et $p(1) > 0$ le sous-espace $\mathrm{Ker}(Q - Id)$ est de dimension 1 et se réduit aux fonctions constantes, il existe par ailleurs une unique mesure de probabilité Q-invariante sur $[0,1]$, absolument continue par rapport à la mesure de Lebesgue, de densité f_p donnée par la formule (3.20).

2. Lorsque $q(0) = 0$ et $p(1) > 0$, le sous-espace $\mathrm{Ker}(Q - Id)$ est de dimension 1 et se réduit aux fonctions constantes et la masse de Dirac δ_0 est l'unique mesure de probabilité Q-invariante sur $[0,1]$.

3. Lorsque $q(0) > 0$ et $p(1) = 0$, le sous-espace $\mathrm{Ker}(Q - Id)$ est de dimension 1 et se réduit aux fonctions constantes. La masse de Dirac δ_1 est l'unique mesure de probabilité Q-invariante sur $[0,1]$.

4. Enfin, lorsque $q(0) = 0$ et $p(1) = 0$, le sous-espace $\mathrm{Ker}(Q - Id)$ est de dimension 2. Il est engendré par la fonction constante $\mathbf{1}$ et une fonction h telle que $h(0) = 0$ et $h(1) = 1$ et les seules mesures de probabilités Q-invariantes sur $[0,1]$ sont les combinaisons convexes $\alpha\delta_0 + (1-\alpha)\delta_1, \ 0 \leq \alpha \leq 1$.

Étape 3 : Contrôle du spectre périphérique de Q dans $\mathcal{H}_\alpha[0,1]$

Nous adaptons ici au cas du noyau Q la stratégie développée dans la section précédente et basée sur le lemme 3.2.

Soit $\phi \in \mathcal{H}_\alpha[0,1]$ telle que $Q\phi = \lambda\phi$ où λ est un nombre complexe de module 1. Pour tout $x \in [0,1]$, la suite $(\lambda^{-n}\phi(X_n))_{n\geq 0}$ est une martingale bornée sous $\mathbb{P}_x$, elle converge donc $\mathbb{P}_x$-presque sûrement vers une v.a. notée Φ^∞. D'après l'inégalité (3.11) et la formule

(3.17), il existe $\Omega_x \subset \Omega$, $\mathbb{P}_x(\Omega_x) = 1$, et $I_0 \subset [0,1]$ de mesure de Lebesgue 1 tels que, pour tout $\omega \in \Omega_x$ et tous $s, t \in I_0$, on ait d'une part (en prenant $q = 1$)

$$\lim_{n\to+\infty} \left|\phi(Z_n(\omega)) - \lambda^{-1}\phi(H_s \cdot Z_n(\omega))\right|^2 p(Z_n(\omega)) = 0 \tag{3.23}$$

et

$$\lim_{n\to+\infty} \left|\phi(Z_n(\omega)) - \lambda^{-1}\phi(A_s \cdot Z_n(\omega))\right|^2 q(Z_n(\omega)) = 0 \tag{3.24}$$

et d'autre part (en prenant $q = 2$)

$$\lim_{n\to+\infty} \left|\phi(Z_n(\omega)) - \lambda^{-2}\phi(H_t H_s \cdot Z_n(\omega))\right|^2 p(Z_n(\omega)) p(H_s \cdot Z_n(\omega)) = 0 \tag{3.25}$$

et

$$\lim_{n\to+\infty} \left|\phi(Z_n(\omega)) - \lambda^{-2}\phi(A_t A_s \cdot Z_n(\omega))\right|^2 q(Z_n(\omega)) q(A_s \cdot Z_n(\omega)) = 0 \tag{3.26}$$

Deux cas sont possibles.

1. *Cas où* $\phi(0) = \phi(1)$.

   Fixons $\omega \in \Omega_x$ et une valeur d'adhérence z_ω de la suite $(Z_n(\omega))_{n\geq 0}$.

   Si $p(z_\omega) \neq 0$, en prenant s arbitrairement proche de 0, l'égalité (3.23) entraîne

   $$\phi(z_\omega) = \lambda^{-1}\phi(0).$$

   Lorsque $p(z_\omega) = 0$, on utilise (3.24) pour conclure de la même façon que

   $$\phi(z_\omega) = \lambda^{-1}\phi(1).$$

   Comme $\phi(0) = \phi(1)$, la suite $(\phi(Z_n(\omega))_{n\geq 0}$ n'a donc qu'une valeur d'adhérence, elle converge vers $\Phi^\infty(\omega) = \lambda^{-1}\phi(0)$ (qui ne dépend pas de ω) et l'on a

   $$\phi(x) = \mathbb{E}_x(\Phi^\infty) = \lambda^{-1}\phi(0).$$

   La fonction ϕ est donc constante et $\lambda = 1$.

2. *Cas où* $\phi(0) \neq \phi(1)$.

   Une des deux valeurs $\phi(0)$ et $\phi(1)$ est non nulle, on suppose $\phi(0) \neq 0$, le cas $\phi(1) \neq 0$ se traite de façon analogue.

(a) Supposons dans un premier temps qu'il existe $x \in [0,1]$ et $\omega_x \in \Omega_x$ tel que la suite $(Z_n(\omega_x))_{n\geq 0}$ possède une valeur d'adhérence z_{ω_x} vérifiant $p(z_{\omega_x}) > 0$. On applique (3.23) avec s arbitrairement proche de 0 puis (3.25) avec s arbitrairement proche de 1 (de façon que $p(H_s(z_{\omega_x})) > 0$) et t arbitrairement proche de 0, il vient

$$\phi(z_{\omega_x}) = \lambda^{-1}\phi(0) = \lambda^{-2}\phi(0).$$

Puisque $\phi(0) \neq 0$, on a nécessairement $\lambda = 1$ et la fonction ϕ est constante.

(b) Supposons maintenant que pour tout $x \in [0,1]$ et tout $\omega \in \Omega_x$, les valeurs d'adhérence z_ω de la suite $(Z_n(\omega))_{n\geq 0}$ vérifient toutes l'égalité $q(z_\omega) = 1$.

En appliquant cette fois-ci (3.24) avec s arbitrairement proche de 0 puis (3.26) avec s arbitrairement proche de 1 (de façon que $q(A_s(z_\omega)) > 0$) et t arbitrairement proche de 0, on obtient

$$\phi(z_\omega) = \lambda^{-1}\phi(1) = \lambda^{-2}\phi(1).$$

Si $\phi(1) \neq 0$, on déduit comme ci-dessus $\lambda = 1$ et la fonction ϕ est constante ($\phi \in \mathbb{C}.\mathbf{1}$) . Si $\phi(1) = 0$, la suite $(\lambda^{-n}\phi(Z_n(\omega))_{n\geq 0}$ converge vers $\Phi^\infty(\omega)$ et puisque $\phi(z_\omega) = 0$ on a nécessairement $\Phi^\infty(\omega) = 0$. L'égalité de martingale $\phi(x) = \mathbb{E}_x(\lambda^{-n_l}\phi(Z_{n_l}))$ permet de conclure que ϕ est identiquement nulle.

En conclusion, l'opérateur Q est quasi-compact sur $\mathcal{H}_\alpha[0,1]$, de rayon spectral 1 et son spectre périphérique est réduit à $\{1\}$ et le sous-espace caractéristique de Q associé à la valeur propre 1 est égal à $\mathrm{Ker}(Q - Id)$. Plus précisément nous avons ces quatre cas.

1. Lorsque $q(0) > 0$ et $p(1) > 0$, il existe un opérateur borné R sur $\mathcal{H}_\alpha[0,1]$ de rayon spectral $\rho < 1$ tel que pour toute fonction $f \in \mathcal{H}_\alpha[0,1]$ et tout $n \geq 0$, on ait

$$Q^n f = \Big(\int_0^1 f(x)\varphi_p(x)dx\Big)\mathbf{1} + R^n f.$$

Dans ce cas, la chaîne $(Z_n)_{n\geq 0}$ est récurrente sur $[0,1]$.

2. Lorsque $q(0) = 0$ et $p(1) > 0$, il existe un opérateur borné R sur $\mathcal{H}_\alpha[0,1]$ de rayon spectral $\rho < 1$ tel que pour toute fonction $f \in \mathcal{H}_\alpha[0,1]$ et tout $n \geq 0$, on ait

$$Q^n f = f(0)\mathbf{1} + R^n f. \tag{3.27}$$

Dans ce cas, pour tout état initial $x \in [0,1]$, la chaîne $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-p.s. vers 0 ; de plus, pour tout $\epsilon \in]0,1[$, l'ensemble $[\epsilon, 1]$ est transient et il existe $\kappa_\epsilon > 0$ tel que pour tout $x \in [\epsilon, 1]$ on ait

$$\mathbb{P}_x(Z_n \in [\epsilon, 1]) \leq \kappa_\epsilon \rho^n.$$

On peut en effet majorer l'indicatrice de l'intervalle $[\epsilon, 1]$ par une fonction $f_\epsilon \in \mathcal{H}_\alpha[0,1]$ qui s'annule en 0 et on applique l'inégalité (3.27).

3. Lorsque $q(0) = 0$ et $p(1) > 0$, il existe un opérateur borné R sur $\mathcal{H}_\alpha[0,1]$ de rayon spectral $\rho < 1$ tel que pour toute fonction $f \in \mathcal{H}_\alpha[0,1]$ et tout $n \geq 0$, on ait

$$Q^n f = f(1)\mathbf{1} + R^n f.$$

Pour tout $x \in [0,1]$, la chaîne $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-p.s. vers 1 ; de plus, pour tout $\epsilon \in]0,1[$, l'ensemble $[0, 1-\epsilon]$ est transient et il existe $\kappa_\epsilon > 0$ tel que pour tout $x \in [0, 1-\epsilon]$ on ait

$$\mathbb{P}_x(Z_n \in [0, 1-\epsilon]) \leq \kappa_\epsilon \rho^n.$$

4. Enfin, lorsque $q(0) = 0$ et $p(1) = 0$, il existe une fonction harmonique h telle que $h(0) = 0$ et $h(1) = 1$ et un opérateur borné R sur $\mathcal{H}_\alpha[0,1]$ de rayon spectral $\rho < 1$ tel que pour toute fonction $f \in$ et tout $n \geq 0$, on ait

$$Q^n f = f(0)(\mathbf{1} - h) + f(1)h + R^n f.$$

Dans ce cas, la fonction h est à valeurs dans $[0,1]$ et pour tout état initial $x \in [0,1]$, la chaîne $(Z_n)_{n\geq 0}$ converge vers 0 avec la probabilité $1-h(x)$ et vers 1 avec la probabilité $h(x)$. En effet, comme dans les cas (2) et (3), pour toute $\epsilon \in]0, \frac{1}{2}[$, l'ensemble $[\epsilon, 1-\epsilon]$ est transient et il existe $\kappa_\epsilon > 0$ tel que pour tout $x \in [0, 1-\epsilon]$ on ait

$$\mathbb{P}_x\left(Z_n \in [\epsilon, 1-\epsilon]\right) \leq \kappa_\epsilon \rho^n.$$

Si bien que pour tout point initial $x \in]0,1[$, les seules valeurs d'adhérence de la suite $(Z_n)_{n\geq 0}$ sont 0 et 1. Par ailleurs, $(h(Z_n))_{n\geq 0}$ est une martingale bornée, elle converge donc et du fait que $h(0) \neq h(1)$ on déduit que la chaîne $(Z_n)_{n\geq 0}$ est en fait convergente, vers une v.a. Z_∞ à valeurs dans $\{0,1\}$. L'égalité de martingale $h(x) = \mathbb{E}_x(h(Z_\infty)) = \mathbb{P}_x(Z_\infty = 1)$ permet de conclure.

Exemple : $p(x) = 1 - x$. Dans ce cas, la fonction $h : x \to x$ est Q-harmonique et l'on peut appliquer le théorème précédent. Notons que pour ce choix de la fonction p, on peut montrer directement que les ensembles $[\epsilon, 1-\epsilon], 0 < \epsilon < \frac{1}{2}$, sont transients en introduisant la quantité Δ définie par : pour tout $0 \leq x \leq 1$

$$\Delta(x) := \mathrm{dist}(x, \{0,1\}) = \inf(x, 1-x).$$

Calculons $\mathbb{E}_x(\Delta(Z_1))$. Nous supposons $0 < x \leq \frac{1}{2}$, le cas $\frac{1}{2} \leq x \leq 1$ se traitant de façon analogue ; on a

$$\begin{aligned} \mathbb{E}_x(\Delta(Z_1)) &= \frac{1-x}{x}\int_0^x y\,dy + \frac{x}{1-x}\int_x^{\frac{1}{2}} y\,dy + \frac{x}{1-x}\int_{\frac{1}{2}}^1 (1-y)dy \\ &= \frac{3x - 4x^2}{4(1-x)} \\ &\leq \frac{3}{4}x, \end{aligned}$$

d'où $\mathbb{E}(\Delta(Z_n)|\mathcal{F}_{n-1}) \leq \dfrac{3}{4}\Delta(Z_{n-1})$ pour tout $n \geq 1$ si bien que par itération

$$\forall n \geq 1, \forall x \in [0,1] \qquad \mathbb{E}_x(\Delta(Z_n)) \leq \left(\frac{3}{4}\right)^n \Delta(x) \leq \left(\frac{3}{4}\right)^n.$$

Il vient $\mathbb{E}_x\Big(\sum_{n\geq 0}\Delta(Z_n)\Big) = \sum_{n\geq 0}\mathbb{E}_x\Big(\Delta(Z_n)\Big) < +\infty$, la suite $(\Delta(Z_n))_{n\geq 1}$ converge donc $\mathbb{P}_x$-p.s. vers 0 et les ensembles $[\epsilon, 1-\epsilon],\ 0 < \epsilon < \frac{1}{2}$ sont bien transients.

La variable aléatoire Z_∞ est à valeurs dans $\{0,1\}$. La martingale $(Z_n)_{n\geq 1}$ étant bornée, elle est fermée par sa limite et pour tout $x \in [0,1]$, on a

$$\mathbb{E}_x(Z_\infty) = \mathbb{P}_x(Z_\infty = 1) = \mathbb{E}_x(Z_0) = x.$$

Ainsi, sous $\mathbb{P}_x$, la suite $(Z_n)_{n\geq 1}$ converge vers 1 avec la probabilité x et vers 0 avec la probabilité $1 - x$.

Chapitre 4

Extensions de la chaîne de Diaconis-Freedman.

Ce chapitre est une introduction à quelques extensions de la chaîne de Diaconis-Freedman $(Z_n)_n$ étudiée dans le chapitre précédent. Nous explicitons quelques idées dans cet esprit.

4.1 Chaîne avec deux points attractifs

Dans cette section nous citons quelques extensions faites et repérées dans la littérature actuelle. Ces extensions peuvent être envisagées de différentes manières. Considérons la chaîne $(Z_n)_n$ sur $[0,1]$, qui décrit le mouvement d'une particule entre deux points attractifs $A_1 = 0$ et $A_2 = 1$, définie par :

$$Z_{n+1} = Z_n(1 - V_{n+1}\mathbb{1}_{\{U_{n+1}\leq p(Z_n)\}}) + (1 - Z_n)W_{n+1}(1 - \mathbb{1}_{\{U_{n+1}\leq p(Z_n)\}}),$$

où $Z_0 = x \in [0,1]$, $(V_n)_n$, $(W_n)_n$ et $(U_n)_n$ sont des suites de variables aléatoires i.i.d. tel que $V_n \rightsquigarrow F_l$, $W_n \rightsquigarrow F_r$ et U_n est de loi uniforme sur $[0,1]$ (i.e. $U_n \rightsquigarrow U[0,1]$).
Lorsque $F_l = F_r$ est une loi de densité h par rapport à la mesure de Lebesgue sur $[0,1]$, la suite $(Z_n)_n$ est une chaîne de Markov d'opérateur de transition Q défini pour toute fonction borélienne bornée $\varphi : [0,1] \to \mathbb{C}$ et tout $x \in [0,1]$ par :

$$Q\varphi(x) = p(x)\int_0^x \frac{\varphi(t)}{x}h(\frac{x-t}{x})\mathrm{d}t + q(x)\int_x^1 \frac{\varphi(t)}{1-x}h(\frac{x-t}{1-x})\mathrm{d}t.$$

Dans le cas où $F_l = F_r = U[0,1]$ et p est une fonction constante sur $[0,1]$, $(Z_n)_n$ est la chaîne étudiée dans [20] et [55].

Dans le cas où $F_l = F_r = U[0,1]$ et p est une fonction dépendante de x et est continue sur $[0,1]$, $(Z_n)_n$ est la chaîne de Diaconis-Freedman étudiée en détail dans le chapitre précédent.

Lorsque par exemple $F_l = F_r = \mathcal{B}(a, a+1), a > 0$ et p est une constante strictement positive, la chaîne $(Z_n)_n$ admet la loi uniforme sur $[0,1]$ comme unique loi stationnaire (voir [49]) pour les détails de ce cas particulier et d'autres cas. Le lecteur peut se référer aussi à [48] pour le cas où p peut varier en fonction de x.

McKinlay et Borovkov [11] ont considéré le cas où $F_l = \mathcal{B}(1,a)$, $F_r = \mathcal{B}(1,b)$, avec $a,\ b > 0$ et p est continue par morceau. Ils montrent en particulier que, pour $F_l = F_r = \mathcal{B}(1,\alpha), \alpha > 0$, et $p(x) = cx + (1-d)(1-x)$, $d, c \in (0,1]$, la chaîne de Markov $(Z_n)_n$ admet la loi $\mathcal{B}(d\alpha, c\alpha)$ comme loi stationnaire.

Tuan-Minh Nguyen dans [63] donne d'autres résultats concernant cette chaîne.

4.2 Chaîne avec trois points attractifs alignés

Nous étudions ici le comportement asymptotique de la chaîne de Markov $(Z_n)_{n\geq 0}$ sur $[0,2]$ décrivant les mouvements d'une particule tel que :

Si la particule se trouve en x à un instant n, son déplacement à venir se déroule en deux temps :

• Dans un premier temps, elle choisit un des 3 points $0, 1$ ou 2 vers lequel elle veut se diriger ; on note ce point A et p, r, q les probabilités respectives des événements $(A = 0)$, $(A = 1)$ et $(A = 2)$, avec les conditions $p, r, q > 0$ et $p + r + q = 1$.

• Une fois le point A déterminé, la particule se déplace en y choisi de façon uniforme dans l'intervalle dont les extrémités sont x et A.

Le noyau de transition Q de la chaîne $(Z_n)_{n\geq 0}$ est donc donné par la formule suivante : pour toute fonction borélienne bornée $\varphi : [0,2] \to \mathbb{R}$, lorsque $0 < x < 1$,

$$Q\varphi(x) = \frac{p}{x}\int_0^x \varphi(y)\mathrm{d}y + \frac{r}{1-x}\int_x^1 \varphi(y)\mathrm{d}y + \frac{q}{2-x}\int_x^2 \varphi(y)\mathrm{d}y,$$

et, lorsque $1 < x < 2$,

$$Q\varphi(x) = \frac{p}{x}\int_0^x \varphi(t)\mathrm{d}y + \frac{r}{x-1}\int_1^x \varphi(y)\mathrm{d}y + \frac{q}{2-x}\int_x^2 \varphi(y)\mathrm{d}y.$$

Donc pour tout $x \in [0,2],\ x \neq 0,1,2$, on a

$$Q\varphi(x) = \frac{p}{x}\int_0^x \varphi(y)\mathrm{d}y + \frac{r}{1-x}\int_x^1 \varphi(y)\mathrm{d}y + \frac{q}{2-x}\int_x^2 \varphi(y)\mathrm{d}y. \tag{4.1}$$

Cette expression "spatiale" du noyau de transition est d'une grande utilité pour déterminer, lorsque c'est possible, une mesure de probabilité invariante possédant une densité par rapport à la mesure de Lebesgue. La recherche d'une telle densité se fait par l'opérateur adjoint de Q par rapport à la mesure de Lebesgue. La forme explicite de l'adjoint découlera directement de (4.1). Nous remarquons que la chaîne $(Z_n)_{n\geq 0}$ s'inscrit dans le cadre des itérations de transformations aléatoires contractantes car son noyau Q peut s'exprimer sous la forme.

$$Q\varphi(x) = p\int_0^1 \varphi(tx)\mathrm{d}t + r\int_0^1 \varphi(x+t(1-x))\mathrm{d}t + q\int_0^1 \varphi(x+t(2-x))\mathrm{d}t. \tag{4.2}$$

L'opérateur Q est markovien sur $[0,2]$, il agit donc sur $\mathbb{L}^\infty[0,2]$. On peut alors considérer l'opérateur Q^* sur $\mathbb{L}^1([0,2],\mathrm{d}x)$, défini pour toute fonction $\varphi \in \mathbb{L}^1[0,2],\ \psi \in \mathbb{L}^\infty[0,2]$, par l'égalité suivante.

$$\int_0^2 \varphi(x)Q\psi(x)\mathrm{d}x = \int_0^2 Q^*\varphi(x)\psi(x)\mathrm{d}x. \tag{4.3}$$

Un calcul direct montre que l'on peut définir $Q^*\varphi$ comme suit : pour tout $x \in [0,2]$,

$$\begin{aligned} Q^*\varphi(x) &= p\int_x^2 \frac{\varphi(y)}{y}\mathrm{d}y + r\mathbb{1}_{[0,1[}(x)\int_0^x \frac{\varphi(y)}{1-y}\mathrm{d}y \\ &\quad + r\mathbb{1}_{[1,2]}(x)\int_x^2 \frac{\varphi(y)}{1-y}\mathrm{d}y + q\int_0^x \frac{\varphi(y)}{(2-y)}\mathrm{d}y. \end{aligned} \tag{4.4}$$

Remarquons que si $f :]0,2[\to \mathbb{R}^+$ est solution de l'équation de point fixe $Q^*f = f$, alors la mesure $f(x)\mathrm{d}x$ sur $[0,2]$ est Q-invariante. Afin de déterminer une telle solution, on peut supposer que f est dérivable sur $]0,1[\cup]1,2[$. L'équation du point fixe entraîne $(Q^*f)' = f'$. Si f est solution de cette dernière équation, alors $Q^*f = f + C$ où C est une constante ; l'égalité (4.3) entraîne $\displaystyle\int_0^2 Q^*f(x)\mathrm{d}x = \int_0^2 f(x)\mathrm{d}x$, ainsi $C = 0$ et $Q^*f = f$.

À partir de l'expression (4.4), en dérivant l'équation $Q^*f = f$, il vient

$$\forall x \in]0,2[,\ x \neq 1 \qquad \varphi'(x) = \Big(\frac{q}{2-x} + \frac{r}{1-x} - \frac{p}{x}\Big)\varphi(x)$$

d'où

$$\varphi(x) = \frac{C}{x^p(\mid 1-x\mid^r)(2-x)^q}$$

où C est une constante positive donnée.

4.3 La chaîne de Diaconis-Freedman sur $[0,1]\times[0,1]$

Il s'agit de la chaîne de Markov $(Z_n)_{n\geq 0}$ sur $[0,1]\times[0,1]$ définie comme suit : si la chaîne est en $x = (x_1, x_2)$ a l'instant n, on choisit à l'instant $n+1$ l'un des pavés $[0,x_1]\times[0,x_2]$, $[0,x_1]\times[x_2,1]$,$[x_1,1]\times[0,x_2]$ ou $[x_1,1]\times[x_2,1]$ avec les probabilités $p_{00}(x_1,x_2)$, $p_{01}(x_1,x_2)$, $p_{10}(x_1,x_2)$ et $p_{11}(x_1,x_2)$ respéctivement, avec $p_{00}(x_1,x_2) + p_{01}(x_1,x_2) + p_{10}(x_1,x_2) + p_{11}(x_1,x_2) = 1$ et l'on se déplace en un point $y = (y_1,y_2)$ distribué selon la loi uniforme dans le pavé ainsi choisi. Le noyau de transition Q de $(Z_n)_{n\geq 0}$ est donc donné pour toute fonction borélienne bornée $\varphi : [0,1]\times[0,1] \to \mathbb{C}$ et tout $(x_1,x_2) \in]0,1[\times]0,1[$, par la formule suivante.

$$\begin{aligned} Q\varphi(x) = Q\varphi(x_1,x_2) &= \frac{p_{00}(x_1,x_2)}{x_1x_2}\iint_{[0,x_1]\times[0,x_2]}\varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \frac{p_{01}(x_1,x_2)}{x_1(1-x_2)}\iint_{[0,x_1]\times[x_2,1]}\varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \frac{p_{10}(x_1,x_2)}{(1-x_1)(x_2)}\iint_{[x_1,1]\times[0,x_2]}\varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \frac{p_{11}(x_1,x_2)}{(1-x_1)(1-x_2)}\iint_{[x_1,1]\times[x_2,1]}\varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2, \end{aligned}$$

$$Q\varphi(0,0) = (p_{00}(0,0) + p_{01}(0,0) + p_{10}(0,0))\varphi(0,0) + p_{11}(0,0)\iint_{[0,1]\times[0,1]}\varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2;$$

$$Q\varphi(1,0) = (p_{00}(1,0) + p_{10}(1,0) + p_{11}(1,0))\varphi(1,0) + p_{01}(1,0) \iint_{[0,1]\times[0,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2;$$

$$Q\varphi(0,1) = (p_{00}(0,1) + p_{01}(0,1) + p_{11}(0,1))\varphi(0,1) + p_{10}(0,1) \iint_{[0,1]\times[0,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2;$$

$$Q\varphi(1,1) = (p_{01}(1,1) + p_{10}(1,1) + p_{11}(1,1))\varphi(1,1) + p_{00}(1,1) \iint_{[0,1]\times[0,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2;$$

$$\begin{aligned} Q\varphi(0,x_2) &= (p_{00}(0,x_2) + p_{01}(0,x_2))\varphi(0,x_2) + \frac{p_{10}(0,x_2)}{x_2} \iint_{[0,1]\times[0,x_2]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \frac{p_{11}(0,x_2)}{(1-x_2)} \iint_{[0,1]\times[x_2,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2, \end{aligned}$$

et

$$\begin{aligned} Q\varphi(x_1,0) &= (p_{00}(x_1,0) + p_{10}(x_1,0))\varphi(x_1,0) + \frac{p_{01}(x_1,0)}{x_1} \iint_{[0,x_1]\times[0,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \frac{p_{11}(x_1,0)}{(1-x_1)} \iint_{[x_1,1]\times[0,1]} \varphi(y_1,y_2)\mathrm{d}y_1\mathrm{d}y_2, \end{aligned}$$

donc pour tout $x = (x_1,x_2) \in [0,1]\times[0.1]$ on a :

$$\begin{aligned} Q\varphi(x) = Q\varphi(x_1,x_2) = p_{00}(x_1,x_2) &\iint_{[0,1]\times[0,1]} \varphi(tx_1,t'x_2)\mathrm{d}t\mathrm{d}t' \\ + p_{01}(x_1,x_2) &\iint_{[0,1]\times[0,1]} \varphi(tx_1,t'x_2+1-t')\mathrm{d}t\mathrm{d}t' \\ + p_{10}(x_1,x_2) &\iint_{[0,1]\times[0,1]} \varphi(tx_1+1-t,t'x_2))\mathrm{d}t\mathrm{d}t' \\ +p_{11}(x_1,x_2) &\iint_{[0,1]\times[0,1]} \varphi(tx_1+1-t,t'x_2+1-t')\mathrm{d}t\mathrm{d}t' \end{aligned} \tag{4.5}$$

Soient $HH_{(t,t')}$, $HA_{(t,t')}$, $AH_{(t,t')}$ et $AA_{(t,t')}$ les transformations $(x_1,x_2) \mapsto (tx_1,t'x_2)$, $(x_1,x_2) \longmapsto (tx_1,t'x_2+1-t')$, $(x_1,x_2) \longmapsto (tx_1+1-t,t'x_2)$ et $(x_1,x_2) \mapsto (tx_1+1-t,t'x_2+1-t')$ respéctivement. Introduisons alors la mesure $\mu_{(x_1,x_2)}$ sur l'espace $\mathbb{L}\mathrm{ip}([0,1]\times[0,1])$ des fonctions lipschitziennes sur $[0,1]\times[0,1]$ définie par

$$\begin{aligned}
\mu_x(\mathrm{d}T) = \mu_{(x_1,x_2)}(\mathrm{d}T) = p_{00}(x) &\iint_{[0,1]\times[0,1]} \delta_{HH_{(t,t')}}(\mathrm{d}T)\mathrm{d}t\mathrm{d}t' \\
+ p_{01}(x) &\iint_{[0,1]\times[0,1]} \delta_{HA_{(t,t')}}(\mathrm{d}T)\mathrm{d}t\mathrm{d}t' \\
+ p_{10}(x) &\iint_{[0,1]\times[0,1]} \delta_{AH_{(t,t')}}(\mathrm{d}T)\mathrm{d}t\mathrm{d}t' \\
+ p_{11}(x) &\iint_{[0,1]\times[0,1]} \delta_{AA_{(t,t')}}(\mathrm{d}T)\mathrm{d}t\mathrm{d}t'
\end{aligned} \tag{4.6}$$

où δ_T désigne la masse de Dirac en T.

L'égalité (4.5) s'écrit alors

$$Q\varphi(x) = Q\varphi(x_1,x_2) = \int_{\mathbb{L}\mathrm{ip}([0,1]\times[0,1])} \varphi(T(x_1,x_2))\mu(\mathrm{d}T).$$

Nous remarquons que la chaîne de Markov $(Z_n)_{n\geq 0}$ est bien associée a un système d'itération de fonctions aléatoires, l'étude de cette chaîne se fait par celle des IFS.

L'opérateur Q est markovien sur $[0,1]\times[0,1]$. Il agit donc sur $\mathbb{L}^\infty([0,1]\times[0,1])$. Considérons l'opérateur Q^* dans $\mathbb{L}^1([0,1]\times[0,1])$ l'adjoint de Q relativement à la mesure de Lebesgue sur $[0,1]\times[0,1]$, défini pour toutes $\varphi\in\mathbb{L}^1([0,1]\times[0,1])$, $\psi\in\mathbb{L}^\infty([0,1]\times[0,1])$, par

$$\iint_{[0,1]\times[0,1]} \varphi(x_1,x_2)Q\psi(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 = \iint_{[0,1]\times[0,1]} Q^*\varphi(x_1,x_2)\psi(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2.$$

Un calcul direct montre que pour toute fonction $\varphi\in\mathbb{L}^1([0,1]\times[0,1])$ et tout $x=(x_1,x_2)\in[0,1]\times[0,1]$ on a

$$\begin{aligned}
Q^*\varphi(x) = Q^*\varphi(x_1,x_2) = &\iint_{[x_1,1]\times[x_2,1]} \frac{p_{00}(z_1,z_2)}{z_1z_2}\varphi(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2 \\
&+ \iint_{[x_1,1]\times[0,x_2]} \frac{p_{01}(z_1,z_2)}{z_1(1-z_2)}\varphi(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2 \\
&+ \iint_{[0,x_1]\times[x_2,1]} \frac{p_{10}(z_1,z_2)}{(1-z_1)(z_2)}\varphi(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2 \\
&+ \iint_{[0,x_1]\times[0,x_2]} \frac{p_{11}(z_1,z_2)}{(1-z_1)(1-z_2)}\varphi(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2.
\end{aligned} \tag{4.7}$$

Notons que (4.7) est valide pour toute fonction boréliennes $\varphi \in \mathbb{L}^1([0,1] \times [0,1])$. Donc si $\varphi : [0,1] \times [0,1] \to \mathbb{R}^+$ est dans $\mathbb{L}^1([0,1] \times [0,1])$ et vérifie $Q^*\varphi = \varphi$, alors la mesure de densité φ relativement à la mesure de Lebesgue sur $[0,1] \times [0,1]$ est Q-invariante. La chaîne $(Z_n)_{n\geq 0}$ décrite ici a été considérée par Leng et Ramli dans [40] et Borovkov et McKinlay dans [11] comme modèle de la marche d'un robot dans une chambre rectangulaire. Borovkov et McKinlay [11] ont considéré dans leur article le cas où les fonctions p_{00}, p_{01}, p_{10} et p_{11} présentent des discontinuités dans le pavé $[0,1] \times [0,1]$.

Avant de donner quelques résultats concernant cette chaîne nous avons les propositions suivantes après quelques définitions.

Définition 4.1. *Le produit tensoriel de deux noyaux P_1 et P_2 définis sur $(E, \mathcal{E})$ et $(F, \mathcal{F})$ respectivement est le noyau $P := P_1 \otimes P_2$ sur $(E \times F, \mathcal{E} \otimes \mathcal{F})$ défini pour tout $(A, B) \in E \otimes \mathcal{F}$ et tout $(x, y) \in E \times F$ par :*

$$P\big((x,y), A \times B\big) := P_1 \otimes P_2\big((x,y), A \times B\big) = P_1(x, A) \times P_2(y, B).$$

Remarque 4.1. En notant encore P l'opérateur associé au noyau P, nous avons immédiatement $\forall z = (x,y) \in E \times F, \forall \psi : E \times F \to \mathbb{R}$ tel que, $\psi(x,y) = (f \otimes g)(x,y) = f(x) \times g(y)$ alors $P\psi(z) = P\psi(x,y) = P_1 f(x) \times P_2 g(y)$ et $\forall k \in \mathbb{N}^\star, P^k\psi(x,y) = P_1^k f(x) \times P_2^k g(y)$.

Proposition 4.1. *Si (X_n) et (Y_n) sont deux chaînes de Markov indépendantes sur les espaces mesurables $(E, \mathcal{E})$ et $(F, \mathcal{F})$ de noyaux de transition P_1 et P_2 et de lois stationnaires ν_1 et ν_2, alors le processus $(Z_n)_n$ défini par $Z_n = (X_n, Y_n)$ est une chaîne de Markov sur l'espace produit $E \times F$, de noyau $P := P_1 \otimes P_2$ et de loi stationnaire $\nu_1 \otimes \nu_2$.*

Démonstration.

(a) Cas où E et F sont dénombrables.

Soit $z_0 = (x_0, y_0),\ z_1 = (x_1, y_1),\ \ldots,\ z_n = (x_n, y_n), z_{n+1} = (x_{n+1}, y_{n+1})$ des états de la chaîne $(Z_n)_n$, montrons que

$$\mathbb{P}[Z_{n+1} = z_{n+1} | Z_n = z_n, \ldots,\ Z_0 = z_0] = \mathbb{P}\left[Z_{n+1} = z_{n+1} | Z_n = z_n\right].$$

on a :

$\mathbb{P}\big[Z_{n+1} = z_{n+1} | Z_n = z_n, ..., \ Z_0 = z_0\big] =$

$= \mathbb{P}\big[(X_{n+1}, Y_{n+1}) = (x_{n+1}, y_{n+1}) | (X_n, Y_n) = (x_n, y_n), ..., (X_0, Y_0) = (x_0, y_0)\big]$

$= \mathbb{P}\big[X_{n+1} = x_{n+1}, Y_{n+1} = y_{n+1} | X_n = x_n, ..., X_0 = x_0, Y_n = y_n, ..., Y_0 = y_0\big],$

par indépendance des chaînes (X_n) et (Y_n), on aura

$\mathbb{P}\big[Z_{n+1} = z_{n+1} | Z_n = z_n, ..., \ Z_0 = z_0\big] =$

$$= \frac{\mathbb{P}\big[X_{n+1} = x_{n+1}, X_n = x_n, ..., X_0 = x_0, Y_{n+1} = y_{n+1}, Y_n = y_n, ..., Y_0 = y_0\big]}{\mathbb{P}\big[X_n = x_n, ..., X_0 = x_0, Y_n = y_n, ..., Y_0 = y_0\big]}$$

$$= \frac{\mathbb{P}\big[X_{n+1} = x_{n+1}, X_n = x_n, ..., X_0 = x_0\big]\mathbb{P}\big[Y_{n+1} = y_{n+1}, Y_n = y_n, ..., Y_0 = y_0\big]}{\mathbb{P}\big[X_n = x_n, ..., X_0 = x_0\big]\mathbb{P}\big[Y_n = y_n, ..., Y_0 = y_0\big]}$$

$= \mathbb{P}\big[X_{n+1} = x_{n+1}, X_n = x_n, ..., X_0 = x_0\big] \times \mathbb{P}\big[Y_{n+1} = y_{n+1}, Y_n = y_n, ..., Y_0 = y_0\big].$

Et comme (X_n) et (Y_n) sont des chaînes de Markov alors,

$$\begin{aligned}
\mathbb{P}\big[Z_{n+1} = z_{n+1} | Z_n = z_n, ..., \ Z_0 = z_0\big] &= \mathbb{P}\big[X_{n+1} = x_{n+1} | X_n = x_n\big] \times \mathbb{P}\big[Y_{n+1} = y_{n+1} | Y_n = y_n\big] \\
&= \mathbb{P}\big[(X_{n+1}, Y_{n+1}) = (x_{n+1}, y_{n+1}) | (X_n, Y_n) = (x_n, y_n)\big] \\
&= \mathbb{P}\left[Z_{n+1} = z_{n+1} | Z_n = z_n\right] \\
&= P_1(x_n, x_{n+1}) \times P_2(y_n, y_{n+1}) \\
&= P_1 \otimes P_2\big((x_n, y_n), (x_{n+1}, y_{n+1})\big) \\
&= P_1 \otimes P_2\big(z_n, z_{n+1}\big).
\end{aligned}$$

Par conséquent $(Z_n)_n$ est une chaîne de Markov de noyau $P_1 \otimes P_2$.

Soit ν_1 (resp. ν_2) une loi stationnaire pour (X_n) (resp. (Y_n)) montrons que $\nu_1 \otimes \nu_2$ est une loi stationnaire pour $(Z_n)_n$.

La mesure ν_1 est stationnaire pour P_1 i.e. $\forall x_1 \in E, \ \nu_1(x_1) = \sum_{x \in E} \nu_1(x) P_1(x, x_1).$

La mesure ν_2 est stationnaire pour P_2 i.e. $\forall y_1 \in F, \ \nu_2(y_1) = \sum_{y \in F} \nu_2(y) P_2(y, y_1).$

On a alors

$$\begin{aligned}
(\nu_1 \otimes \nu_2)P(x_1, y_1) &= \sum_{(x,y)\in E\times F} \nu_1(x)\nu_2(y)P\big((x,y),(x_1,y_1)\big) \\
&= \sum_{x\in E, y\in F} \nu_1(x)P_1(x,x_1)\nu_2(y)P_2(y,y_1) \\
&= \sum_{x\in E} \nu_1(x)P_1(x,x_1)\sum_{y\in E}\nu_2(y)P_2(y,y_1) \\
&= \nu_1(x_1)\nu_2(y_1) = \nu_1\otimes\nu_2(x_1,y_1).
\end{aligned}$$

Donc $\nu_1 \otimes \nu_2$ est bien une loi stationnaire pour P.

(b) Cas où E et F sont quelconques.

Par définition de l'espérance conditionnelle, on a d'une part, pour $A\times B$, $A_n\times B_n$, ..., $A_0\times B_0 \in \mathcal{E}\otimes\mathcal{F}$:

$$\int_{\{Z_n\in A_n\times B_n,\ldots,Z_0\in A_0\times B_0\}} \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n,\ldots,Z_0]\mathrm{d}\mathbb{P} = \int_{\{\mathrm{Z_n\in A_n\times B_n),\ldots,Z_0\in A_0\times B_0}\}} \mathbb{1}_{\mathrm{A\times B}}(\mathrm{Z_{n+1}})\mathrm{d}\mathbb{P}$$

Par indépendance des chaînes (X_n) et (Y_n) on a :

$$\int_{\{Z_n\in A_n\times B_n,\ldots,Z_0\in A_0\times B_0\}} \mathbb{1}_{A\times B}(Z_{n+1})\mathrm{d}\mathbb{P} = \int_{\{\mathrm{(X_n,Y_n)\in A_n\times B_n,\ldots,(X_0,Y_0)\in A_0\times B_0}\}} \mathbb{1}_{\mathrm{A\times B}}(\mathrm{X_{n+1},Y_{n+1}})\mathrm{d}\mathbb{P}$$

$$= \int_{\{\mathrm{X_n\in A_n,\ldots,X_0\in A_0}\}\cap\{\mathrm{Y_n\in B_n,\ldots,Y_0\in B_0}\}} \mathbb{1}_{\mathrm{A}}(\mathrm{X_{n+1}})\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\mathrm{d}\mathbb{P}$$

$$= \int_{\{\mathrm{X_n\in A_n,\ldots,X_0\in A_0}\}} \mathbb{E}[\mathbb{1}_{\mathrm{A}}(\mathrm{X_{n+1}})\,|\,\mathrm{X_n,\ldots,X_0}]\mathrm{d}\mathbb{P}\times\int_{\{\mathrm{Y_n\in B_n,\ldots,Y_0\in B_0}\}} \mathbb{E}[\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\,|\,\mathrm{Y_n,\ldots,Y_0}]\mathrm{d}\mathbb{P}$$

En utilisant la propriétés de Markov on aura :

$$\int_{\{Z_n\in A_n\times B_n,\ldots,Z_0\in A_0\times B_0\}} \mathbb{1}_{A\times B}(Z_{n+1})\mathrm{d}\mathbb{P} = \int_{\{\mathrm{X_n\in A_n,\ldots,X_0\in A_0}\}} \mathbb{E}[\mathbb{1}_{\mathrm{A}}(\mathrm{X_{n+1}})\,|\,\mathrm{X_n}]\mathrm{d}\mathbb{P}\times\int_{\{\mathrm{Y_n\in B_n,\ldots,Y_0\in B_0}\}} \mathbb{E}[\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\,|\,\mathrm{Y_n}]\mathrm{d}\mathbb{P}.$$

En appliquant Fubini, on aura :

$$\int_{\{Z_n\in A_n\times B_n,\ldots,Z_0\in A_0\times B_0\}} \mathbb{1}_{A\times B}(Z_{n+1})\mathrm{d}\mathbb{P} = \int_{\{\mathrm{Z_n\in A_n\times B_n,\ldots,Z_0\in A_0\times B_0}\}} \mathbb{E}[\mathbb{1}_{\mathrm{A}}(\mathrm{X_{n+1}})\,|\,\mathrm{X_n}]\times\mathbb{E}[\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\,|\,\mathrm{Y_n}]\mathrm{d}\mathbb{P}.$$

Par unicité de l'espérance conditionnelle on obtient :

$$\mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n,...,Z_0] = \mathbb{E}[\mathbb{1}_A(X_{n+1})\,|\,X_n] \times \mathbb{E}[\mathbb{1}_B(Y_{n+1})\,|\,Y_n] \tag{4.8}$$

D'autre part, montrons que

$$\mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n] = \mathbb{E}[\mathbb{1}_A(X_{n+1})\,|\,X_n] \times \mathbb{E}[\mathbb{1}_B(Y_{n+1})\,|\,Y_n].$$

On a pour $A\times B, C\times D \in \mathcal{E}\otimes\mathcal{F}$:

$$\begin{aligned}\int_{\{Z_n\in C\times D\}} \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n]\mathrm{d}\mathbb{P} &= \int_{\{Z_n\in C\times D\}} \mathbb{1}_{A\times B}(Z_{n+1})\mathrm{d}\mathbb{P} = \int_{\{(\mathrm{X_n},\mathrm{Y_n})\in \mathrm{C}\times\mathrm{D}\}} \mathbb{1}_{\mathrm{A}\times\mathrm{B}}(\mathrm{X_{n+1}},\mathrm{Y_{n+1}})\mathrm{d}\mathbb{P}\\ &= \int_{\{X_n\in C\}\cap\{Y_n\in D\}} \mathbb{1}_A(X_{n+1})\mathbb{1}_B(Y_{n+1})\mathrm{d}\mathbb{P}.\end{aligned}$$

Par indépendance des chaînes (X_n) et (Y_n) on aura :

$$\begin{aligned}\int_{\{Z_n\in C\times D\}} \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n]\mathrm{d}\mathbb{P} &= \int_{\{X_n\in C\}} \mathbb{1}_A(X_{n+1})\mathrm{d}\mathbb{P} \times \int_{\{\mathrm{Y_n}\in\mathrm{D}\}} \mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\mathrm{d}\mathbb{P}\\ &= \int_{\{X_n\in C\}} \mathbb{E}[\mathbb{1}_A(X_{n+1})\,|\,X_n]\mathrm{d}\mathbb{P} \times \int_{\{\mathrm{Y_n}\in\mathrm{D}\}} \mathbb{E}[\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\,|\,\mathrm{Y_n}]\mathrm{d}\mathbb{P}.\end{aligned}$$

En utilisant Fubini, on obtient :

$$\int_{\{Z_n\in C\times D\}} \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n]\mathrm{d}\mathbb{P} = \int_{\{\mathrm{X_n}\in\mathrm{C},\mathrm{Y_n}\in\mathrm{D}\}} \mathbb{E}[\mathbb{1}_{\mathrm{A}}(\mathrm{X_{n+1}})\,|\,\mathrm{X_n}]\times\mathbb{E}[\mathbb{1}_{\mathrm{B}}(\mathrm{Y_{n+1}})\,|\,\mathrm{Y_n}]\mathrm{d}\mathbb{P}.$$

D'où

$$\mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n] = \mathbb{E}[\mathbb{1}_A(X_{n+1})\,|\,X_n] \times \mathbb{E}[\mathbb{1}_B(Y_{n+1})\,|\,Y_n]. \tag{4.9}$$

Des équations (4.8) et (4.9), on déduit que

$$\forall A\times B \in \mathcal{E}\otimes\mathcal{F},\quad \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n,...,Z_0] = \mathbb{E}[\mathbb{1}_{A\times B}(Z_{n+1})\,|\,Z_n].$$

Donc $(Z_n)_n$ est une chaîne de Markov sur $E\times F$ de noyau $P_1\otimes P_2$.

Si ν_1 (resp. ν_2) est une loi stationnaire pour (X_n) (resp. (Y_n)) alors $\nu_1 \otimes \nu_2$ est une loi stationnaire pour $(Z_n)_n$. En effet, pour tout $A \times B \in \mathcal{E} \otimes \mathcal{F}$, on a

$$\begin{aligned}
(\nu_1 \otimes \nu_2)P(A \times B) &= \int_{E\times F} \nu_1(\mathrm{d}x) \times \nu_2(\mathrm{d}y)P((x,y), A \times B) \\
&= \int_{E\times F} \nu_1(\mathrm{d}x) \times \nu_2(\mathrm{d}y)P_1 \otimes P_2((x,y), A \times B) \\
&= \int_{E\times F} \nu_1(\mathrm{d}x)P_1(x,A) \times \nu_2(\mathrm{d}y)P_2(y,B) \\
&= \int_{E} \nu_1(\mathrm{d}x)P_1(x,A) \times \int_F \nu_2(\mathrm{d}y)P_2(y,B) \\
&= \nu_1(A) \times \nu_2(B) = (\nu_1 \otimes \nu_2)(A \times B)
\end{aligned}$$

□

Proposition 4.2. *Soient $(X_n)_n$ et $(Y_n)_n$ deux chaînes de Markov sur E et F, respect. de noyaux de transition P_1 et P_2 respect. et $(Z_n)_n$ la chaîne de Markov sur l'espace produit $E \times F$ définie par $Z_n := (X_n, Y_n), \forall n$ et de noyau de transition P tel que $P = P_1 \otimes P_2$. Alors $(X_n)_n$ et $(Y_n)_n$ sont indépendantes.*

Démonstration. Comme $(Z_n) := ((X_n, Y_n))_n$ est une chaîne de Markov alors pour toutes fonctions mesurables $f : E \to \mathbb{R}$ et $g : F \to \mathbb{R}$ et $\forall \;\; z_0 = (x_0, y_0), ..., z_n = (x_n, y_n) \in E \times F$, on a

$$\mathbb{E}_{z_0}\big[f(X_1,...,X_n)g(Y_1,...,Y_n)\big] = \int f(x_1,...,x_n)g(y_1,...,y_n)P(z_0, dz_1)P(z_1, \mathrm{d}z_2)...P(z_{n-1}, \mathrm{d}z_n).$$

Comme $P = P_1 \otimes P_2$ alors on a

$$\begin{aligned}
&\mathbb{E}_{z_0}\big[f(X_1,...,X_n)g(Y_1,...,Y_n)\big] = \\
&= \int f(x_1,...,x_n)g(y_1,...,y_n)P_1(x_0, dx_1)P_2(y_0, \mathrm{d}y_1)...P_1(x_{n-1}, \mathrm{d}x_n)P_2(y_{n-1}, \mathrm{d}y_n) \\
&= \int f(x_1,...,x_n)P_1(x_0, dx_1)...P_1(x_{n-1}, \mathrm{d}x_n) \int g(y_1,...,y_n)P_1(y_0, dy_1)...P_1(y_{n-1}, \mathrm{d}y_n) \\
&= \mathbb{E}_{x_0}\big[f(X_1,...,X_n)\big]\mathbb{E}_{y_0}\big[g(Y_1,...,Y_n)\big].
\end{aligned}$$

Donc $(X_n)_n$ et $(Y_n)_n$ sont indépendantes.

Grâce à la propriété connue que si X et Y sont deux variables aléatoires réelles tel que pour toutes fonctions f et g mesurables et bornées sur $\mathbb{R}$, $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$, alors Xet Y sont indépendantes. □

Exemples.

Nous donnons ici quelques exemples et résultats partiels concernant la chaîne de Markov $(Z_n)_{n\geq 0}$ sur $[0,1]\times[0,1]$ décrite au début de cette section. Cette chaîne modélise les mouvements d'une particule entre quatre points attractifs formant un rectangle.

1. Lorsque les fonctions poids p_{00}, p_{01}, p_{10} et p_{11} vérifient : $p_{00}(x_1,x_2)=x_1x_2$, $p_{01}(x_1,x_2)=x_1(1-x_2), p_{10}(x_1,x_2)=(1-x_1)(x_2)$ et $p_{11}(x_1,x_2)=(1-x_1)(1-x_2)$ pour tout $x=(x_1,x_2)\in[0,1]\times[0,1]$. La chaîne $(Z_n)_{n\geq 0}$ est une suite de variables aléatoires indépendantes et de loi uniforme sur $[0,1]\times[0,1]$.

2. On munit $[0,1]\times[0,1]$ de la distance euclidienne d. Soit $C([0,1]\times[0,1])$ l'espace des fonctions continues de $[0,1]\times[0,1]$ dans $\mathbb{C}$. On note $\mathcal{H}_\alpha([0,1]\times[0,1])$ l'espace des fonctions α-Hölder continues de $[0,1]\times[0,1]$ dans $\mathbb{C}$ défini par

$$\mathcal{H}_\alpha([0,1]\times[0,1]) := \{f\in C([0,1]\times[0,1]) \mid \|f\|_\alpha := |f|_\infty + m_\alpha(f) < +\infty\}$$

   où $m_\alpha(f) := \sup\limits_{\substack{x,y\in[0,1]\times[0,1]\\ x\neq y}} \dfrac{|f(x)-f(y)|}{d(x,y)^\alpha} < \infty$ et $|\cdot|_\infty$ est la norme de la convergence uniforme.

   Supposons à présent que la fonction p_{00} appartient à $\mathcal{H}_\alpha([0,1]\times[0,1])$ et vérifie

$$\forall x\in[0,1]\times[0,1] \qquad p_{00}(x)>0. \tag{4.10}$$

   Dans ce cas, la chaîne $(Z_n)_{n\geq 0}$ admet une unique mesure de probabilité invariante sur $[0,1]\times[0,1]$ car les hypothèses **H1**, **H2** et **H3** de la Proposition 3.4 sont satisfaites (avec la famille de mesures $(\mu_x)_{x\in[0,1]\times[0,1]}$ donnée par la formule (4.6)), en effet

(a) *Hypothèse* ***H1*** : Il existe $\alpha = 1$ tel que pour tous $x, y \in [0,1], x \neq y$, on a

$$\begin{aligned}\int_{\mathbb{L}\mathrm{ip}([0,1]\times[0,1])} \Big(\frac{d(T\cdot x, T\cdot y)}{d(x,y)}\Big)^{\alpha} \mu_x(\mathrm{d}T) \quad \leq \quad & p_{00}(x) \iint_{[0,1]\times[0,1]} \sqrt{t^2+t'^2}\mathrm{d}t\mathrm{d}t' + \\ & p_{01}(x) \iint_{[0,1]\times[0,1]} \sqrt{t^2+t'^2}\mathrm{d}t\mathrm{d}t' + \\ & p_{10}(x) \iint_{[0,1]\times[0,1]} \sqrt{t^2+t'^2}\mathrm{d}t\mathrm{d}t' + \\ & p_{11}(x) \iint_{[0,1]\times[0,1]} \sqrt{t^2+t'^2}\mathrm{d}t\mathrm{d}t' \\ < \quad & 1,\end{aligned}$$

car le coefficient de lipschitz des applications $HH_{(t,t')}, HA_{(t,t')}, AH_{(t,t')}$ et $AA_{(t,t')}$ est égal à $\sqrt{t^2+t'^2}$.

(b) *Hypothèse* ***H2*** : pour tous $x, y \in [0,1]\times[0,1], x \neq y$, on a

$$\begin{aligned}\frac{|\mu_x - \mu_y|}{d(x,y)^{\alpha}} \quad \leq \quad & \frac{|p_{00}(x) - p_{00}(y)|}{d(x,y)^{\alpha}} \Big| \iint_{[0,1]\times[0,1]} \delta_{HH_{(t,t')}} \mathrm{d}t\mathrm{d}t' \Big| \\ + \quad & \frac{|p_{01}(x) - p_{01}(y)|}{d(x,y)^{\alpha}} \Big| \iint_{[0,1]\times[0,1]} \delta_{HA_{(t,t')}} \mathrm{d}t\mathrm{d}t' \Big| \\ + \quad & \frac{|p_{10}(x) - p_{10}(y)|}{d(x,y)^{\alpha}} \Big| \iint_{[0,1]\times[0,1]} \delta_{AH_{(t,t')}} \mathrm{d}t\mathrm{d}t' \Big| \\ + \quad & \frac{|p_{11}(x) - p_{11}(y)|}{d(x,y)^{\alpha}} \Big| \iint_{[0,1]\times[0,1]} \delta_{AA_{(t,t')}} \mathrm{d}t\mathrm{d}t' \Big| \\ \leq \quad & m_{\alpha}(p_{00}) + m_{\alpha}(p_{01}) + m_{\alpha}(p_{10}) + m_{\alpha}(p_{11}) < \infty.\end{aligned}$$

(c) *Hypothèse* ***H3*** : pour tout $x \in [0,1]\times[0,1]$, on a

$$\mu_x \geq \delta \iint_{[0,1]\times[0,1]} \delta_{HH_{(t,t')}} \mathrm{d}t\mathrm{d}t'$$

avec $\delta := \inf_{x\in[0,1]\times[0,1]} p_{00}(x) > 0$.

La fonction $HH_{(0,0)} : x \mapsto (0,0)$ appartient au support de $\mu := \iint_{[0,1]\times[0,1]} \delta_{HH_{(t,t')}} \mathrm{d}t\mathrm{d}t'$ et fournit de façon évidente une suite contractante au semi-groupe T_{μ}.

Remarquons que nous avons le même résultat lorsque l'une des propriétés suivantes est vérifie.

- La fonction p_{01} appartient à $\mathcal{H}_\alpha([0,1]\times[0,1])$ et $\forall x \in [0,1]\times[0,1] \qquad p_{01}(x) > 0.$
- La fonction p_{10} appartient à $\mathcal{H}_\alpha([0,1]\times[0,1])$ et $\forall x \in [0,1]\times[0,1] \qquad p_{10}(x) > 0.$
- La fonction p_{11} appartient à $\mathcal{H}_\alpha([0,1]\times[0,1])$ et $\forall x \in [0,1]\times[0,1] \qquad p_{11}(x) > 0.$

3. Lorsque $p_{00}(x_1,x_2) = p_1(x_1)\times p_2(x_2)$, $p_{01}(x_1,x_2) = p_1(x_1)\times q_2(x_2)$, $p_{10}(x_1,x_2) = q_1(x_1)\times p_2(x_2)$ et $p_{11}(x_1,x_2) = q_1(x_1)\times q_2(x_2)$ avec $p_1(x_1)+q_1(x_1) = 1$ et $p_2(x_1) + q_2(x_1) = 1$, la chaîne $(Z_n)_n$ est de la forme $Z_n = (Z_n^1, Z_n^2)$ avec $(Z_n^1)_n$et $(Z_n^2)_n$ sont deux chaînes indépendantes de noyaux Q_1 et Q_2 définis : pour toutes fonctions boréliennes bornées $f, g : [0,1] \to \mathbb{C}$ et tous $x_1, x_2 \in [0,1]$ par :
$$Q_1 f(x_1) = p_1(x_1)\int_0^1 f(tx_1)\mathrm{d}t + q_1(x_1)\int_0^1 f(tx_1+1-t)\mathrm{d}t.$$
et
$$Q_2 g(x_2) = p_2(x_2)\int_0^1 g(t'x_2)\mathrm{d}t' + q_2(x_2)\int_0^1 g(t'x_2+1-t')\mathrm{d}t'.$$
Les mesures invariantes de la chaîne (Z_n) dans ce cas sont donc des produits de mesures invariantes et les résultats obtenus dans le chapitre précédent suffisent pour conclure.

4. Lorsque les fonctions poids p_{00}, p_{01}, p_{10} et p_{11} sont constantes, la chaîne $(Z_n)_{n\geq 0}$ admet dans ce cas une unique mesure de probabilité invariante sur $[0,1]\times[0,1]$ car la condition (3.2) de la proposition 3.3 est satisfaite et nous avons :

   (a) Si $p_{11} = 1$ l'unique mesure de probabilité invariante pour $(Z_n)_{n\geq 0}$ est $\delta_{(1,1)}$ la masse de Dirac en $(1,1)$. De plus la chaîne $(Z_n)_{n\geq 0}$ converge $\mathbb{P}_x$-presque sûrement vers $(1,1)$.
   En effet dans ce cas la chaîne $(Z_n)_{n\geq 0}$ est un produit de deux chaînes $(Z_n^1)_n$ et $(Z_n^2)_n$ de Diaconis-Freedman indépendantes convergentes vers 1. Puisque,
$$\begin{aligned}\forall x \in [0,1]\times[0,1] \quad \mathbb{E}_x(1-Z_1^1) &= 1-\mathbb{E}_x(Z_1^1) = 1-\Big(\frac{1+x_1}{2}\Big)\\ &= \frac{1-x_1}{2}\end{aligned}$$

d'où $\mathbb{E}_x(1 - Z_{n+1}/\mathcal{F}_n) = \dfrac{1 - Z_n}{2}$ pour tout $n \geq 0$ par itération on obtient

$$\forall n \geq 0, \forall x \in [0,1] \times [0,1] \qquad \mathbb{E}_x(1 - Z_{n+1}) = \left(\frac{1}{2}\right)^{n+1} (1 - x_1) \leq \left(\frac{1}{2}\right)^{n+1}.$$

et donc $\sum_{n\geq 0} \mathbb{E}_x\Big(1 - Z_n^1\Big) = \mathbb{E}_x\Big(\sum_{n\geq 0} 1 - Z_n^1\Big) < +\infty$, par suite la suite (Z_n^1) converge donc $\mathbb{P}_x$-a.s. vers 1.

(b) Si $p_{00} = p_{01} = 0$ et $p_{10},\ p_{11} > 0$ (ou, de façon symétrique $p_{10} = p_{11} = 0$ et $p_{00},\ p_{01} > 0$) l'unique mesure de probabilité invariante pour $(Z_n)_{n\geq 0}$ est $\delta_1 \otimes \mathcal{B}(p_{11}, p_{10})$ où $\mathcal{B}(p_{11}, p_{10})$ est la loi Beta de paramètres p_{11} et p_{10}. Puisque $(Z_n)_{n\geq 0}$ est un produit de deux chaînes $(Z_n^1)_n$ et $(Z_n^2)_n$ de Diaconis-Freedman indépendantes. En effet, d'après les résultats obtenus dans le chapitre précédent $(Z_n^1)_n$ converge $\mathbb{P}_x$-presque sûrement vers 1 et $(Z_n^1)_n$ converge vers ν_2 la loi $\mathcal{B}(p_{11}, p_{10})$. Pour toutes fonctions continues bornées f et g, nous avons :

$$\begin{aligned}
\mid \mathbb{E}_x(f(Z_n^1)g(Z_n^2) - f(1)\nu_2(g)) \mid &\leq \mid \mathbb{E}_x(f(Z_n^1)g(Z_n^2) - f(1)g(Z_n^2)) \\
&- \mathbb{E}_x(f(1)g(Z_n^2) - f(1)\nu_2(g)) \mid \\
&\leq |g|_\infty \mathbb{E}_x \mid f(Z_n^1) - f(1) \mid \\
&+ f(1)\mathbb{E}_x \mid f(Z_n^2) - \nu_2(g) \mid .
\end{aligned}$$

Comme $\lim_{n\to\infty} \mathbb{E}_x \mid f(Z_n^2) - \nu_2(g) \mid = 0$ et $\lim_{n\to\infty} \mathbb{E}_x \mid f(Z_n^1) - f(1) \mid = 0$, alors

$$\lim_{n\to\infty} \mathbb{E}_x(f(Z_n^1)g(Z_n^2)) = f(1)\nu_2(g),$$

et donc $(Z_n)_n$ converge en loi vers $\delta_1 \otimes \mathcal{B}(p_{11}, p_{10})$.

(c) Si $p_{01} = p_{10} = 0$ et $p_{00},\ p_{11} > 0$ (ou, de façon symétrique $p_{00} = p_{11} = 0$ et $p_{01},\ p_{10} > 0$) nous conjecturons que l'unique mesure de probabilité Q-invariante admet une densité par rapport à la mesure de Lebesgue .

La recherche de cette densité revient à la résolution de l'équation intégro-différentielle suivante obtenue à partir de l'équation (4.7).

$$\varphi(x_1, x_2) = \iint_{[x_1,1]\times[x_2,1]} \frac{p_{00}}{z_1 z_2}\varphi(z_1, z_2)\mathrm{d}z_1\mathrm{d}z_2 + \iint_{[0,x_1]\times[0,x_2]} \frac{p_{11}}{(1-z_1)(1-z_2)}\varphi(z_1, z_2)\mathrm{d}z_1\mathrm{d}z_2.$$

.

De façon similaire au cas unidimensionnel, nous conjecturons la proposition suivante.

Proposition 4.3. *Soit $(Z_n)_{n\geq 0}$ la chaîne de Markov sur $[0,1]\times[0,1]$ d'opérateur de transition Q défini par la formule (4.5) avec des fonctions poids p_{00}, p_{01},p_{10} et p_{11} α-Hölder. Alors*

1. *lorsque $p_{00}(0,0) < 1$,$p_{01}(0,1) < 1$, $p_{10}(1.0) < 1$ et $p_{11}(1,1) < 1$ il existe une unique mesure de probabilité Q-invariante sur $[0,1]\times[0,1]$. Celle-ci est absolument continue par rapport à la mesure de Lebesgue sur $[0,1]\times[0,1]$.*

2. *Lorsque $p_{00}(0,0) = 1$,$p_{01}(0,1) > 0$, $p_{10}(1.0) > 0$ et $p_{11}(1,1) > 0$, il existe une unique mesure de probabilité Q-invariante sur $[0,1]\times[0,1]$ qui est la masse de Dirac $\delta_{(0,0)}$.*

3. *Lorsque $p_{01}(0,1) = 1$,$p_{00}(0,0) > 0$, $p_{10}(1.0) > 0$ et $p_{11}(1,1) > 0$, la masse de Dirac $\delta_{(0,1)}$ est l'unique mesure de probabilité Q-invariante sur $[0,1]\times[0,1]$.*

4. *Lorsque $p_{10}(1.0) = 1$, $p_{00}(0,0) > 0$, $p_{01}(0,1) > 0$ et $p_{11}(1,0) > 0$, la masse de Dirac $\delta_{(1,0)}$ est l'unique mesure de probabilité Q-invariante sur $[0,1]\times[0,1]$.*

5. *Lorsque $p_{11}(1.1) = 1$, $p_{00}(0,0) > 0$, $p_{01}(0,1) > 0$ et $p_{10}(1,0) > 0$, la masse de Dirac $\delta_{(1,1)}$ est l'unique mesure de probabilité Q-invariante sur $[0,1]\times[0,1]$.*

Conclusion et perspectives

Après l'étude de la chaîne de Diaconis-Freedman sur [0, 1] et les quelques idées de son extensions, nous concluons ce mémoire de thèse en notant l'apport des méthodes d'analyse fonctionnelle au Calcul de Probabilités et à l'étude des processus itératifs, à partir du théorème de Ionescu Tulcea et Marinescu .

Au chapitre des applications utiles dans la vie quotidienne, nous signalons que ce type de chaînes de Markov, est à la base de nombreuses applications, en particulier, en robotique. En effet il est à noter, que cela permet la modélisation de la marche d'un robot dans une chambre carrée (cf. McKinlay et Borovkov [11]).

Nous signalons aussi que la chaîne de Diaconis-Freedman permet la modélisation du jeu donnant-donnant.

Les perspectives immédiates ouvertes par ce travail résident dans l'extension de la chaîne de Diaconis-Freedman que nous résumons en deux points :

▶ Modification du noyau de transition en remplaçant la loi uniforme par une loi admettant une densité, dans le cas où la particule est attirée par deux points attractifs (0 et 1 dans le cas introduit par Diaconis-Freedman).

▶ Étude du cas où le mouvement de la particule a lieu en dimension supérieure avec au moins 3 points attractifs ou bien en remplaçant la loi uniforme en dimension 1 par son analogue en dimension supérieure).

Au plan global, par rapport à notre première motivation, ce travail nous fournit un exemple d'étude de chaîne itérative. Par rapport à notre motivation de départ, il s'agit de voir l'impact de ces résultats au plan méthodologique sur l'étude des théories de l'apprentissage.

Annexe : Programmes de simulation

On donne ici les programmes utilisés au chapitre 1 pour simuler les lois des chaînes de Markov des exemples 1 et 2. Ces programmes ont été écrits en Scilab sous environnement ubuntu 16.04 LTS.

Programme 1

```
clear;
clc;
close;
a=1/2;
x=0;
n=1000000;
xx=zeros(n,1);
for i=1 :n, u= grand(1,1,'def');
if u<(1/2) then y(i)=0; else y(i)=1/2;
end
x=a*x+y(i);
xx(i)=x;
end
xtitle(" n=1000000 et p=0.5" )
histplot(-0.1 :.01 :1.1,xx,style=2), xgrid
```

Programme 2

```
clear;
clc;
close;
function y = f(x)
if x >= 0 & x <= 1 then y = x/(sqrt(2));
end if x >= 1&x <= 2 then y = (sqrt(2*x - (x*x)/2 - 1));
```

end

endfunction

function $z = g(x)$

if $x >= 0$ &$x <= 1$ then $z = 2 - (sqrt(1 - (x * x/2)));$

end

if $x >= 1\&x <= 2$ then $z = 2 - (sqrt(2 - 2 * x + (x * x/2)));$

end

endfunction

$n = 1000;$

$x = 0;$

$xx = zeros(n, 1);$

for $i = 1 : n, u = grand(1, 1,' def')$

if $u < (1/2)$ then $x(i) = f(x); else\ x(i) = g(x)$; end

$x = x(i),$

$xx(i) = x;$

end

xtitle("$n = 1000$ ")

histplot(-0.1 :.01 :2.1, xx, style=2), xgrid

Bibliographie

[1] Alsmeyer, G. & Fuh, C.D. (2001). *Limit theorems for iterated random functions by regenerative methods,* Stochastic Process. Appl., 96(1) : 123–142.

[2] Athreya, K. B. & Stenflo, O. (2003). *Perfect sampling for Doeblin chains,* Sankhya, 65(4) : 763–777.

[3] Barnsley, M. F., Demko S. G., Elton, J. H. & Geronimo, J. S. (1988). *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities* Ann. Inst. H. Poincaré Probab. Statist. 24 , no. 3, 367–394.

[4] Barnsley, M. F. & Elton, J.H. (1988). *A new class of Markov processes for image encoding* Adv. in Appl. Probab. 20 , no. 1, 14–32.

[5] Barnsley, M.F., Elton, J. H. & Hardin, D. P. (1989). *Recurrent iterated function systems. Fractal approximation* Constr. Approx. 5 , no. 1, 3–31.

[6] Billingsley, P. (1968). *Convergence of probability Measures,* Wiley, New York.

[7] Benaïm, M. & El Karaoui, N. (2007). *Promenades aléatoire, Chaines de Markov et simulation, Martingales et stratégies,* Editions de l'Ecole Polytechnique, Paris.

[8] Ben Alaya, M. (1992). *Les Théorèmes ergodiques en simulation,* Thèse, Ecole Nationale des Ponts et Chaussées .

[9] Bhattacharya, R. & Majumdar, M. (2007). *Random dynamical systems : theory and applications.* Cambridge University Press, Cambridge.

[10] Borovkov, A. A. (1998) *Ergodicity and Stability of Stochastic Processes,* Wiley, New York.

[11] Borovkov, K. & McKinlay, S. (2016). *On explicit form of the stationary distributions for a class of bounded Markov chains,* J. Appl. Probab. vol. 53, no. 1, 231–243.

[12] Boudiba, M.A. (1985).*La chaîne de Markov* $X_n = |X_{n-1} - Y_n|$ *et les chaînes associées, avec* $(Y_n)_n$ *une suite i.i.d. de variables aléatoires,* $\mathbb{E}[Y_n] < \infty$ Compt. Rend. Acad. Sciences, Paris.

[13] Boudiba, M.A. (1986).*Sur la chaîne de Feller* $X_n = |X_{n-1} - Y_n|$ *et les chaînes associées,* Thèse, Université Paul Sabatier de Toulouse.

[14] Leo Breiman (1992). *Probability.* Society for Industrial and Applied Mathematics Philadelphia.

[15] Brezis, H. (1983). *Analyse fonctionnelle ,Théorie et applications,* Masson, Paris.

[16] Brunel, A. & Revuz, D. (1974). *Quelques applications probabilistes de la quasi-compacité,* Annales de l'institut Henri Poincarré, Section B, Vol. X. no. 3.

[17] Conze, J.P. et Raugi, A. (1990). *Fonctions harmoniques pour un opérateur de transition et applications,* Bull. Soc.math. France., vol.118, p.273-310.

[18] Dunford, N. & Schwarz, J. (1958). *Linear Operators,* Vol. 1. Interscience.

[19] Dudley, R. M. (1989). *Real Analysis and Probability.* Wadsworth & Brooks, Belmont (CA).

[20] Diaconis, P. & Freedman, D. (1999). *Iterated Random Functions* SIAM review, Vol. 41, no. 1.

[21] Diaconis, P. & Shahshahani, M. (1984). *Products of random matrices as they arise in the study of random walks on groups* Technical report no. 229, Departement of Statistics, Stanford University.

[22] Doob, J. L. (1953.) *Stochastic process,* Wiley. New-York.

[23] Elton, J. H. (1987).*An ergodic theorem for iterated maps* Ergodic Theory Dynam. Systems 7 , no. 4, 481–488.

[24] Feller, W. (1966). *Introduction to Probability Theory and Stochastic Process with Applications* 2nd Edition Willey and Sons, N-Y.

[25] Guivarc'h, Y. & Hardy J. (1988). *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov.* Ann. Inst. Henri Poincaré, Vol. 24, no. 1, p. 73-98.

[26] Guivarc'h, Y. & Raugi, A. (1986). *Products of random matrices and convergence theorems,* Contempory Math., A.M.S., vol. 50, 31–54.

[27] Fortet, R. (1978). *Condition de Doeblin et quasi-compacité.* Ann. Inst. Henri Poincaré, XIV, 4, 379-390.

[28] Hennion, H. (1993). *Sur un théorème spectral et son application aux noyaux lipschitziens,* Proceedings of the American Mathematical Society, vol. 118 , 627–634.

[29] Hennion, H. (1995) *Quasi-compacité cas des noyaux lipschitiens et des noyaux markoviens,* Séminaire de Probabilités de Rennes, p. 1–50.

[30] Hennion, H. & Hervé, L. (2001). *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness,* Lecture Notes in Math., vol. 1766, Springer.

[31] Hennion, H. (2006) *Quasi-compactness and absolutely continuous kernels, applications to Markov chains,* ArXiv Mathematics e-prints, (arXiv :math/0606680v1), .

[32] Hervé, L. (1994). *Étude d'opérateurs quasi-compacts positifs. Applications aux opérateurs de transfert,* Ann. Inst. Henri Poincaré, section B, vol. 30, no. 3, p. 437-466.

[33] Hervé, L. (1996). *Transformée en paquets d'ondelettes des signaux stationnaires : comportement asymptotique des densités spectrales,* Rev. Mat. Iberoamericana, vol. 12 no. 3, 653–667.

[34] Hervé, L. (2008). *Quasi-Compactness and mean ergodicity for Markov kernels acting on weighted supremum normed spaces,* Ann. Inst. H. Poincaré Probab. Statist. 44 (6), 1090-1095.

[35] Ionescu-Tulcea, C.T. & Marinescu, G. (1950). *Théorie ergodique pour des classes d'opérations non complètement continues*, Annals of Mathematics, vol. 52 , 140-147.

[36] Kaijser, T. (1981). *On a new contraction condition for random systems with complete connections,* Rev. Rou. de Math. Pure et Appli., vol 26, no. 8, 1075–1117.

[37] Kapica, R. & Sleczka, M. (2017). *Random iteration with place dependent probabilities*, arXiv :1107.0707v3.

[38] Kifer, Y. (1986). *Ergodic theory of random transformations,* Progress in Probability and Statistics, Vol. 10, Birkhäuser Boston Inc.

[39] Leontiev, A. N. (1978). *Activity, Consciousness and Personality* , Prentice-Hall.

[40] Leng, G. & Ramli, M. A. (2010). *The stationary probability density of a class of bounded Markov processes,* Adv. in Appl. Probab, vol. 42, no. 4, 986–993.

[41] Lerma, O. & Lasserre, J.B. (2003). *Markov Chains and Invariant Proba- bilities,* Birkhäuser, Basel.

[42] Letac, G. (1986). *A contraction principle for certain Markov Chains and its applications.* Contemporary Mathematics, Volume 50, American Mathematecal Society, Rhodes Island, N-Y.

[43] Ladjimi, F. & Peigné, M.(2017). *Iterated function systems with place dependent probabilities and application to the Diaconis-Freedman's chain on* $[0,1]$. ArXiv e-prints, arXiv :1707.07237.

[44] Ladjimi, F. & Peigné, M.(2019).*On the asymptotic behavior of the Diaconis-Freedman chain on [0.1].* Stat. & Probab. Letters, Vol. 145, 1–11.

[45] Lin, M.(1974) *On quasi-compact Markov operators,* The Ohio State University, Columbus The Annals of Probability , Vol. 2, no.3, 464–475.

[46] Neveu, J. (1964). *Bases mathématiques du calcul des probabilités.* Masson et Cie.

[47] Norman, M.F. (1972). *Markov processes and learning models.* Academic Press.

[48] Pacheco-Gonzalez, C. G. (2009). *Ergodicity of a bounded Markov chain with attractiveness towards the centre*, Statist. Prob. Lett. 79,2177-2181.

[49] Pacheco-Gonzalez, C. G. & Stoyanov, J. (2008). *A class of Markov chains with beta ergodic distributions*, Math. Scientist 33, 110–119.

[50] Peigné, M. (1993). *Iterated function systems and spectral decomposition of the associated Markov operator*, in Fascicule de probabilités, Publ. Inst. Rech. Math. Rennes, vol. 1993, Univ. Rennes I, http ://www.lmpt.univ-tours.fr/ peigne/fichiers/ifs.pdf.

[51] Peigné M. (1992). *Marches de Markov sur le semi-groupe des contractions de* $\mathbb{R}^d$*. Cas de la marche aléatoire à pas markoviens sur* $(\mathbb{R}^+)^d$ *avec chocs élastiques sur les axes*, Ann. Inst. Henri Poincaré, section B, vol. 28, no. 1, p. 63-94.

[52] Peigné, M. & Woess, W. (2011). *Stochastic dynamical systems with weak contractivity I. Strong and local contractivity* , Colloquium Mathematicum 125 , 1-54.

[53] Peigné, M. & Woess, W.(2011). *Stochastic dynamical systems with weak contractivity properties II. Iteration of Lipschitz mappings,* Colloquium Mathematicum 125 , 55–81.

[54] Mirek, M. (2011). *Heavy tail phenomenom and convergence to stable laws for iterated Lipschitz random maps,* Probability Theory and Related Fields, vol. 151, 705–734.

[55] Pirinsky, C. & Stoyanov, J. (2000). *Random motions, classes of ergodic Markov chains and beta distributions*, Stat. & Probab. Letters 50, 293–304.

[56] Propp, J. & Wilson, D. (1998). *Exact sampling with coupled Markov chains, Random Structures* Algorithms, 9, pp. 223–252.

[57] Propp, J. & Wilson, D.(1998). *How to get a perfectly random sample from a generic Markov chain and generate a random spanning tree of a directed graph* Journal of Algorithms, No 27.

[58] Meyn, S.P. & Tweedie, R.L. (1993). *Markov Chains and Stochastic Stability,* Springer Verlag.

[59] Revuz, A. (1975). *Markov Chains* North Holland.

[60] Rudin, W. (1990). *Functional Analysis,* 2nd. edition, McGraw Hill .

[61] Stenflo, Ö. (2012). *A survey of average contractive iterated function systems,* J. Difference Equ. Appl. 18, no. 8, 1355–1380.

[62] Stenflo, Ö. (2012) *Iterated function systems with a given continuous stationary distribution,* Fractals, Vol 20, no.3 & 4, 197–202.

[63] Tuan-Minh Nguyen (2017) *Random Geometry and Reinforced Jump Processes,* Thèse, Lund University.

[64] Wu, W. B. & Shao, X. (2004) *Limit theorems for iterated random functions,* J. Appl. Probab. 41(2) :425 ? 436.

[65] Yosida, K. (1980). *Functional Analysis.* Springer-Verlag Berlin Heiderlberg.