HAL
open science

Vitesse de mélange et théorèmes limites pour les systèmes dynamiques aléatoires et non-autonomes

Romain Aimino

To cite this version:

Romain Aimino. Vitesse de mélange et théorèmes limites pour les systèmes dynamiques aléatoires et non-autonomes. Mathématiques générales [math.GM]. Université de Toulon, 2014. Français. NNT : 2014TOUL0005 . tel-01090409

HAL Id: tel-01090409

https://theses.hal.science/tel-01090409v1

Submitted on 3 Dec 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE
l'UNIVERSITÉ DE TOULON

Spécialité

Mathématiques

École doctorale Mer et Sciences 548

Présentée par

Romain AIMINO

Pour obtenir le grade de

DOCTEUR en MATHÉMATIQUES

Sujet de la thèse :

Vitesse de mélange et théorèmes limites pour les systèmes dynamiques aléatoires et non-autonomes

soutenue le 23 Octobre 2014

devant le jury composé de :

M. Sandro Vaienti	Directeur de thèse
M. Sébastien Gouëzel	Rapporteur
M. Mark Pollicott	Rapporteur
M. Carlangelo Liverani	Examinateur
M. Claude-Alain Pillet	Examinateur
M. Benoît Saussol	Examinateur
M. Dalibor Volný	Examinateur

Remerciements

Je tiens tout d'abord à remercier Sandro Vaienti pour avoir dirigé cette thèse. Grâce à sa passion débordante pour les sciences et les mathématiques en général, et les systèmes dynamiques en particulier, ainsi que son intérêt et enthousiasme portés à mes recherches, il m'a permis d'aimer et d'apprécier à mon tour ce sujet passionnant. L'idée que partager son bureau avec son directeur de thèse puisse être une expérience agréable pourra sembler surprenante à de nombreux étudiants, mais cela a pourtant été le cas, et je lui suis reconnaissant d'avoir rendu ces trois années aussi appréciables, par sa bienveillance et ses qualités humaines.

Je remercie le Conseil Régional Provence-Alpes-Côte d'Azur pour avoir financé cette thèse.

Mes remerciements vont à Sébastien Gouëzel et Mark Pollicott pour voir accepté la lourde et ingrate tâche de relire et de rapporter ma thèse, dans un intervalle de temps relativement étroit, ainsi que pour leurs nombreuses remarques constructives qui m'ont permis d'améliorer cette thèse. Je remercie également Carlangelo Liverani, Claude-Alain Pillet, Benoît Saussol et Dalibor Volný de me faire l'honneur d'assister à ma soutenance en tant que membres du jury.

J'adresse mes sincères remerciements à mes différents coauteurs : Huyi Hu, Matthew Nicol, Jérôme Rousseau, Michael Todd, Andrew Török et Sandro Vaienti, sans qui mes progrès auraient été tout autre. Collaborer avec eux aura été extrêmement enrichissant, et j'ai beaucoup appris à leurs côtés.

Pour les multiples discussions à caractère mathématique que j'ai pu tenir avec eux, je remercie José Alvès, Wael Bahsoun, Jean-René Chazottes, Pierre Collet, Jean-Pierre Conze, Bastien Fernandez, Jorge Freitas, Stefano Galatolo, Michele Gianfelice, Ian Melbourne, Glenn Merlet, Mikko Stenlund, Damien Thomine, Hongkun Zhang et tout particulièrement Carlangelo Liverani pour avoir débloqué de nombreuses situations par ses conseils avisés.

Je remercie le personnel enseignant du département de mathématiques à l'USTV pour l'attention portée à mon égard, et particulièrement Philippe Briet pour ses nombreux conseils et pour m'avoir aiguillé en master vers les systèmes dynamiques et mon directeur de thèse. Je suis également reconnaissant au personnel administratif et informatique du Centre de Physique Théorique pour leur sympathie, leur compétence et leur disponibilité. J'adresse mes remerciements aux doctorants du Centre de Physique Théorique, du département de mathématiques de l'USTV et de l'IML pour leur bonne humeur et les différents bons moments

passés ensemble.

Je remercie chaleureusement ma famille et mes amis, qui sauront se reconnaître, pour leur soutien sans faille. J'ai une pensée émue pour mes grands parents, Andréa et Yvon, qui nous ont quitté durant l'élaboration de cette thèse. Mes derniers remerciements, et non les moindres, vont à ma maman pour tout ce qu'elle a fait pour moi, durant mais aussi avant cette thèse.

Table des matières

Introduction **7**

1 Théorèmes limites pour les systèmes aléatoires **17**
1.1 Systèmes dynamiques aléatoires : cadre de travail et exemples 17
1.1.1 Cadre abstrait 17
1.1.2 Applications dilatantes, lisses par morceaux, en dimension 1 22
1.1.3 Applications dilatantes et lisses par morceaux en dimension supérieure 25
1.1.4 Applications dilatantes et affines par morceaux 27
1.2 Théorèmes limites par la méthode spectrale 28
1.2.1 Décroissance des corrélations 28
1.2.2 Théorème de la limite centrale avec vitesse de convergence 29
1.2.3 Principe des grandes déviations 36
1.2.4 Théorème de la limite locale 39
1.2.5 Principe d'invariance presque sure 41
1.3 Le théorème de la limite centrale par la méthode des martingales 43
1.4 Lemmes dynamiques de Borel-Cantelli 46
1.4.1 Convergence presque sure 46
1.4.2 Étude des fluctuations 47
1.4.3 Application au problème des cibles rétrécissantes 49
1.5 Version quenched du théorème de la limite centrale 51
1.5.1 Un résultat général 51
1.5.2 Le cas où la mesure μ est préservée par toutes les applications T_ω . . 53
1.5.3 Le cas où la mesure μ n'est pas nécessairement préservée par les applications T_ω 57
1.6 Remarques et questions ouvertes 63

2 Théorèmes limites pour les systèmes non-autonomes **65**
2.1 Systèmes dynamiques non-autonomes 65
2.1.1 Quelques généralités 65
2.1.2 Un cadre fonctionnel abstrait 67
2.1.3 Perte de mémoire exponentielle 68

2.1.4 Application aux systèmes unidimensionnels 69
2.1.5 Théorèmes limites 70
2.1.6 La condition **(Min)** 71
2.2 Applications de l'intervalle avec un point fixe neutre 73
2.2.1 Le modèle et le résultat principal 73
2.2.2 Un cône invariant 73
2.2.3 Une perturbation aléatoire et son noyau associé 75
2.2.4 Distorsion : démonstration du lemme 2.2.8 78
2.3 Remarques et questions ouvertes 85

3 Inégalités de concentration 87
3.1 Inégalités de concentration : définition 87
3.2 Inégalités de concentration pour les systèmes dynamiques non-autonomes . . 92
3.2.1 Le résultat et sa démonstration 92
3.2.2 Applications 99
3.3 Inégalités de concentration pour les systèmes dynamiques aléatoires 106
3.3.1 Le résultat et sa démonstration 106
3.3.2 Applications 111
3.4 Remarques et questions ouvertes 112

4 Lemmes de Borel-Cantelli l'induction de Rauzy 114
4.1 L'induction de Rauzy-Veech-Zorich 114
4.1.1 Échanges d'intervalle 114
4.1.2 Induction de Rauzy-Veech et classes de Rauzy 115
4.1.3 Induction de Zorich 117
4.1.4 Une nouvelle induction 118
4.1.5 L'opérateur de transfert de $\mathcal{T}_2$ 119
4.2 Lemmes dynamiques de Borel-Cantelli 121
4.2.1 Lemmes dynamiques de Borel-Cantelli pour $\mathcal{T}_2$ 122
4.2.2 Lemmes dynamiques de Borel-Cantelli pour $\mathcal{T}_1$ 122
4.2.3 Lemmes de Borel-Cantelli pour des suites de fonctions quelconques pour $\mathcal{T}_1$ 123
4.2.4 Appendice : apériodicité et mélange faible 128
4.3 Statistiques de récurrence pour l'induction de Rauzy-Veech-Zorich 130
4.3.1 Théorie des valeurs extrêmes 130
4.3.2 Temps d'entrée et de retour 132
4.4 Remarques et questions ouvertes 133

Bibliographie 135

Introduction

Les propriétés statistiques des systèmes dynamiques déterministes autonomes uniformément dilatants sont de nos jours relativement bien comprises, des plus élémentaires, avec l'existence d'une mesure de probabilité invariante et absolument continue [99, 26, 129], la décroissance exponentielle des corrélations et les théorèmes limites [66, 79, 127], jusqu'au plus raffinées, avec les lemmes dynamiques de Borel-Cantelli [41, 95], les versions presque sures des théorèmes limites [38] ou les inégalités de concentration [44, 39]. La démonstration de la plupart de ces résultats s'appuie sur un même dénominateur commun : l'existence d'un trou spectral pour l'opérateur de transfert du système, sur un espace fonctionnel convenablement choisi. Les livres de Boyarsky et Gora [27], de Baladi [19] et de Hennion et Hervé [76] présentent une élégante synthèse de cette approche, ainsi que de nombreuses perspectives historiques. Néanmoins, la situation est comprise de façon bien moins satisfaisante pour les systèmes dynamiques aléatoires ou non autonomes.

Systèmes dynamiques aléatoires

Un système dynamique aléatoire peut être vu comme une composition aléatoire d'applications agissant sur un même espace X, où les applications sont choisies selon un processus stationnaire. Lorsque ce processus consiste en une suite indépendante et identiquement distribuée d'applications, on obtient une chaîne de Markov d'espace d'états X. Plus précisément, si $(\Omega, \mathbb{P})$ est un espace probabilisé, et $T_\omega : X \to X$, $\omega \in \Omega$, est une famille d'applications telle que $(\omega, x) \mapsto T_\omega x$ est mesurable, on peut définir une chaîne de Markov (Z_n) par

$$\begin{cases} Z_0 \sim \mu \\ Z_{n+1} = T_{\omega_{n+1}}(Z_n), \end{cases}$$

où (ω_n) est une suite iid de variables aléatoires à valeurs dans Ω, de loi commune $\mathbb{P}$, et μ est une mesure de probabilité quelconque sur X. Cette dernière sera dite stationnaire si Z_n est distribuée selon μ pour tout n. On s'intéressera aux orbites aléatoires $(T_{\omega_n} \dots T_{\omega_1} x)$, où x est choisi aléatoirement suivant la loi μ, et les ω_k sont choisis aléatoirement et indépendamment suivant la loi $\mathbb{P}$. On notera par la suite $T^n_{\underline{\omega}} = T_{\omega_n} \dots T_{\omega_1}$ pour une suite $\underline{\omega} = (\omega_1, \omega_2, \dots)$. Le cas d'un processus iid a été étudié de façon intensive dans le livre de Kifer [91], tandis que le cas général est abordé dans le livre d'Arnold [8].

L'étude des systèmes dynamiques aléatoires est nécessaire dans la plupart des applications physiques concrètes, où il est très peu vraisemblable qu'une même application soit itérée au cours du temps, et où il est plus réaliste que des applications différentes, mais toutes très proches d'une application donnée, soient composées aléatoirement. Ce sujet des perturbations stochastiques est très bien traité dans le livre de Kifer [92]. Les systèmes dynamiques aléatoires apparaissent aussi naturellement dans l'étude des systèmes de particules sur un réseau, où sur un seul site, la particule est assujettie à une dynamique déterministe locale, mais peut sauter d'un site à un autre aléatoirement, voir [97, 139].

Lorsque le système est contractant en moyenne, l'opérateur de transition de la chaîne de Markov admet un trou spectral sur un espace de fonctions régulières, ce qui permet d'en déduire les théorèmes limites, comme c'est par exemple le cas dans [76, 77]. Inversement, si le système est dilatant en moyenne, comme ce sera le cas pour les systèmes considérés dans cette thèse, l'opérateur de transition est de nature singulière et il devient nécessaire d'étudier les propriétés spectrales de son préadjoint, qui peut être interprété comme la moyenne des opérateurs de transfert des applications T_ω.

En suivant cette approche, l'existence d'une mesure stationnaire absolument continue par rapport à la mesure de Lebesgue a d'abord été examinée par Pelikan [122] et Morita [115] pour des système unidimensionnels, et dans la thèse de Hsieh [81] pour le cas multidimensionnel. La littérature concernant les théorèmes limites est assez vaste. Il doit d'abord être remarqué que pour les systèmes aléatoires, les théorèmes limites se divisent en deux catégories : d'une part, les résultats de type annealed, où l'aléa porte à la fois sur le choix de la condition initiale et le choix des applications itérées, et les résultats de type quenched, où l'aléa porte seulement sur le choix de la condition initiale, et qui sont valides pour presque tout choix des applications itérées.

Les théorèmes limites annealed découlent de l'analyse spectrale de l'opérateur de transfert en moyenne, généralisant ainsi l'approche utilisée pour les systèmes déterministes. Dans cette direction, on peut citer les articles de Baladi [18], Baladi et Young [16], Ishitani [84] ou la thèse de Tümel [139]. Les résultats de type quenched sont généralement plus difficiles à démontrer. La décroissance quenched des corrélations a été étudiée grâce aux cônes de Birkhoff dans [17, 34, 94], tandis qu'un théorème de la limite centrale et une loi du logarithme itéré quenched sont démontrées par Kifer [93] à l'aide d'une approximation par des martingales. Ces résultats concernent des processus stationnaires quelconques, non nécessairement iid, où l'existence d'une mesure stationnaire absolument continue n'est plus assurée. Dans le cadre iid, on peut citer les articles [10, 11] sur les automorphismes aléatoires du tore, ainsi qu'un récent travail de Stenlund et Sulku [135], où une décroissance exponentielle quenched des corrélations et un principe d'invariance presque sure annealed sont démontrés pour des applications aléatoires lisses dilatantes du cercle, par la méthode du couplage.

Dans cette thèse, on établit un cadre fonctionnel abstrait, qui englobe de nombreux systèmes aléatoires dilatants en dimension 1 et supérieure, sous lequel les théorèmes limites annealed sont vérifiés. On résume ces résultats dans le théorème suivant, et on renvoie au chapitre 1 pour les hypothèses exactes sur l'espace de Banach $\mathcal{B}$, la classe d'observables $\mathcal{B}_0$

ainsi que la signification précises des termes "trou spectral" et "apériodicité". On mentionnera seulement que ce théorème s'applique pour des observables à variation bornée, dans le cas d'une famille finie d'applications dilatantes et C^2 par morceaux, qui possède la propriété du recouvrement aléatoire, c'est-à-dire que tout intervalle non trivial recouvrira l'espace tout entier avec probabilité non nulle en un temps fini.

Théorème. *Soit $\{T_\omega\}_{\omega\in\Omega}$ un système dynamique aléatoire. On suppose que l'opérateur de transfert en moyenne du système, défini par rapport à une mesure de référence m, admet un trou spectral sur un espace de Banach $\mathcal{B} \subset L^1(m)$ qui satisfait à certaines hypothèses. Alors :*

1. Mesure stationnaire : *il existe une unique mesure de probabilité μ stationnaire et absolument continue par rapport à m, et sa densité appartient à $\mathcal{B}$;*
2. Décroissance des corrélations : *il existe $0 < \theta < 1$ et $C > 0$ telles que pour tout $f \in \mathcal{B}_0$, tout $g \in L^\infty(m)$ et tout $n \geq 1$:*
$$\left|\int_{\Omega^{\mathbb{N}}}\int_X f(x)g(T^n_{\underline{\omega}}x)d\mu(x)d\mathbb{P}^{\otimes\mathbb{N}}(\underline{\omega}) - \int_X f\,d\mu\int_X g\,d\mu\right| \leq C\theta^n\|f\|_{\mathcal{B}_0}\|g\|_{L^\infty_m},$$
pour une certaine classe d'observables $\mathcal{B}_0$.

Soit $\varphi \in B_0$ borné et à valeurs réelles. On pose $S_n(\underline{\omega}, x) = \sum_{k=0}^{n-1}\varphi(T^k_{\underline{\omega}}x)$. On a :

3. Théorème de la limite centrale : *il existe $\sigma^2 \geq 0$, appelée variance asymptotique de φ, telle que*
$$\frac{S_n - n\int_X \varphi\,d\mu}{\sqrt{n}} \Longrightarrow \mathcal{N}(0,\sigma^2),$$
en loi sous la probabilité $\mathbb{P}^{\otimes\mathbb{N}} \otimes \mu$, et $\sigma^2 = 0$ si et seulement si φ est un cobord L^2 aléatoire, c'est-à-dire si et seulement si il existe $\psi \in L^2(\mu)$ tel que $\varphi = \psi - \psi \circ T_\omega$, μ-presque partout, pour $\mathbb{P}$-presque tout $\omega \in \Omega$;
4. Inégalité de Berry-Esséen : *si $\sigma^2 > 0$, il existe $C > 0$ telle que pour tout $n \geq 1$ et tout $t \in \mathbb{R}$,*
$$\left|\mathbb{P}^{\otimes\mathbb{N}} \otimes \mu\left(\frac{S_n - n\int_X \varphi\,d\mu}{\sqrt{n}} \leq t\right) - \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^t e^{-\frac{u^2}{2\sigma^2}}\,du\right| \leq \frac{C}{\sqrt{n}};$$
5. Principe des grandes déviations : *si $\sigma^2 > 0$, alors il existe $\epsilon_0 > 0$ et une fonction taux $c :]-\epsilon_0, +\epsilon_0[\to [0, +\infty[$ strictement convexe, ne s'annulant qu'en 0, telle que pour tout $0 < \epsilon < \epsilon_0$,*
$$\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}^{\otimes\mathbb{N}} \otimes \mu\left(S_n - n\int_X \varphi\,d\mu > n\epsilon\right) = -c(\epsilon);$$
6. Théorème de la limite locale : *si $\sigma^2 > 0$ et si φ est apériodique, alors pour tout intervalle $I \subset \mathbb{R}$ borné, on a*
$$\lim_{n\to\infty}\sup_{s\in\mathbb{R}}\left|\sigma\sqrt{n}\,\mathbb{P}^{\otimes\mathbb{N}} \otimes \mu\left(s + S_n - n\int_X \varphi\,d\mu \in I\right) - \frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2n\sigma^2}}|I|\right| = 0;$$

7. Principe d'invariance presque sure : *Pour tout $\lambda > \frac{1}{4}$, quitte à élargir l'espace de probabilité, il existe une suite (Y_n) de variables aléatoires iid, suivant la loi $\mathcal{N}(0, \sigma^2)$, telle que*

$$\left| S_n - \sum_{k=0}^{n-1} Y_k \right| = o(n^\lambda), \quad \text{p.s.}$$

Ces théorèmes limites sont pour la plupart démontrés à l'aide de la méthode perturbative de Nagaëv, telle que détaillée dans [76]. Le principe d'invariance presque sure est un renforcement important du théorème de la limite centrale, et a de nombreuses conséquences, parmi lesquelles on peut trouver la loi du logarithme itéré, les versions fonctionnelles et presque sure du théorème de la limite centrale, ... On démontrera celui-ci par la méthode spectrale développée par Gouëzel [64]. Par un argument fondé sur le lemme de Borel-Cantelli, on établira une borne supérieure quenched pour le principe des grandes déviations, mais la question de savoir si le principe des grandes déviations quenched au sens propre est valable reste ouverte.

Changeant notre approche, on adaptera la méthode d'approximation par des martingales, due à Gordin [59], et on donnera une preuve alternative du théorème de la limite centrale annealed. Cela nécessitera l'introduction d'un système déterministe de nature symbolique, pour lequel la procédure standard d'approximation par des martingales pourra être appliquée. L'ingrédient clé de la démonstration ici sera la décroissance annealed des corrélations qui nous permettra de montrer que les sommes de Birkhoff aléatoires peuvent être exprimées comme somme d'une martingale inversée et d'un cobord. Le théorème de la limite centrale découle alors immédiatement du résultat analogue pour les martingales.

On considèrera ensuite les lemmes dynamiques de Borel-Cantelli, ainsi que des inégalités de concentration annealed, qui seront présentés plus loin dans l'introduction.

En important une technique issue du domaine des marches aléatoires en environnement aléatoire, Ayyer, Liverani et Stenlund [11] ont démontré un théorème de la limite centrale quenched pour les automorphismes aléatoires du tore. Leur démarche consiste à établir l'existence d'un trou spectral à la fois pour le système initial, et pour un système dit "doublé" agissant sur X^2 via les applications $\hat{T}_\omega(x, y) = (T_\omega x, T_\omega y)$, et dirigé par le même processus aléatoire donné par $(\Omega, \mathbb{P})$ pour le choix des applications. L'introduction de ce système doublé permet de démontrer un théorème de la limite centrale quenched pour des sous-suites convenables par un argument de type Borel-Cantelli, et le principe de grandes déviations permet d'estimer l'erreur entre deux termes consécutifs de la sous-suite. Malheureusement, cette méthode nécessite une relation précise entre la variance asymptotique de l'observable considéré et celle d'un observable auxiliaire défini sur le système doublé. Cette relation est aisée à démontrer lorsque toutes les applications préservent une mesure commune, mais est plus difficile à établir, et peut-être même fausse, dans le cas général. Dans cette thèse, on montrera que l'argument de [11] peut être appliqué aux systèmes unidimensionnels dilatants, et on donnera une condition concrète nécessaire et suffisante pour démontrer la condition requise pour la variance :

Théorème. *Soit $\{T_\omega\}_{\omega \in \Omega}$ une famille finie d'applications dilatantes et C^2 par morceaux*

sur $[0,1]$*. On suppose que le système possède la propriété de recouvrement, et on note* μ *l'unique mesure stationnaire absolument continue par rapport à Lebesgue. Soit* $\varphi : [0,1] \to \mathbb{R}$ *à variation bornée. Alors* φ *satisfait le théorème de la limite centrale quenched, c'est-à-dire*

$$\frac{\sum_{k=0}^{n-1} \varphi(T_{\underline{\omega}}^k x) - n \int_X \varphi \, d\mu}{\sqrt{n}} \Longrightarrow \mathcal{N}(0, \sigma^2),$$

sous la probabilité μ*, pour* $\mathbb{P}^{\otimes \mathbb{N}}$*-presque tout* $\underline{\omega}$*, où* σ^2 *est la variance asymptotique annealed de* φ*, si et seulement si*

$$\lim_{n\to\infty} \frac{1}{n} \int_{\Omega^{\mathbb{N}}} \left[\sum_{k=0}^{n-1} \left(\int_X \varphi(T_{\underline{\omega}}^k x) d\mu(x) - \int_X \varphi \, d\mu \right) \right]^2 d\mathbb{P}^{\otimes \mathbb{N}}(\underline{\omega}) = 0.$$

On donnera aussi un exemple où les applications préservent toutes la mesure de Lebesgue, mais ne sont pas toutes linéaires par morceaux, et où le théorème de la limite centrale quenched est vérifié.

Systèmes dynamiques non-autonomes

Les systèmes dynamiques non-autonomes, aussi appelés systèmes séquentiels, consistent en une composition de différentes applications d'un espace X dans lui même, et correspondent à des dynamiques où la loi d'évolution change au cours du temps, éventuellement de façon très lente. Si $T_n : X \to X$ est une suite d'applications, on définit l'orbite de x par $(T_n \dots T_1 x)$, et on notera par la suite $T_1^n = T_n \dots T_1$. De tels systèmes ont été étudiés soit d'un point de vue topologique, avec par exemple l'introduction de la notion d'entropie topologique [98] et du chaos de Devaney [40], ou du point de vue de la théorie ergodique : [21] a développé les notions d'ergodicité et de mélange, et [89] a défini l'entropie métrique.

Il est important de remarquer que, en général, il n'existe pas de mesure invariante commune à toutes les applications T_n, bien que l'on dispose souvent d'une mesure de probabilité naturelle m sur l'espace, telle que la mesure de Lebesgue, permettant de définir des processus aléatoires $(\varphi \circ T_1^n)$ à partir d'un observable $\varphi : X \to \mathbb{R}$. Néanmoins, ces processus sont non-stationnaires, du fait que la mesure ne soit pas invariante, et que les applications changent avec le temps, ce qui ajoute des difficultés à leur étude.

Il est à noter que l'étude des propriétés statistiques des systèmes non-autonomes présente de nombreuses similarités avec l'étude des propriétés quenched des systèmes aléatoires, la différence majeure étant qu'un résultat quenched donne des informations pour presque toute suite d'applications, mais ne permet pas de discriminer si une suite particulière vérifie la propriété ou non, tandis qu'un résultat sur les systèmes non-autonomes cible une suite particulière, au prix d'hypothèses renforcées.

Jusqu'à présent, la compréhension des propriétés statistiques des systèmes séquentiels reste très incomplète, et l'accent a été principalement mis sur le phénomène de perte de mémoire, qui est une généralisation de la notion classique de décroissance des corrélations :

si on démarre de deux densités de probabilité, toutes deux appartenant à une même classe de fonctions régulières, celles-ci vont s'attirer mutuellement sous l'effet de la dynamique si le système est suffisamment mélangeant. Ceci conduit à l'idée que l'ensemble des densités lisses sont attirées indifféremment par la même cible mobile dans l'espace des densités, d'où le terme de perte de mémoire.

Une question naturelle est alors d'étudier à quelle vitesse cette perte de mémoire a lieu. De nombreux travaux ont été consacrés à cette question. Parmi eux, on peut citer [121] pour les applications lisses dilatantes, [121] et [46] pour les applications dilatantes par morceaux en dimension 1, [68] pour les applications dilatantes par morceaux en dimension supérieure, [133] pour les difféomorphismes d'Anosov en dimension 2, [134] pour les billards avec diffuseurs convexes et [114] pour les systèmes ouverts en dimension 1.

On remarque que jusqu'à maintenant, ce problème a été examiné uniquement pour des systèmes présentant une dilatation (et éventuellement une contraction) uniforme, pour lesquels la vitesse de perte de mémoire est exponentiellement rapide. Dans cette thèse, on analysera un modèle constitué d'applications intermittentes de l'intervalle $X = [0,1]$: celles-ci sont uniformément dilatantes partout, hormis en 0 où elles ont un point fixe neutre en commun. Plus précisément, pour $0 < \alpha < 1$, on définit $T_\alpha : [0,1] \to [0,1]$ par

$$T_\alpha(x) = \begin{cases} x + \dfrac{3^\alpha}{2^{1+\alpha}} x^{1+\alpha} & \text{si } 0 \le x < \dfrac{2}{3} \\ 3x - 2 & \text{si } \dfrac{2}{3} \le x \le 1. \end{cases}$$

On est alors en mesure de montrer que la perte de mémoire pour la composition d'applications de ce type est polynomiale :

Théorème. *Soit $0 < \alpha < 1$. Alors, pour toute suite (β_n) avec $0 < \beta_n \le \alpha$, on a, en notant $T_1^n = T_{\beta_n} \dots T_{\beta_1}$, pour tout ϕ, ψ de classe C^1 avec $\int_X \phi \, dm = \int_X \psi \, dm$ et tout $\varphi \in L^\infty(m)$:*

$$\left| \int_X \phi(x)\varphi(T_1^n x)\, dm(x) - \int_X \psi(x)\varphi(T_1^n x)\, dm(x) \right| \le C \frac{(\log n)^{\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}-1}} \mathcal{F}(\|\phi\|_{C^1} + \|\psi\|_{C^1}) \|\varphi\|_{L^\infty_m}$$

pour une certaine constante $C > 0$ et une fonction affine $\mathcal{F} : \mathbb{R} \to \mathbb{R}$.

La démonstration de ce résultat est basée sur la technique de perturbation stochastique introduite par Liverani, Saussol et Vaienti [104], et qui consiste à perturber le système aléatoirement afin de "détruire" le point fixe neutre, de telle sorte à retrouver un système qui mélange exponentiellement, et ensuite à estimer l'erreur commise en remplaçant le système initial par le système perturbé.

Au-delà de la perte de mémoire, seul le théorème de la limite centrale a été réellement étudié pour les systèmes séquentiels, avec [14, 15] pour certaines classes d'applications hyperboliques, et [74, 119, 46] pour des applications uniformément dilatantes en dimension 1. Il est

intéressant de remarquer que [46] étudie aussi la loi des grands nombres et les lemmes dynamiques de Borel-Cantelli. On peut aussi citer le récent travail [72] sur le principe d'invariance presque sure.

Malgré cette pauvreté relative de la littérature sur les théorèmes limites pour les systèmes non-autonomes, comme rappelé par Stenlund [133], la majorités des propriétés statistiques des systèmes dynamiques non-stationnaires peuvent être étendues aux systèmes non-stationnaires suffisamment chaotiques. Suivant cette philosophie, on s'intéressera dans cette thèse aux inégalités de concentration pour les systèmes séquentiels.

Inégalités de concentration

Les inégalités de concentration sont une généralisation des inégalités de grandes déviations à des fonctionnelles non additives. Plus précisément, leur intérêt fondamentale est d'estimer la probabilité qu'une variable aléatoire de la forme $Y_n = K(Z_0, \dots, Z_{n-1})$ dévie de son espérance, pour des processus (Z_n) suffisamment indépendants et des fonctions K assez régulières, bien que la dépendance donnée par K puisse être compliquée ou implicite. Il est assez courant de supposer le processus (Z_n) à valeurs dans un espace métrique (X, d) et de considérer des fonctions $K : X^n \to \mathbb{R}$ séparément Lipschitz, c'est-à-dire telles que les quantités

$$\mathrm{Lip}_i(K) = \sup_{x_0,\dots,x_i,\tilde{x}_i,\dots,x_{n-1}} \frac{|K(x_0, \dots, x_i, \dots, x_{n-1}) - K(x_0, \dots, \tilde{x}_i, \dots, x_{n-1})|}{d(x_i, \tilde{x}_i)}$$

soient finies pour tout $i = 0, \dots, n-1$. On dira alors qu'un processus (Z_n) satisfait une inégalité de concentration exponentielle s'il existe $C > 0$ telle que

$$\mathbb{E}(e^{K(Z_0,\dots,Z_{n-1}) - \mathbb{E}(K(Z_0,\dots,Z_{n-1}))}) \leq e^{C \sum_{i=0}^{n-1} \mathrm{Lip}_i^2(K)},$$

pour tout $n \geq 1$ et toute fonction K séparément Lipschitz, ce qui entraine

$$\mathbb{P}(K(Z_0, \dots, Z_{n-1}) - \mathbb{E}(K(Z_0, \dots, Z_{n-1})) > t) \leq e^{-\frac{Ct^2}{\sum_{i=0}^{n-1} \mathrm{Lip}_i^2(K)}},$$

pour tout $t > 0$ et une constante C éventuellement différente. Il est important de noter que ces inégalités sont non-asymptotiques, car valables pour tout $n \geq 1$ fini, contrairement au principe des grandes déviations, mais qu'à la différence de celui-ci, elles ne donnent pas de bornes asymptotiques optimales.

Ces inégalités ont trouvé de multiples applications dans des domaines aussi variés que les statistiques, la géométrie des espaces de grande dimension, les mathématiques discrètes, les graphes aléatoires, l'analyse probabiliste des algorithmes, ... On renvoie à [25, 100, 110, 111] pour une présentation précise et élégante de ce sujet dans un cadre purement probabiliste.

Les inégalités de concentration ont été introduites dans le contexte des systèmes dynamiques par Collet, Martinez et Schmitt [44], qui ont démontré une telle inégalité pour des applications uniformément dilatantes de l'intervalle. Plus exactement, ils ont établi que si

T est une application de classe C^2 par morceaux, uniformément dilatante et topologiquement mélangeante de $[0,1]$, alors le processus (T^n), défini sur l'espace probabilisé $([0,1],\mu)$, où μ est l'unique probabilité T-invariante et absolument continue, satisfait une inégalité de concentration exponentielle, à condition que la densité de μ soit bornée inférieurement loin de 0.

Un nombre important de travaux ont alors suivi [35, 36, 39], visant essentiellement à étendre ce résultat à des systèmes non uniformément dilatants ou hyperboliques, avec éventuellement des inégalités de concentration polynomiale plus faibles. Une description relativement complète de la situation pour les systèmes modélisés par une tour de Young avec une queue des temps de retour exponentielle ou polynomiale est donnée par Chazottes et Gouëzel [39], et complétée par Gouëzel et Melbourne [65].

Dans cette thèse, on démontrera une inégalité de concentration pour la classe de systèmes séquentiels considérée par Conze et Raugi [46]. Ces systèmes consistent en une composition d'applications C^2 par morceaux et uniformément dilatantes de l'intervalle, qui admettent une perte de mémoire exponentielle pour la norme BV. On renvoie aux chapitres 2 et 3 pour l'énoncé précis et la signification des hypothèses.

Théorème. *Soit $(T_n)_n$ un système séquentiel sur $[0,1]$, constitué d'applications de classe C^2 par morceaux, uniformément dilatantes, et qui satisfont certaines conditions techniques non restrictives. On suppose que le système admet une perte de mémoire exponentielle pour la norme BV et que la densité de la mesure de Lebesgue, sous l'effet de la dynamique, reste uniformément bornée loin de 0. Alors le processus (Z_n) défini sur $([0,1],m)$ par $Z_n(x) = T_1^n x$ satisfait une inégalité de concentration exponentielle.*

Pour les systèmes dynamiques aléatoires, il semble n'exister jusqu'à présent aucun résultat sur le phénomène de concentration, hormis le travail de Maldonado [106] portant sur le bruit observationnel. Dans cette thèse, on démontrera une inégalité de concentration annealed pour des systèmes aléatoires dilatants de l'intervalle.

Théorème. *Soit $\{T_\omega\}_{\omega\in\Omega}$ une famille finie d'applications de classe C^2 par morceaux et uniformément dilatantes sur $[0,1]$. On suppose que le système aléatoire possède la propriété de recouvrement, et on note μ l'unique probabilité stationnaire et absolument continue. Alors le processus (Z_n) défini sur $(\Omega^{\mathbb{N}} \times [0,1], \mathbb{P}^{\otimes\mathbb{N}} \otimes \mu)$ par $Z_n(\underline{\omega}, x) = T_{\underline{\omega}}^n x$ satisfait une inégalité de concentration exponentielle.*

La démonstration de ces deux résultats s'appuie sur la méthode des différences bornées de McDiarmid [110, 111], qui consiste à exprimer la variable aléatoire $K(Z_0,\dots,Z_{n-1})$ comme une martingale, et à estimer ses incréments en vue d'appliquer l'inégalité de Azuma-Hoeffding. On décrira aussi des applications de ces deux résultats à la convergence de la mesure empirique, à l'approximation des orbites ainsi qu'à la version presque sure du théorème de la limite centrale.

Lemmes dynamiques de Borel-Cantelli

Soit (X, T, μ) un système dynamique où μ est une probabilité T-invariante et ergodique. Par ergodicité, pour tout ensemble mesurable A avec $\mu(A) > 0$, μ-presque tout $x \in X$ vérifie $T^n x \in A$ infiniment souvent. On peut chercher à généraliser cette propriété à une suite d'ensembles (A_n). On dira qu'une suite d'ensembles mesurables (A_n) avec $\sum_n \mu(A_n) = \infty$ vérifie la propriété de Borel-Cantelli si μ-presque tout $x \in X$ satisfait $T^n x \in A_n$ infiniment souvent. Le nom de cette propriété se justifie car il s'agit de la réciproque du lemme de Borel-Cantelli classique en théorie des probabilités appliqué aux événement $(T^{-n}(A_n))$. Cette question est particulièrement intéressante lorsque A_n est une suite de boules rétrécissant autour d'un point $x_0 \in X$, et partage alors de nombreuses relations avec les notions de récurrence. Une version quantitativement plus précise de cette propriété s'énonce de la façon suivante : on dira que la suite (A_n) vérifie la propriété forte de Borel-Cantelli si

$$\frac{\sum_{k=0}^{n-1} \mathbb{1}_{A_k}(T^k x)}{\sum_{k=0}^{n-1} \mu(A_k)} \to 1,$$

pour μ-presque tout $x \in X$.

Dans le cadre des systèmes dynamiques uniformément hyperboliques, ces questions ont pour la première fois été abordées par Philipp [123], Kleinbock et Margulis [96], Chernov et Kleinbock [41], et par Dolgopyat [50] pour des systèmes partiellement hyperboliques. Plus récemment, des lemmes dynamiques de Borel-Cantelli ont été démontrés pour certains systèmes non-uniformément dilatants, avec par exemple les travaux de Kim [95], Gouëzel [63], Gupta, Nicol et Ott [67] ou Haydn, Nicol, Persson et Vaienti [70]. La clé de toutes ces démonstrations réside dans un contrôle suffisant de la décroissance des corrélations pour les fonctions indicatrices.

Dans cette thèse, on démontrera des lemmes dynamiques de Borel-Cantelli pour des systèmes aléatoires, ce qui n'avait pas été considéré auparavant. On énonce ici un théorème en dimension 1, bien que l'on démontre un résultat général valable en toute dimension :

Théorème. *Soit $\{T_\omega\}_{\omega \in \Omega}$ une famille finie d'applications de classe C^2 par morceaux et uniformément dilatantes sur $[0, 1]$. On suppose que le système aléatoire possède la propriété de recouvrement, et on note μ l'unique probabilité stationnaire et absolument continue. Soit (I_n) une suite d'intervalles telle que $\sum_n \mu(I_n) = \infty$. Alors, pour μ-presque tout $x \in X$ et $\mathbb{P}^{\otimes \mathbb{N}}$-presque tout $\underline{\omega} \in \Omega^{\mathbb{N}}$,*

$$\frac{\sum_{k=0}^{n-1} \mathbb{1}_{I_k}(T_{\underline{\omega}}^k x)}{\sum_{k=0}^{n-1} \mu(I_k)} \to 1.$$

Ce résultat peut être interprété comme une extension de la loi des grands nombres, et il est alors naturel d'étudier la nature des fluctuations dans cette convergence. Il s'avère que dans de nombreux cas, on peut démontrer un théorème de la limite centrale en adaptant la démonstration de [71] et [46] :

Théorème. *On conserve les mêmes hypothèses qu'au théorème précédent, et on suppose de plus que* $(\log n)\mu(I_n) \to 0$. *Alors* $\frac{S_n}{a_n} \Rightarrow \mathcal{N}(0,1)$ *en loi sous la probabilité* $\mathbb{P}^{\otimes\mathbb{N}} \otimes \mu$, *où* $S_n(\underline{\omega}, x) = \sum_{k=0}^{n-1} \left(\mathbb{1}_{I_k}(T_{\underline{\omega}}^k x) - \mu(I_k)\right)$, *et* $a_n^2 \to \infty$ *est la variance de* S_n.

On s'intéressera aussi aux lemmes de Borel-Cantelli pour l'induction de Rauzy-Veech-Zorich. Cette application, qui intervient naturellement dans l'étude des échanges d'intervalles et du flot de Teichmüller, peut être vue comme une généralisation multidimensionnelle de l'algorithme de développement en fractions continues. Ses propriétés statistiques ont fait l'objet d'une attention importante récemment, avec la démonstration de la décroissance des corrélations par Avila et Bufetov [9], du théorème de la limite centrale par Bufetov [32] et du principe d'invariance presque sure par Pollicott [124] pour des observables Hölder.

En utilisant l'existence d'une tour de Young avec une queue des temps de retour exponentielle, on sera en mesure de démontrer le résultat suivant. On renvoie au chapitre 4 pour la définition de $\mathcal{T}_1$ et de la condition de Keane, qui est satisfaite pour presque tout point.

Théorème. *Soit* $\mathcal{T}_1$ *l'induction de Rauzy-Veech-Zorich. Soit* p^* *un point satisfaisant la condition de Keane. Alors toute suite de boules rétrécissant autour de* p^* *vérifie la propriété forte de Borel-Cantelli pour* $\mathcal{T}_1$.

Grâce à la théorie du renouvellement pour les opérateurs, on étudiera aussi directement la décroissance des corrélations pour $\mathcal{T}_1$ pour des fonctions caractéristiques, et on démontrera un résultat semblable au précédent pour des suites non nécessairement décroissantes. On présentera également des résultats sur les lois des valeurs extrêmes et les statistiques de temps d'entrée et de retour pour $\mathcal{T}_1$.

Chapitre 1

Théorèmes limites annealed et quenched pour les systèmes dynamiques aléatoires dilatants

Dans ce chapitre, on présentera des résultats sur les théorèmes limites pour les systèmes aléatoires issus de l'article [3] écrit en collaboration avec Matthew Nicol et Sandro Vaienti, et accepté pour publication dans *Probability Theory and Related Fields*. On définira d'abord la notion de système dynamique aléatoire, avant de poser un cadre fonctionnel abstrait sous lequel on démontrera de nombreux théorèmes limites annealed. On étudiera aussi les lemmes dynamiques de Borel-Cantelli pour les systèmes aléatoires, et on discutera de la validité du théorème de la limite centrale quenched pour les systèmes unidimensionnels.

1.1 Systèmes dynamiques aléatoires : cadre de travail et exemples

1.1.1 Cadre abstrait

Soit $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ un espace de probabilité, et $\theta : \tilde{\Omega} \to \tilde{\Omega}$ une transformation qui préserve la mesure $\tilde{\mathbb{P}}$. Soit $(X, \mathcal{A})$ un espace mesurable. Pour chaque $\underline{\omega} \in \tilde{\Omega}$, on associe une transformation $T_{\underline{\omega}} : X \to X$ telle que l'application $(\underline{\omega}, x) \mapsto T_{\underline{\omega}}(x)$ est mesurable. On s'intéresse alors aux orbites aléatoires $T_{\theta^n \underline{\omega}} \circ \ldots \circ T_{\underline{\omega}} x$. On suppose que X est un espace polonais et que $\mathcal{A}$ est la tribu borélienne associée.
On peut à présent définir la transformation produit-gauche associée au système aléatoire : $F : \tilde{\Omega} \times X \to \tilde{\Omega} \times X$ est défini par $F(\underline{\omega}, x) = (\theta \underline{\omega}, T_{\underline{\omega}} x)$. On dira qu'une mesure de probabilité μ sur $(X, \mathcal{A})$ est une mesure stationnaire si $\tilde{\mathbb{P}} \otimes \mu$ est invariante par F.
La situation la plus courante est celle du cas i.i.d. : $(\tilde{\Omega}, \tilde{\mathcal{T}}, \tilde{\mathbb{P}})$ est alors un espace produit dénombrable, c'est-à-dire $\tilde{\Omega} = \Omega^{\mathbb{N}}$, $\tilde{\mathcal{T}} = \mathcal{T}^{\otimes \mathbb{N}}$, $\tilde{\mathbb{P}} = \mathbb{P}^{\otimes \mathbb{N}}$ et θ est le décalage à gauche. Si à

chaque $\omega \in \Omega$ est associé une application T_ω sur X telle que $(\omega, x) \mapsto T_\omega(x)$ est mesurable, alors on pose $T_{\underline{\omega}} = T_{\omega_1}$ pour chaque $\underline{\omega} \in \tilde{\Omega}$, avec $\underline{\omega} = (\omega_1, \omega_2, \ldots)$. Ce cas particulier rentre bien dans le cadre détaillé précédemment. A partir de maintenant, on ne considèrera plus que le cas i.i.d.
On a la caractérisation suivante des mesures stationnaires :

Lemme 1.1.1. *Une mesure de probabilité μ sur $(X, \mathcal{A})$ est stationnaire si et seulement si $\mu(A) = \int_\Omega \mu(T_\omega^{-1}(A)) d\mathbb{P}(\omega)$ pour tout $A \in \mathcal{A}$.*

Démonstration. L'invariance de $\tilde{\mathbb{P}} \otimes \mu$ par F est équivalente à $\tilde{\mathbb{P}} \otimes \mu(F^{-1}(K \times A)) = \tilde{\mathbb{P}}(K)\mu(A)$ pour tout ensemble $A \in \mathcal{A}$ et $K \in \tilde{\mathcal{T}}$. Or, pour de tels ensembles, on a $F^{-1}(\hat{\Omega} \times A) = \{(\underline{\omega}, x) \,/\, \underline{\omega} \in \theta^{-1}(K), x \in T_{\omega_1}^{-1}(A)\}$, et $\theta^{-1}(K) = \Omega \times K$. On a alors

$$\begin{aligned}
\tilde{\mathbb{P}} \otimes \mu(F^{-1}(K \times A)) &= \int_{\theta^{-1}(K)} \mu(T_{\omega_1}^{-1}(A)) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_\Omega d\mathbb{P}(\omega_1) \int_K d\tilde{\mathbb{P}}(\underline{\omega}) \mu(T_{\omega_1}^{-1}(A)) \\
&= \tilde{\mathbb{P}}(K) \int_\Omega \mu(T_\omega^{-1}(A)) d\mathbb{P}(\omega),
\end{aligned}$$

d'où le lemme. □

De plus, si l'on pose $X_n(\underline{\omega}, x) = T_{\theta^n \underline{\omega}} \circ \ldots \circ T_{\underline{\omega}} x$, cela définit une chaîne de Markov d'espace d'états $(X, \mathcal{A})$ et d'opérateur de transition $U(x, A) = \mathbb{P}(\{\omega \,/\, T_\omega x \in A\})$ pour tout $x \in X$ et $A \in \mathcal{A}$. La probabilité μ est alors stationnaire pour le système aléatoire si et seulement si elle est stationnaire pour la chaîne de Markov (X_n) dans le sens usuel.
On suppose à présent que $(X, \mathcal{A})$ est muni d'une probabilité m telle que chaque transformation T_ω est non-singulière par rapport à m, c'est-à-dire que $m(T^{-1}(A)) = 0$ pour tout $A \in \mathcal{A}$ tel que $m(A) = 0$. On s'intéresse à l'existence et aux propriétés des mesures stationnaires absolument continues par rapport à m. Pour cela, on introduit les opérateurs de transfert et de Koopman moyennés.
Comme chaque transformation T_ω est non-singulière par rapport à m, l'opérateur de transfert P_ω et celui de Koopman U_ω de T_ω sont bien définis, et agissent respectivement sur $L^1(m)$ et $L^\infty(m)$. On rappelle ici brièvement leur définition, et on renvoie vers les livres de Baladi [19] et de Boyarsky et Göra [27] pour de plus amples détails. L'opérateur de Koopman U_ω est défini par $U_\omega f = f \circ T_\omega$ pour tout $f \in L^\infty(m)$. Son action sur les classes d'équivalence modulo m de fonctions est bien définie grâce à la non-singularité. L'opérateur de transfert P_ω, appelé aussi opérateur de Perron-Frobenius, agit sur $L^1(m)$ de la façon suivante : pour $f \in L^1(m)$, soit m_f la mesure complexe vérifiant $dm_f = f dm$. Alors $P_\omega f$ est défini comme étant la dérivée de Radon-Nikodym de $(T_\omega)_\star m_f = m_f \circ T_\omega^{-1}$ par rapport à m. La principale relation entre ces deux opérateurs est donnée par la formule de dualité

$$\int_X P_\omega f(x) g(x) dm(x) = \int_X f(x) U_\omega g(x) dm(x),$$

qui est valable pour tout $f \in L^1(m)$ et $g \in L^\infty(m)$.
On peut maintenant définir les versions moyennées de ces opérateurs : pour $f \in L^1(m)$ et $g \in L^\infty(m)$, on pose $Pf(x) = \int_\Omega P_\omega f(x)\, d\mathbb{P}(\omega)$ et $Ug(x) = \int_\Omega U_\omega g(x)\, d\mathbb{P}(\omega)$. L'opérateur U ainsi défini coïncide avec l'opérateur de transition de la chaîne de Markov (X_n). On remarque que pour tout $n \geq 0$, on a $P^n f(x) = \int_{\tilde{\Omega}} P_{\omega_n} \dots P_{\omega_1} f(x) d\tilde{\mathbb{P}}(\underline{\omega})$ et $U^n g(x) = \int_{\tilde{\Omega}} U_{\omega_1} \dots U_{\omega_n} g(x) d\tilde{\mathbb{P}}(\underline{\omega})$. Par ailleurs, on vérifie aisément que U est l'opérateur dual de P, c'est-à-dire que

$$\int_X Pf(x)g(x)\, dm(x) = \int_X f(x)Ug(x)\, dm(x)$$

pour tout $f \in L^1(m)$ et $g \in L^\infty(m)$. L'une des propriétés fondamentales de l'opérateur P est qu'il permet de caractériser les mesures stationnaires absolument continues : en effet, une probabilité absolument continue est stationnaire si et seulement si sa densité est un point fixe de P.
On va exiger que P admette de bonnes propriétés spectrales sur un certain espace de Banach. Plus précisément, on suppose qu'il existe un espace de Banach $(\mathcal{B}, \|.\|)$ inclus dans $L^1(m)$ tel que :

(H1) l'injection de $\mathcal{B}$ dans $L^1(m)$ est compacte et d'image dense ;

(H2) les fonctions constantes appartiennent à $\mathcal{B}$;

(H3) pour tout $f \in \mathcal{B}$, $|f|$ et $\bar{f}$ appartiennent $\mathcal{B}$;

(H4) $\mathcal{B}$ est invariant par P : $P(\mathcal{B}) \subset \mathcal{B}$, et P agit continument sur $\mathcal{B}$;

(H5) P vérifie une inégalité de Lasota-Yorke : il existe $N \geq 1$, $\rho < 1$ et $K \geq 0$ tels que $\|P^N f\| \leq \rho \|f\| + K\|f\|_{L^1_m}$ pour tout $f \in \mathcal{B}$.

D'après le théorème de Hennion [75] (voir aussi la version antérieure de Ionescu-Tulcea et Marinescu [83]), ces hypothèses entrainent que le rayon spectral essentiel de P vu comme opérateur sur $\mathcal{B}$ est inférieur ou égal à $\rho < 1$. D'autre part, le rayon spectral de P sur $\mathcal{B}$ est égal à 1 : en effet, l'inégalité de Lasota-Yorke implique que le rayon spectral de P est inférieur ou égal à 1, et comme la mesure m, qui appartient au dual topologique de $\mathcal{B}$, est un point fixe de $P^\star$, on en déduit que le rayon spectral de P est plus grand ou égal à 1, d'où l'égalité. On en déduit que l'opérateur P est quasi-compact de type diagonal sur $\mathcal{B}$ (voir par exemple [76]) et admet en particulier la décomposition spectrale suivante :

$$P = \sum_i \lambda_i \Pi_i + Q,$$

où les λ_i sont des valeurs propres de P de module 1, Π_i sont des projecteurs de rang fini, sur les espaces propres associés, Q est un opérateur borné avec un rayon spectral strictement inférieur à 1. Ils satisfont

$$\Pi_i \Pi_j = \delta_{ij} \Pi_i,\ Q\Pi_i = \Pi_i Q = 0.$$

En particulier, P admet un nombre fini de valeurs propres isolées, de multiplicité finies, de module 1, tandis que le reste du spectre est inclus dans un disque de rayon strictement inférieur à 1.

D'après le corollaire III.4 de [76], la suite $h_n = \frac{1}{n}\sum_{k=0}^{n-1} P^k \mathbb{1}$ converge dans $\mathcal{B}$ vers un élément $h \in \mathcal{B}$ qui vérifie $Ph = h$, $h \geq 0$ et $\int h dm = 1$. On en déduit donc l'existence d'une mesure stationnaire absolument continue, dont la densité appartient à $\mathcal{B}$.
Des arguments standards de Schaefer [131] montrent que l'ensemble des valeurs propres de module 1 consiste en un nombre fini de groupes cycliques. On fait une hypothèse interdisant la possibilité d'avoir du spectre périphérique autre que 1 :

(H6) 1 est une valeur propre (isolée) simple de P, et il n'y a pas d'autre valeur propre sur le cercle unité.

Cette hypothèse entraine l'unicité de la mesure stationnaire absolument continue. On la dénotera par μ, et sa densité par h.
Sous ces hypothèses, on peut démontrer que le système $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ est exact. On aura besoin des deux lemmes suivants, dus à Morita [116, Lemme 3.2] :

Lemme 1.1.2. *Soit $\mathcal{L}$ l'opérateur de transfert de F par rapport à la mesure $\tilde{\mathbb{P}} \otimes m$. Pour tout $\phi \in L^1(m)$, on définit $\Phi(\underline{\omega}, x) = \phi(x)$. Alors $\mathcal{L}\Phi(\underline{\omega}, x) = P\phi(x)$, p.p.*

Démonstration. Soit $A \in \mathcal{A}$ et $\Gamma \in \tilde{\mathcal{T}}$. On a

$$\begin{aligned}\int_{\Gamma\times A} \mathcal{L}\Phi \, d\tilde{\mathbb{P}} \otimes m = \int \mathbb{1}_{\Gamma\times A} \circ F \, \Phi \, d\tilde{\mathbb{P}} \otimes m &= \int_X \int_{\tilde{\Omega}} \mathbb{1}_A(T_{\omega_1}x)\mathbb{1}_\Gamma(\sigma\underline{\omega})\phi(x) d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\ &= \tilde{\mathbb{P}}(\Gamma) \int_X \phi(x) U \mathbb{1}_A(x) dm(x) \\ &= \iint_{\Gamma\times A} P\phi(x) d\tilde{\mathbb{P}}(\underline{\omega}) dm(x),\end{aligned}$$

d'où le lemme. □

Lemme 1.1.3. *Soit $\mathcal{L}$ l'opérateur de transfert de F par rapport à la mesure $\tilde{\mathbb{P}} \otimes m$. Pour tout $\phi \in L^1(m)$ et tout $\psi : \Omega^n \to \mathbb{C}$ mesurable et bornée, on définit*

$$\Phi(\underline{\omega}, x) = \phi(x)\psi(\omega_1, \ldots, \omega_n).$$

Alors $\mathcal{L}^n\Phi$ admet une version dans $L^1(m)$: il existe $\chi \in L^1(m)$ tel que $\mathcal{L}^n\Phi(\underline{\omega}, x) = \chi(x)$ p.p.

Démonstration. Soit $A \in \mathcal{A}$ et $\Gamma \in \tilde{\mathcal{T}}$. On a

$$\begin{aligned}\int_{\Gamma\times A} \mathcal{L}^n\Phi \, d\tilde{\mathbb{P}} \otimes m &= \int \mathbb{1}_{\Gamma\times A} \circ F^n \, \Phi \, d\tilde{\mathbb{P}} \otimes m \\ &= \int_X \int_{\tilde{\Omega}} \mathbb{1}_\Gamma(\sigma^n\underline{\omega})\mathbb{1}_A(T^n_{\underline{\omega}}x)\phi(x)\psi(\omega_1, \ldots, \omega_n) d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\ &= \tilde{\mathbb{P}}(\Gamma) \int_X \phi(x) \left(\int_{\tilde{\Omega}} \mathbb{1}_A(T^n_{\underline{\omega}}x)\psi(\omega_1, \ldots, \omega_n) d\tilde{\mathbb{P}}(\underline{\omega}) \right) dm(x) \\ &=: \tilde{\mathbb{P}}(\Gamma)\nu(A).\end{aligned}$$

On vérifie facilement que ν est une mesure complexe finie, et on a

$$\begin{aligned}|\nu(A)| \leq \|\psi\|_{\sup} \int_X \int_{\tilde{\Omega}} \mathbb{1}_A(T^n_{\underline{\omega}} x) d\tilde{\mathbb{P}}(\underline{\omega}) |\phi(x)| dm(x) &= \|\psi\|_{\sup} \int_X U^n \mathbb{1}_A \, |\phi| \, dm \\ &= \|\psi\|_{\sup} \int_A P^n |\phi| \, dm,\end{aligned}$$

ce qui montre que ν est absolument continue par rapport à m. Il existe donc $\chi \in L^1(m)$ tel que

$$\int_{\Gamma \times A} \mathcal{L}^n \Phi \, d\tilde{\mathbb{P}} \otimes m = \int_\Gamma \int_X \chi(x) dm(x) d\tilde{\mathbb{P}}(\underline{\omega}),$$

pour tout $A \in \mathcal{A}$ et tout $\Gamma \in \tilde{\mathcal{T}}$, ce qui prouve le lemme. □

Proposition 1.1.4. *Le système $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ est exact, et donc en particulier ergodique et mélangeant.*

Démonstration. Par les hypothèses (H1) à (H6), on sait que P^n converge pour la norme d'opérateur sur $\mathcal{B}$ vers Π le projecteur sur l'espace engendré par h donné par $\Pi(f) = (\int_X f \, dm) \, h$. Comme $\{P^n - \Pi\}_n$ forme une famille équicontinue d'opérateurs sur $L^1(m)$ qui converge vers 0 sur le sous-espace dense $\mathcal{B}$, on déduit que

$$\left\| P^n f - \left(\int_X f \, dm \right) h \right\|_{L^1_m} \to 0$$

pour tout $f \in L^1(m)$. Soit à présent $\Phi \in L^1(\tilde{\mathbb{P}} \otimes \mu)$ donné par

$$\Phi(\underline{\omega}, x) = \phi(x) \psi(\omega_1, \ldots, \omega_k),$$

où $\phi \in L^1(m)$ et ψ est mesurable bornée. D'après le lemme 1.1.3, il existe $\chi \in L^1(m)$ tel que $\mathcal{L}^k \Phi(\underline{\omega}, x) = \chi(x)$ p.p. Par le lemme 1.1.2, on obtient que $\mathcal{L}^{n+k} \Phi(\underline{\omega}, x) = P^n \chi(x)$ p.p. Comme $\int_X \chi \, dm = \int_{\tilde{\Omega} \times X} \mathcal{L}^k \Phi \, d\tilde{\mathbb{P}} \otimes m = \int_{\tilde{\Omega} \times X} \Phi \, d\tilde{\mathbb{P}} \otimes m$, on trouve que $\mathcal{L}^n \Phi$ converge vers $\left(\int_{\tilde{\Omega} \times X} \Phi \, d\tilde{\mathbb{P}} \otimes m \right) h$ pour la norme $L^1(\tilde{\mathbb{P}} \otimes m)$. Comme l'espace engendré par les Φ considérés est dense dans $L^1(\tilde{\mathbb{P}} \otimes m)$, le même argument d'équicontinuité que précédemment montre que cette convergence a lieu pour n'importe quel $\Phi \in L^1(\tilde{\mathbb{P}} \otimes m)$. En particulier, $\mathcal{L}^n \Phi$ converge vers 0 pour tout Φ d'intégrale nulle par rapport à $\tilde{\mathbb{P}} \otimes m$, ce qui montre que $\tilde{\mathbb{P}} \otimes m$ est exacte pour $(\tilde{\Omega} \times X, F)$, et termine la preuve puisque $\tilde{\mathbb{P}} \otimes \mu \ll \tilde{\mathbb{P}} \otimes m$. □

Dans de nombreux cas, les hypothèses (H4) et (H5) peuvent être déduites des hypothèses correspondantes pour les opérateurs P_ω si les constantes apparaissant dans les inégalités de Lasota-Yorke sont uniformes vis-à-vis de ω. Néanmoins, il est possible de les établir lorsque l'un des opérateurs P_ω ne satisfait pas une inégalité de Lasota-Yorke (généralement lorsque T_ω n'est pas uniformément dilatante), comme on va le voir avec la classe d'exemples suivante.

1.1.2 Applications dilatantes, lisses par morceaux, en dimension 1

On appellera application de Lasota-Yorke une application $T : [0,1] \to [0,1]$ de classe C^2 par morceaux pour laquelle $\lambda(T) := \inf |T'| > 0$. Son opérateur de transfert P_T (par rapport à la mesure de Lebesgue m) vérifie

$$P_T f(x) = \sum_{Ty=x} \frac{f(y)}{|T'(y)|}$$

pour tout $f \in L^1(m)$. On va étudier les propriétés spectrales de P_T sur l'espace des fonctions à variation bornée, dont on rappelle la définition : une fonction $f : [0,1] \to \mathbb{C}$ est dite à variation bornée si sa variation totale, définie par

$$\mathrm{Var}(f) = \sup \sum_{i=0}^{n-1} |f(x_{i+1}) - f(x_i)|,$$

où la borne supérieure est prises sur toutes les suites finies $0 = x_0 < \ldots < x_n = 1$, est finie. Pour une classe d'équivalence $f \in L^1(m)$, on définit

$$\mathrm{Var}(f) = \inf\{\mathrm{Var}(g) \,/\, f = g \ m\text{−p.p.}\}.$$

L'espace $\mathrm{BV} = \{f \in L^1(m) \,/\, \mathrm{Var}(f) < \infty\}$ est muni de la norme $\|f\| = \|f\|_{L^1_m} + \mathrm{Var}(f)$, pour laquelle il est un espace de Banach satisfaisant les hypothèses (H1), (H2) et (H3) ci-dessus. De plus, c'est une algèbre de Banach qui s'injecte continument dans $L^\infty(m)$.
Pour les applications de Lasota-Yorke, on dispose de l'inégalité suivante :

Proposition 1.1.5 (Inégalité de Lasota-Yorke [99])**.** *Pour tout $f \in \mathrm{BV}$, on a*

$$\mathrm{Var}(P_T f) \leq \frac{2}{\lambda(T)} \mathrm{Var}(f) + A(T) \|f\|_{L^1_m},$$

où $A(T)$ est une constante finie ne dépendant que T.

Soient Ω un ensemble fini, $\mathbb{P} = \{p_\omega\}_{\omega \in \Omega}$ un vecteur de probabilité et $T = \{T_\omega\}_{\omega \in \Omega}$ une famille finie d'applications de Lasota-Yorke. On suppose que $p_\omega > 0$ pour tout $\omega \in \Omega$. On appellera par la suite le triplet $(\Omega, \mathbb{P}, T)$ un système de Lasota-Yorke aléatoire.

Définition 1.1.6. *Le système de Lasota-Yorke aléatoire $(\Omega, \mathbb{P}, T)$ est dilatant en moyenne si*

$$\Lambda := \sum_{\omega \in \Omega} \frac{p_\omega}{\lambda(T_\omega)} < 1.$$

L'opérateur de transfert moyenné du système $(\Omega, \mathbb{P}, T)$ est $P = \sum_{\omega \in \Omega} p_\omega P_{T_\omega}$. Il vérifie $P^n = \sum_{\underline{\omega} \in \Omega^n} p^n_{\underline{\omega}} P_{T^n_{\underline{\omega}}}$, où $\underline{\omega} = (\omega_1, \ldots, \omega_n)$, $p^n_{\underline{\omega}} = p_{\omega_1} \ldots p_{\omega_n}$ et $T^n_{\underline{\omega}} = T_{\omega_n} \circ \ldots \circ T_{\omega_1}$.

Proposition 1.1.7. *Si $(\Omega, \mathbb{P}, T)$ est dilatant en moyenne, alors les hypothèses 4 et 5 sont vérifiées.*

Démonstration. D'après la proposition 1.1.5 et la sous-additivité de la variation totale, on a

$$\mathrm{Var}(P^n f) \leq 2\theta_n \mathrm{Var}(f) + A_n \|f\|_{L^1_m}$$

pour tout $n \geq 1$ et tout $f \in \mathrm{BV}$, avec $\theta_n = \sum_{\underline{\omega} \in \Omega^n} \frac{p^n_{\underline{\omega}}}{\lambda(T^n_{\underline{\omega}})}$ et $A_n = \sum_{\underline{\omega} \in \Omega^n} p^n_{\underline{\omega}} A(T^n_{\underline{\omega}})$. Comme $\lambda(T^n_{\underline{\omega}}) \geq \lambda(T_{\omega_1}) \dots \lambda(T_{\omega_n})$, on obtient $\theta_n \leq \sum_{\underline{\omega} \in \Omega^n} \frac{p_{\omega_1} \dots p_{\omega_n}}{\lambda(T_{\omega_1}) \dots \lambda(T_{\omega_n})} = \Lambda^n$. Ainsi, $2\theta_n < 1$ pour n assez grand, tandis que le terme A_n correspondant est fini. Cela termine la preuve. □

Remarque 1.1.8.

1. Ce résultat entraîne en particulier l'existence d'une mesure stationnaire absolument continue dont la densité appartient à BV.
2. Pelikan [122] a montré que la proposition précédente reste valable sous l'hypothèse plus faible $\sup_x \sum_{\omega \in \Omega} \frac{p_\omega}{|T'_\omega(x)|} < 1$.
3. Le résultat reste encore vrai si Ω est infini, sous des hypothèses d'intégrabilité adéquates pour la distorsion, voir la remarque 5.1 de Morita [115]

On donne à présent un critère simple permettant de vérifier l'hypothèse (H6).

Définition 1.1.9. *Le système de Lasota-Yorke aléatoire* $(\Omega, \mathbb{P}, T)$ *possède la propriété de recouvrement si pour tout intervalle non-trivial* $I \subset [0,1]$*, il existe* $n \geq 1$ *et* $\underline{\omega} \in \Omega^n$ *tel que* $T^n_{\underline{\omega}}(I) = [0,1]$ *mod* m*.*

Proposition 1.1.10. *Si* $(\Omega, \mathbb{P}, T)$ *est dilatant en moyenne et possède la propriété de recouvrement, alors l'hypothèse* (H6) *est vérifiée, et de plus la densité de l'unique mesure stationnaire absolument continue est bornée inférieurement par une constante* $c > 0$*.*

Démonstration. Comme le spectre périphérique de P est constitué d'une union finie de groupes cycliques finis, il existe $k \geq 1$ tel que 1 est l'unique valeur propre périphérique de P^k. Il suffit donc de montrer que l'espace propre associé est de dimension 1. Un argument standard montre que l'on peut trouver une base de cet espace propre constitué de fonctions positives à support disjoints. Soit $h \in \mathrm{BV}$ non nulle et satisfaisant $h \geq 0$ et $P^k h = h$. Il existe un intervalle $I \subset [0,1]$ non trivial et $\alpha > 0$ tel que $h \geq \alpha \mathbb{1}_I$. Comme le système possède la propriété de recouvrement, on peut choisir $n \geq 1$ et $\underline{\omega}^\star$ tel que $T^{nk}_{\underline{\omega}^\star}(I) = [0,1]$. Pour tout $x \in [0,1]$, on a :

$$h(x) = P^{nk} h(x) \geq \alpha P^{nk} \mathbb{1}_I(x) = \alpha \sum_{\underline{\omega} \in \Omega^{nk}} p^{nk}_{\underline{\omega}} \sum_{T^{nk}_{\underline{\omega}} y = x} \frac{\mathbb{1}_I(y)}{|(T^{nk}_{\underline{\omega}})'(y)|} \geq \alpha p^{nk}_{\underline{\omega}^\star} \sum_{T^{nk}_{\underline{\omega}^\star} y = x} \frac{\mathbb{1}_I(y)}{|(T^{nk}_{\underline{\omega}^\star})'(y)|}.$$

Cela montre que $h(x) \geq \alpha \frac{p^{nk}_{\underline{\omega}^\star}}{\|(T^{nk}_{\underline{\omega}^\star})'\|_{\sup}} > 0$, puisqu'il existe toujours un $y \in I$ tel que $T^{nk}_{\underline{\omega}^\star} y = x$. On en déduit que h a pour support l'intervalle $[0,1]$ ce qui termine la preuve. □

Certaines propriétés statistiques des systèmes aléatoires unidimensionnels ont été étudiés à l'aide de l'espace BV dans la thèse de Tümel [139].

Exemple 1.1.11. On décrit ici un exemple montrant que l'on peut aller au delà des systèmes dilatants en moyenne. On utilisera pour cela un système aléatoire très simple introduit par Pelikan [122]. On considère les application $T_0(x) = 2x \mod 1$ et $T_1(x) = \frac{x}{2}$, itérées respectivement avec probabilité p et $q = 1-p$. Ce système est dilatant en moyenne si et seulement si $p > \frac{2}{3}$, et comme il possède la propriété de recouvrement, il rentre alors dans le cadre décrit précédemment.
Néanmoins, Pelikan a montré qu'il existe une mesure de probabilité stationnaire absolument continue dès que $p > \frac{1}{2}$, mais que celle-ci n'est pas bornée au voisinage de 0 lorsque $\frac{1}{2} < p < \frac{2}{3}$, interdissant donc l'existence d'un trou spectral sur BV. L'argument clé de Pelikan réside dans l'existence d'un changement de variable adapté, montrant que le système aléatoire est conjugué à un système dilatant en moyenne.
Plus précisément, si on pose $y = \Phi(x) = x^{1-r}$, pour $0 < r < 1$ tel que $1 < 2^{1-r} < \frac{p}{q}$, ce qui est possible puisque $\lim_{r\to 1} 2^{1-r} = 1$ et $\frac{p}{q} > 1$, alors le système aléatoire défini par les applications $\tilde{T}_i = \Phi \circ T_i \circ \Phi^{-1}$, $i = 0, 1$, itérées avec probabilités respectives p et q, est dilatant en moyenne.
En effet, un simple calcul montre que $\lambda(\tilde{T}_0) \geq 2^{1-r}$ et $\lambda(\tilde{T}_1) = 2^{r-1}$, d'où le résultat puisque $\frac{p}{x} + qx < 1$ dès que $1 < x < \frac{p}{q}$.
Si on note $\tilde{P}$ l'opérateur de transfert en moyenne du nouveau système aléatoire, on a d'après l'inégalité de Lasota-Yorke

$$\|\tilde{P}^n f\|_{\mathrm{BV}} \leq C\theta^n \|f\|_{\mathrm{BV}} + D\|f\|_{L^1_m},$$

pour tout $f \in \mathrm{BV}$ et $n \geq 1$, avec $C, D > 0$ et $\theta < 1$.
Comme l'opérateur de transfert P du système aléatoire initial vérifie $\tilde{P}^n = P_\Phi \circ P^n \circ P_\Phi^{-1}$, où P_Φ désigne l'opérateur de transfert de $\Phi : [0,1] \to [0,1]$ par rapport à la mesure de Lebesgue, on est amené à introduire l'espace

$$\mathcal{B} = \left\{ f \in L^1(m) \,:\, P_\Phi f \in \mathrm{BV} \right\},$$

muni de la norme $\|f\|_{\mathcal{B}} = \|P_\Phi f\|_{\mathrm{BV}}$. En utilisant l'expression explicite $P_\Phi f(x) = \frac{x^{\frac{r}{1-r}}}{1-r} f(x^{\frac{1}{1-r}})$, on prouve le lemme élémentaire suivant :

Lemme 1.1.12. *$f \in \mathcal{B}$ si et seulement si $X^r f \in \mathrm{BV}$, où $X(x) = x$ désigne la fonction identité, et la norme $\|.\|_{\mathcal{B}}$ est équivalente à la norme $f \mapsto \|X^r f\|_{\mathrm{BV}}$. De plus, $\mathcal{B}$ s'injecte de façon compacte dans $L^1(m)$.*

Comme $X^r \in \mathrm{BV}$, on en déduit que $\mathrm{BV} \subset \mathcal{B}$, et donc les hypothèses (H1), (H2) et (H3) sont vérifiées. D'autre part, pour tout $f \in \mathcal{B}$ et $n \geq 1$, on a

$$\begin{aligned}
\|P^n f\|_{\mathcal{B}} = \|P_\Phi P^n f\|_{\mathrm{BV}} = \|P_\Phi P_\Phi^{-1} \tilde{P}^n P_\Phi f\|_{\mathrm{BV}} &= \|\tilde{P}^n P_\Phi f\|_{\mathrm{BV}} \\
&\leq C\theta^n \|P_\Phi f\|_{\mathrm{BV}} + D\|P_\Phi f\|_{L^1_m} \\
&\leq C\theta^n \|f\|_{\mathcal{B}} + D\|f\|_{L^1_m}
\end{aligned}$$

et donc les hypothèses (H4) et (H5) sont vérifiées.
Finalement, toute fonction positive et non nulle appartenant à $\mathcal{B}$ est bornée uniformément loin de 0 sur un intervalle, et donc on peut reproduire la démonstration de la proposition 1.1.10 pour démontrer (H6) puisque le système aléatoire possède la propriété de recouvrement.

1.1.3 Applications dilatantes et lisses par morceaux en dimension supérieure

On décrit une classe introduite par Saussol [129] d'applications dilatantes et lisses par morceaux en dimension finie quelconque. Soit m_d la mesure de Lebesgue d-dimensionnelle, $d(.,.)$ la distance euclidienne sur $\mathbb{R}^d$ et γ_d le volume de la boule unité de $\mathbb{R}^d$. Soit M un compact régulier de $\mathbb{R}^d$ et $T : M \to M$ une application T telle qu'il existe une famille finie d'ouverts disjoints $U_i \subset M$ et des ouverts V_i avec $\bar{U}_i \subset V_i$ et des applications $T_i : V_i \to \mathbb{R}^d$ satisfaisant pour un certain $0 < \alpha \leq 1$ et ϵ_0 :

1. $m_d(M \setminus \cup_i U_i) = 0$;
2. pour tout i, les restrictions de T et T_i à U_i coïncident, et $B_{\epsilon_0}(TU_i) \subset T_i(V_i)$,
3. pour tout i, T_i est un C^1-difféomorphisme de V_i sur T_iV_i, et pour tout $x, y \in V_i$ avec $d(T_ix, T_iy) \leq \epsilon_0$, on a
$$|\det DT_i(x) - \det DT_i(y)| \leq c\,|\det DT_i(x)|\, d(T_ix, T_iy)^\alpha,$$
où c est une constante indépendante de i, x et y ;
4. il existe $s < 1$ tel que pour tout $x, y \in V_i$ avec $d(T_ix, T_iy) \leq \epsilon_0$, on a $d(x, y) \leq sd(T_ix, T_iy)$;
5. Supposons que les bords des ouverts U_i sont inclus dans des sous-variétés compactes de classe C^1 par morceaux et de codimension 1. On définit
$$Y = \sup_x \sum_i \sharp\{\text{morceaux lisses intersectant } \partial U_i \text{ et contenant } x\},$$
et
$$\eta_0 = s^\alpha + \frac{4s}{1-s} Y \frac{\gamma_{d-1}}{\gamma_d}.$$
Alors $\eta_0 < 1$.

Les conditions ci-dessus peuvent être affaiblies afin d'autoriser un nombre infini dénombrable de domaines d'injectivité, et aussi la possibilité pour les U_i d'admettre un bord fractal. On renvoie à Saussol [129] pour plus de détails. On appellera de telles applications des applications dilatantes et lisses par morceaux sur M.
L'analyse de l'opérateur de transfert de ces applications requiert l'introduction d'un espace fonctionnel appelé Quasi-Hölder. Cet espace a été introduit et étudié par Keller [90] pour des

applications unidimensionnelles avant d'être utilisé par Saussol [129] en dimension supérieure. Rappelons-en la définition. Soit $f : \mathbb{R}^d \to \mathbb{C}$ une fonction mesurable. Pour tout borélien $A \subset \mathbb{R}^d$, on définit $\mathrm{osc}(f, A) = \operatorname*{ess\,sup}_{x,y\in A} |f(x) - f(y)|$. Pour tout $\epsilon > 0$, la fonction $x \mapsto \mathrm{osc}(f, B_\epsilon(x))$ est positive et s.c.i., donc l'intégrale $\int_{\mathbb{R}^d} \mathrm{osc}(f, B_\epsilon(x))dx$ est bien définie.
Pour tout $f \in L^1(\mathbb{R}^d)$ et $0 < \alpha \leq 1$, on définit

$$|f|_\alpha = \sup_{0<\epsilon\leq\epsilon_0} \frac{1}{\epsilon^\alpha} \int_{\mathbb{R}^d} \mathrm{osc}(f, B_\epsilon(x))dx.$$

Pour un compact régulier $M \subset \mathbb{R}^d$, on pose

$$V_\alpha(M) = \{f \in L^1(\mathbb{R}^d) \,/\, \mathrm{supp}(f) \subset M,\ |f|_\alpha < \infty\},$$

muni de la norme $\|f\|_\alpha = \|f\|_{L^1_m} + |f|_\alpha$, où m est la mesure de Lebesgue normalisée de telle sorte que $m(M) = 1$. Bien que la norme dépende explicitement de ϵ_0, le choix de deux ϵ_0 différents conduira au même espace et à deux normes équivalentes. $V_\alpha(M)$ muni de cette norme devient une algèbre de Banach qui s'injecte continument dans $L^\infty(m)$, de façon compacte dans $L^1(m)$, et vérifie les hypothèses (H1), (H2) et (H3).
Si T est une application dilatante et lisse par morceaux sur M satisfaisant les conditions rappelées ci-dessus, alors son opérateur de transfert P_T est donné par

$$P_f(x) = \sum_{Ty=x} \frac{f(y)}{|\det DT(y)|}.$$

Il vérifie l'inégalité de Lasota-Yorke suivante :

Proposition 1.1.13. [129, Lemme 4.1] *Si ϵ_0 est assez petit, il existe $\eta < 1$ et $D < \infty$ tels que pour tout $f \in V_\alpha$,*

$$|P_T f|_\alpha \leq \eta |f|_\alpha + D\|f\|_{L^1_m}.$$

Soit Ω un ensemble fini, $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ un vecteur de probabilité et $\{T_\omega\}_{\omega\in\Omega}$ une famille finie d'applications dilatantes et lisses par morceaux sur M. On chosit ϵ_0 assez petit de telle sorte que les hypothèses rappelées auparavant soient satisfaites par toutes les applications T_ω pour le même ϵ_0 et tel que $|P_{T_\omega} f|_\alpha \leq \eta_\omega |f|_\alpha + D_\omega \|f\|_{L^1_m}$ pour tout $f \in V_\alpha$, avec $\eta_\omega < 1$ et $D_\omega < \infty$. On pose $\eta = \max \eta_\omega$ et $D = \max D_\omega$. Comme $P = \sum_\omega p_\omega P_{T_\omega}$, on obtient immédiatement que $|Pf|_\alpha \leq \eta |f|_\alpha + \|f\|_{L^1_m}$ pour tout $f \in V_\alpha$, ce qui entraine les hypothèses (H4) et (H5). Pour vérifier l'hypothèse (H6), on peut procéder comme dans le cas unidimensionnel, en introduisant l'hypothèse de recouvrement : en effet, toute fonction positive et non nulle $h \in V_\alpha$ est bornée uniformément loin de 0 sur une boule d'après le lemme 3.1 de [129], ce qui permet d'imiter la preuve de la proposition 1.1.10 et d'obtenir :

Proposition 1.1.14. *Si $(\Omega, \mathbb{P}, T)$ possède la propriété de recouvrement, c'est-à-dire si pour toute boule $B \subset M$, il existe $n \geq 1$ et $\underline{\omega} \in \Omega^n$ tels que $T^n_{\underline{\omega}}(B) = M$, alors l'hypothèse* (H6) *est vérifiée, et de plus la densité de l'unique mesure stationnaire absolument continue est bornée inférieurement par une constante $c > 0$.*

Il existe d'autres espaces fonctionnels susceptibles d'être utilisés pour étudier les applications dilatantes multidimensionnelles. Une des possibilités est l'espace des fonctions à variation bornée, en dimension quelconque. Cet espace a été intensément étudié pour ses applications aux systèmes dynamiques dilatants, voir par exemple [26, 105] parmi bien d'autres. On peut aussi mentionner la thèse de Hsieh [81], consacrée à l'étude des systèmes aléatoires en dimension supérieure, dans le même cadre que nous, bien que son travail se restreint à des applications de type Jabłoński, pour lesquelles la partition de régularité est constituée de rectangles. Néanmoins, l'espace BV en dimension supérieure présente certains inconvénients : il n'est pas inclus dans L^∞, ce n'est pas une algèbre de Banach, et il contient des fonctions positives non nulles qui ne sont bornées inférieurement sur aucune boule, rendant l'utilisation de la notion de recouvrement difficile, à la différence de Quasi-Hölder. Une autre alternative est d'utiliser les espaces de Sobolev fractionnaires, comme cela a été fait par Thomine [137] dans un cadre déterministe.

1.1.4 Applications dilatantes et affines par morceaux

En s'appuyant sur un travail de Tsujii [138], on va considérer le cas d'applications dilatantes et affines par morceaux. Premièrement, on rappelle la définition :

Définition 1.1.15. *Soit U un polyèdre borné de $\mathbb{R}^d$ d'intérieur non vide. Une application dilatante et affine par morceaux est une application $T : U \to U$ pour laquelle il existe une famille $\mathcal{U} = \{U_k\}_{k=1}^l$ de polyèdres $U_k \subset U$, $k = 1, \dots, l$ telle que :*

1. *les intérieurs des polyèdres U_k sont disjoints deux à deux ;*
2. *$\cup_{k=1}^l U_k = U$;*
3. *la restriction de l'application T à chaque U_k est une application affine ;*
4. *il existe $\rho > 1$ tel que $\|DT(x).v\| \geq \rho\|v\|$ pour tout $x \in \cup_{k=1}^l \mathrm{int}(U_k)$ et tout $v \in \mathbb{R}^d$.*

Une conséquence de Tsujii [138] est que l'opérateur de transfert P_T d'une telle application satisfait une inégalité de Lasota-Yorke sur l'espace Quasi-Hölder V_1, c'est-à-dire avec $\alpha = 1$:

Proposition 1.1.16. *Pour toute application dilatante et affine par morceaux $T : U \to U$, il existe $\epsilon_0 > 0$, $\theta < 1$, $C, K > 0$ tels que pour tout $n \geq 0$ et $f \in V_1$:*

$$|P_{\mathcal{T}}^n f|_1 \leq C\theta^n |f|_1 + K\|f\|_{L^1_m}.$$

Soit $T = \{T_\omega\}_{\omega\in\Omega}$ une famille finie d'applications dilatantes et affines par morceaux sur U, et $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ un vecteur de probabilité. Les mêmes arguments que ceux de Tsujii [138] permettent de montrer que le système aléatoire associé satisfait les hypothèses (H1) jusqu'à (H5) pour l'espace Quasi-Hölder, pour un choix adapté de ϵ_0. Si de plus, on suppose que le système possède la propriété de recouvrement, alors l'hypothèse (H6) est elle aussi vérifiée.

1.2 Théorèmes limites par la méthode spectrale

On suppose à partir de maintenant qu'il existe un espace de Banach $\mathcal{B}_0 \subset L^1(m)$, muni de sa norme $\|.\|_0$ tel que pour une certaine constante $C > 0$, on ait $\|fg\| \leq C\|f\|_0\|g\|$ pour tout $f \in \mathcal{B}_0$ et $g \in \mathcal{B}$. Comme $\mathbb{1} \in \mathcal{B}$, on voit que cela implique que $\mathcal{B}_0$ s'injecte continument dans $\mathcal{B}$. Cette hypothèse est clairement satisfaite par $\mathcal{B}_0 = \mathcal{B}$ lorsque $\mathcal{B}$ est une algèbre de Banach, ce qui est le cas lorsque $\mathcal{B}$ est l'espace des fonctions à variation bornée en dimension 1, ou l'espace Quasi-Hölder en dimension quelconque. Si $\mathcal{B}$ est l'espace des fonctions à variation bornée en dimension plus grande que 1, on peut prendre $\mathcal{B}_0 = \mathrm{Lip}$, comme le montre le lemme 6.4 de [137]. Dans le cas de l'espace $\mathcal{B}$ introduit pour l'exemple 1.1.11, on peut prendre $\mathcal{B}_0 = \mathrm{BV}$, puisque $\|fg\|_{\mathcal{B}} \leq C\|X^r fg\|_{\mathrm{BV}} \leq C\|f\|_{\mathrm{BV}}\|X^r g\|_{\mathrm{BV}} \leq C\|f\|_{\mathrm{BV}}\|g\|_{\mathcal{B}}$.
Les éléments de $\mathcal{B}_0$ joueront le rôle des observables dans toute la suite.

1.2.1 Décroissance des corrélations

Les hypothèses (H1) à (H6) faites sur l'opérateur P permettent d'écrire que $P = \Pi + Q$, où Π est le projecteur sur l'espace engendré par la densité h, donné par $\Pi f = (\int_X f dm)\, h$, et Q est un opérateur sur $\mathcal{B}$ de rayon spectral strictement inférieur à 1 et qui satisfait $\Pi Q = Q\Pi = 0$. Il s'ensuit que pour tout $n \geq 0$, on a $P^n = \Pi + Q^n$, avec $\|Q^n\| \leq C\lambda^n$, $C > 0$ et $\lambda < 1$. On en déduit classiquement la décroissance exponentielle des corrélations :

Proposition 1.2.1. *On a :*

1. *Pour tout $f \in \mathcal{B}_0$ et $g \in L^\infty(m)$,*
$$\left|\int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu\right| \leq C\lambda^n \|f\|_0 \|g\|_{L^\infty_m};$$
2. *Si $\mathcal{B}$ s'injecte continument dans $L^\infty(m)$, alors pour tout $f \in \mathcal{B}_0$ et $g \in L^1(m)$,*
$$\left|\int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu\right| \leq C\lambda^n \|f\|_0 \|g\|_{L^1_m},$$
3. *Si $\mathcal{B}$ s'injecte continument dans $L^\infty(m)$ et si de plus la densité de μ est bornée uniformément loin de 0, alors pour tout $f \in \mathcal{B}_0$ et $g \in L^1(\mu)$,*
$$\left|\int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu\right| \leq C\lambda^n \|f\|_0 \|g\|_{L^1_\mu}.$$

Démonstration. La preuve est classique, on adapte par exemple l'appendice C.4 de [7]. On pose $c_n(f,g) = \int_X f\, U^n g\, d\mu - \int_X f\, d\mu \int_X g\, d\mu$. En utilisant les propriétés de l'opérateur de transfert, on a
$$c_n(f,g) = \int_X \left[P^n(fh) - \left(\int_X fh\, dm\right) h\right] g\, dm.$$

Si $f \in \mathcal{B}_0$, on a $fh \in \mathcal{B}$ et donc la décomposition spectrale de P donne

$$\left\| P^n(fh) - \left(\int_X fh\, dm \right) h \right\| = \|P^n(fh) - \Pi(fh)\| \leq C\lambda^n \|fh\| \leq C\|h\|\lambda^n \|f\|_0.$$

Ainsi,

$$\begin{aligned} |c_n(f,g)| &\leq \left\| P^n(fh) - \left(\int_X fh\, dm \right) h \right\|_{L^1_m} \|g\|_{L^\infty_m} \\ &\leq C \left\| P^n(fh) - \left(\int_X fh\, dm \right) h \right\| \|g\|_{L^\infty_m} \\ &\leq C\lambda^n \|f\|_0 \|g\|, \end{aligned}$$

ce qui démontre le premier point.
Pour le point 2, on a

$$\begin{aligned} |c_n(f,g)| &\leq \left\| P^n(fh) - \left(\int_X fh\, dm \right) h \right\|_{L^\infty_m} \|g\|_{L^1_m} \\ &\leq C \left\| P^n(fh) - \left(\int_X fh\, dm \right) h \right\| \|g\|_{L^1_m} \\ &\leq C\lambda^n \|f\|_0 \|g\|, \end{aligned}$$

par l'injection continue de L^∞_m dans $\mathcal{B}$.
Le point 3 se déduit du précédent en remarquant que pour tout $g \in L^1(\mu)$, on a $\|g\|_{L^1_m} \leq (\inf h)^{-1} \|g\|_{L^1_\mu}$. □

1.2.2 Théorème de la limite centrale avec vitesse de convergence

Soit $\varphi \in \mathcal{B}_0$ un observable borné à valeurs réelles. On s'intéresse aux théorèmes limites satisfaits par les sommes de Birkhoff $S_n = \sum_{k=0}^{n-1} X_k$, avec $X_k(\underline{\omega}, x) = \varphi(T_{\omega_k} \dots T_{\omega_1} x)$ défini sur $\tilde{\Omega} \times X$. Pour cela, nous emploierons la méthode perturbative de Nagaëv. On renvoie au livre de Hennion et Hervé [76] pour un exposé détaillé de cette méthode.
On remarque dans un premier temps que S_n correspond à la somme de Birkhoff classique du système produit-gauche déterministe $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ pour l'observable φ (on l'a identifié φ avec $(\underline{\omega}, x) \mapsto \varphi(x)$).
Sous nos hypothèses, le système déterministe $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ est ergodique par la proposition 1.1.4, ainsi une application directe du théorème ergodique de Birkhoff donne la loi des grands nombres, qui est en fait valable pour tout $\varphi \in L^1(\mu)$:

Proposition 1.2.2 (Loi forte des grands nombres)**.** *On a*

$$\frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k_{\underline{\omega}} x) \to \int_X \varphi\, d\mu,$$

pour μ-presque tout $x \in X$ et $\tilde{\mathbb{P}}$-presque tout $\underline{\omega} \in \Omega^{\mathbb{N}}$.

On s'intéresse à présent au théorème de la limite centrale. On supposera sans perdre de généralité que $\int_X \varphi \, d\mu = 0$. La première étape est de démontrer l'existence de la variance asymptotique.

Proposition 1.2.3 (Formule de Green-Kubo)**.** *La limite* $\sigma^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) \geq 0$ *existe et est égale à*

$$\sigma^2 = \int_X \varphi^2 \, d\mu + 2\sum_{n=1}^{+\infty} \int_X \varphi \, U^n \varphi \, d\mu.$$

Démonstration. On développe le carré S_n^2 pour obtenir $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) = \sum_{k,l=0}^{n-1} \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l)$.

Lemme 1.2.4. *Pour tout* $k, l \geq 0$*, on a* $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l) = \int_X \varphi \, U^{|k-l|} \varphi \, d\mu$.

Démonstration du lemme. Par symétrie, on peut supposer $k \geq l$. On a

$$\begin{aligned}
&\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k X_l) = \int_X h(x) \int_{\tilde{\Omega}} X_k(\underline{\omega}, x) X_l(\underline{\omega}, x) \, d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\
&= \int_X h(x) \int_{\tilde{\Omega}} (\varphi \circ T_{\omega_k} \circ \ldots \circ T_{\omega_{l+1}})(T_{\omega_l} \ldots T_{\omega_1} x) \varphi(T_{\omega_l} \ldots T_{\omega_1} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) dm(x) \\
&= \int_X h(x) \int_{\tilde{\Omega}} U^l(\varphi(\varphi \circ T_{\omega_k} \circ \ldots \circ T_{\omega_{l+1}}))(x) \, d\tilde{\mathbb{P}}(\omega_{l+1}, \ldots) dm(x) \\
&= \int_{\tilde{\Omega}} \int_X P^l h(x) \varphi(x) \varphi(T_{\omega_k} \ldots T_{\omega_{l+1}} x) \, dm(x) d\tilde{\mathbb{P}}(\omega_{l+1}, \ldots) = \int_X \varphi(x) U^{k-l} \varphi(x) \, d\mu(x)
\end{aligned}$$

□

En appliquant ce lemme, on obtient

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) = \sum_{k,l=0}^{n-1} \int_X \varphi \, U^{|k-l|} \varphi \, d\mu = n \int_X \varphi^2 \, d\mu + 2\sum_{k=1}^{n} (n-k) \int_X \varphi \, U^k \varphi \, d\mu.$$

Comme $\int_X \varphi \, U^k \varphi \, d\mu$ décroit exponentiellement vite, on en déduit immédiatement que $\frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2)$ converge vers $\int_X \varphi^2 \, d\mu + 2\sum_{n=1}^{+\infty} \int_X \varphi \, U^n \varphi \, d\mu$. □

On dispose du critère suivant, qui caractérise la situation dans laquelle la variance σ^2 est nulle.

Proposition 1.2.5. *La variance asymptotique satisfait* $\sigma^2 = 0$ *si et seulement si il existe* $\psi \in L^2(\mu)$ *tel que pour* $\mathbb{P}$*-presque tout* $\omega \in \Omega$*, on a* $\varphi = \psi - \psi \circ T_\omega$*,* μ*-presque partout.*

Démonstration. On suit la preuve du lemme 4.1 de [11]. Les calculs effectués lors de la preuve de la proposition 1.2.3 montrent que

$$\left| \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) - n\sigma^2 \right| \leq 2\sum_{k=1}^{n} k \left| \int_X \varphi U^k \varphi \, d\mu \right| + 2\sum_{k=n+1}^{\infty} n \left| \int_X \varphi U^k \varphi \, d\mu \right| \leq C.$$

Si $\sigma^2 = 0$, alors cela entraine que S_n est bornée dans $L^2(\tilde{\mathbb{P}} \otimes \mu)$. D'après le théorème de Banach-Alaoglu, on peut en extraire une sous-suite faiblement convergente : il existe $(n_j)_j$ et $Y \in L^2(\tilde{\mathbb{P}} \otimes \mu)$ tel que pour tout $Z \in L^2(\tilde{\mathbb{P}} \otimes \mu)$,

$$\lim_j \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_{n_j} Z) = \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(YZ).$$

Si l'on ne considère que des fonctions Z qui ne dépendent que de $x \in X$, on obtient que pour tout $\phi \in L^2(\mu)$,

$$\lim_j \mathbb{E}_\mu \left(\phi \sum_{k=0}^{n_j-1} U^k \varphi \right) = \mathbb{E}_\mu(\phi\psi),$$

où $\psi = \mathbb{E}_{\tilde{\mathbb{P}}}(Y) \in L^2(\mu)$.
D'autre part, pour tout $\phi \in L^2(\mu)$,

$$\begin{aligned}\mathbb{E}_\mu(\phi(\varphi - \psi + U\psi)) &= \lim_j \mathbb{E}_\mu \left(\phi \left(\varphi - \sum_{k=0}^{n_j-1} U^k \varphi + \sum_{k=0}^{n_j-1} U^{k+1} \varphi \right) \right) \\ &= \lim_j \mathbb{E}_\mu(\phi U^{n_j} \varphi) = 0.\end{aligned}$$

On obtient donc que $\varphi = \psi - U\psi$, μ-presque partout, pour $\psi \in L^2(\mu)$.
On introduit à présent les variables aléatoires $G_n(\underline{\omega}, x) = \psi(T^n_{\underline{\omega}} x)$ et $M_n = \sum_{k=0}^{n-1}(X_k + G_{k+1} - G_0) = G_n - G_0 + S_n$. Soit $\mathcal{E}_k$ la tribu sur $\tilde{\Omega} \times X$ engendrée par les coordonnées $x, \omega_1, \ldots, \omega_k$. Le processus (M_n) est une martingale par rapport à la filtration $\{\mathcal{E}_k\}_k$. En effet, en utilisant le fait que $\tilde{\mathbb{P}} \otimes \mu$ est une mesure produit (dénombrable) sur l'espace $\Omega^{\mathbb{N}} \times X$, on a

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(G_{k+1}|\mathcal{E}_k)(\omega_1, \ldots, \omega_k, x) &= \int_\Omega \psi(T_{\omega_{k+1}} T_{\omega_k} \ldots T_{\omega_1} x) d\mathbb{P}(\omega_{k+1}) \\ &= U\psi(T^k_{\underline{\omega}} x) = \psi(T^k_{\underline{\omega}} x) - \varphi(T^k_{\underline{\omega}} x),\end{aligned}$$

c'est-à-dire que $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(G_{k+1}|\mathcal{E}_k) = G_k - X_k$. On a donc $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(X_k + G_{k+1} - G_k|\mathcal{E}_k) = 0$, ce qui prouve que (M_n) est une martingale.
Comme une martingale dans L^2 a ses accroissements orthogonaux, on obtient par stationnarité

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(M_n^2) = \sum_{k=1}^n \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}((X_k + G_{k+1} - G_k)^2) = n\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}((\varphi_p + \psi_p \circ F - \psi_p)^2),$$

où on a noté φ_p la fonction qui à $(\underline{\omega}, x)$ associe $\varphi(x)$, et similairement pour ψ_p, voir le paragraphe 1.3 pour les notations.
D'autre part, $M_n = G_n - G_0 + S_n$ est bornée dans L^2, comme somme de termes bornés dans L^2, d'où l'on déduit que $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}((\varphi_p + \psi_p \circ F - \psi_p)^2) = 0$, et donc $\varphi(x) = \psi(x) - \psi(T_\omega x)$ pour $\mathbb{P}$-presque tout ω et μ-presque tout x.

□

Soit $\mathcal{M}_{\mathcal{B}}$ l'ensemble des mesures de probabilités sur $(X, \mathcal{A})$ qui sont absolument continues par rapport à m et dont la densité appartient à $\mathcal{B}$. Pour une mesure $\nu \in \mathcal{M}_{\mathcal{B}}$, on dénotera par $\|\nu\|$ la $\mathcal{B}$-norme de sa densité $\frac{d\nu}{dm}$. On peut à présent énoncer le théorème de la limite centrale :

Théorème 1.2.6 (Théorème de la limite centrale)**.** *Pour toute mesure de probabilité $\nu \in \mathcal{M}_{\mathcal{B}}$, le processus $(\frac{S_n}{\sqrt{n}})_n$ converge en loi vers $\mathcal{N}(0, \sigma^2)$, sous la probabilité $\tilde{\mathbb{P}} \otimes \nu$.*

En particulier, le théorème est valide pour la mesure stationnaire μ et la mesure de référence m, avec la même variance asymptotique σ^2.
Afin de démontrer le théorème, on introduit les opérateurs de Laplace. On étudie le cas général d'un paramètre complexe $z \in \mathbb{C}$, plutôt que de se limiter à un paramètre imaginaire pur (et donc aux opérateurs de Fourier), car cela nous sera utile par la suite pour démontrer le principe de grandes déviations. Pour $z \in \mathbb{C}$, on définit P_z par $P_z(f) = P(e^{z\varphi} f)$. Grâce à notre hypothèse sur $\mathcal{B}_0$, P_z est bien défini et est un opérateur borné sur $\mathcal{B}$, et de plus l'application $z \mapsto P_z$ est complexe-analytique sur $\mathbb{C}$. En effet, si on définit $C_n(f) = P(\varphi^n f)$, on a $P_z = \sum_{n \geq 0} \frac{z^n}{n!} C_n$ et cette série est convergente sur $\mathbb{C}$ car $\|C_n\| \leq (C\|\varphi\|_0)^n \|P\|$.
On a la relation fondamentale suivante :

Lemme 1.2.7. *Pour tout $n \geq 0$ et tout $f \in \mathcal{B}$, on a*

$$\int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega}, x)} f(x)\, dm(x)\, d\tilde{\mathbb{P}}(\underline{\omega}) = \int_X P_z^n(f)\, dm.$$

Démonstration. On procède par récurrence sur n. Le cas $n = 0$ est trivial. Supposons maintenant que la relation est vraie pour un certain $n \geq 0$ et tout $f \in \mathcal{B}$. Soit f un élément de $\mathcal{B}$. Comme $P_z(f)$ appartient lui aussi à $\mathcal{B}$, l'hypothèse de récurrence donne

$$\begin{aligned}
\int_X P_z^{n+1}(f)\, dm = \int_X P_z^n(P_z(f))\, dm &= \int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega}, x)} P_z(f)(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_{\tilde{\Omega}} \int_X e^{zS_n(\underline{\omega}, x)} P(e^{z\varphi} f)(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_{\tilde{\Omega}} \int_X U(e^{zS_n(\underline{\omega}, \cdot)})(x) e^{z\varphi(x)} f(x)\, dm(x) d\tilde{\mathbb{P}}(\underline{\omega}) \\
&= \int_X \int_{\tilde{\Omega}} \int_{\Omega} e^{z(\varphi(x) + S_n(\underline{\omega}, T_\omega x))}\, d\mathbb{P}(\omega) d\tilde{\mathbb{P}}(\underline{\omega}) f(x) dm(x)
\end{aligned}$$

Mais $\varphi(x) + S_n(\underline{\omega}, T_\omega x) = S_{n+1}(\omega\underline{\omega}, x)$, où $\omega\underline{\omega}$ désigne la concaténation $(\omega, \omega_1, \omega_2, \ldots)$ si $\underline{\omega} = (\omega_1, \omega_2, \ldots)$. Comme $\int_{\tilde{\Omega}} \int_{\Omega} e^{zS_{n+1}(\omega\underline{\omega}, x)}\, d\mathbb{P}(\omega) d\tilde{\mathbb{P}}(\underline{\omega}) = \int_{\tilde{\Omega}} e^{zS_{n+1}(\underline{\omega}, x)} d\tilde{\mathbb{P}}(\underline{\omega})$ par la structure de mesure produit de $\tilde{\mathbb{P}}$, on obtient la formule voulue pour $n + 1$ et f. □

Si $f \in \mathcal{B}$ est la densité de $\nu \in \mathcal{M}_{\mathcal{B}}$ par rapport à m, alors d'après le lemme précédent, la fonction génératrice des moments de S_n sous la probabilité $\tilde{\mathbb{P}} \otimes \nu$ est donnée par $\int_X P_z^n(f)\, dm$, ce qui nous incite à étudier le comportement asymptotique des itérés des opérateurs de Laplace P_z. Comme $z \mapsto P_z$ est une perturbation analytique de l'opérateur quasi-compact P, on peut appliquer la théorie classique de perturbations des opérateurs, voir par exemple Kato [88] ou le théorème III.8 dans [76], et obtenir le lemme suivant :

Lemme 1.2.8. *Il existe $\epsilon_1 > 0$, $\eta_1 > 0$, $\eta_2 > 0$ et des fonctions complexes-analytiques $\lambda(.)$, $h(.)$, $m(.)$, $Q(.)$, toutes définies sur $\mathbb{D}_{\epsilon_1} = \{z \in \mathbb{C} \,/\, |z| < \epsilon_1\}$, prenant leurs valeurs respectivement dans $\mathbb{C}$, $\mathcal{B}$, $\mathcal{B}^\star$, $\mathcal{L}(\mathcal{B})$ et satisfaisant pour tout $z \in \mathbb{D}_{\epsilon_1}$:*

1. $\lambda(0) = 1, h(0) = h, m(0) = m, Q(0) = Q$;
2. $P_z(f) = \lambda(z)\langle m(z), f\rangle h(z) + Q(z)f$ *pour tout* $f \in \mathcal{B}$;
3. $\langle m(z), h(z)\rangle = 1$;
4. $Q(z)h(z) = 0$ *et* $m(z)Q(z) = 0$;
5. $|\lambda(z)| > 1 - \eta_1$;
6. $\|Q(z)^n\| \leq C(1 - \eta_1 - \eta_2)^n$.

De plus, $|\langle m, Q(z)^n f\rangle| \leq C|z|(1 - \eta_1 - \eta_2)^n \|f\|$ *pour tout* $f \in \mathcal{B}$ *et* $z \in \mathbb{D}_{\epsilon_1}$.

Pour $n \geq 0$, on a $P_z^n(f) = \lambda(z)^n \langle m(z), f\rangle h(z) + Q(z)^n f$. Ainsi, le comportement asymptotique de P_z^n est clairement lié à celui de la valeur propre dominante $\lambda(z)$ dans un voisinage de 0.

Lemme 1.2.9. *La valeur propre dominante $\lambda(.)$ satisfait $\lambda'(0) = \int \varphi d\mu = 0$ et $\lambda''(0) = \sigma^2 \geq 0$.*

Démonstration. D'après le corollaire III.11 de [76], $\lambda'(0) = \langle m(0), P'(0)h(0)\rangle$. Comme $m(0) = m$, $h(0) = h = \frac{d\mu}{dm}$ et $P'(0)f = C_1(f) = P(\varphi f)$ pour tout $f \in \mathcal{B}$, puisque $P(z) = \sum_{n\geq 0} \frac{z^n}{n!} C_n$ avec $C_n(f) = P(\varphi^n f)$, la formule pour $\lambda'(0)$ devient

$$\lambda'(0) = \langle m, P(\varphi h)\rangle = \langle m, \varphi h\rangle = \int \varphi d\mu = 0.$$

En utilisant à nouveau le corollaire III.11 de [76], on a

$$\lambda''(0) = \langle m(0), P''(0)h(0)\rangle + 2\langle m(0), P'(0)\tilde{h}\rangle,$$

où $\tilde{h}$ est l'unique élément de $\mathcal{B}$ satisfaisant $\langle m(0), \tilde{h}\rangle = 0$ et $(\lambda(0) - P(0))\,\tilde{h} = (P'(0) - \lambda'(0))\,h(0)$. Cela implique que $\tilde{h}$ est l'unique élément de $\mathcal{B}$ tel que $\langle m, \tilde{h}\rangle = 0$ et $(I - P)\tilde{h} = P(\varphi h)$. D'après le corollaire III.6 de [76], $\tilde{h}$ est donné par $\tilde{h} = \sum_{n\geq 0} Q(0)^n(\varphi h)$. Mais $Q(0)^n(\varphi h) = P^n(\varphi h) - \langle m, \varphi h\rangle h = P^n(\varphi h)$, puisque $\langle m, \varphi h\rangle = \int \varphi d\mu = 0$. Ainsi $\tilde{h} = \sum_{n\geq 0} P^n(P(\varphi h)) = \sum_{n\geq 1} P^n(\varphi h)$.
D'une part, on a $\langle m(0), P''(0)h(0)\rangle = \langle m, C_2(h)\rangle = \langle m, P(\varphi^2 h)\rangle = \int \varphi^2 d\mu$. D'autre part,

$$\begin{aligned}\langle m(0), P'(0)\tilde{h}\rangle = \langle m, P(\varphi\tilde{h})\rangle = \langle m, \varphi\tilde{h}\rangle &= \sum_{n\geq 1} \langle m, \varphi P^n(\varphi h)\rangle \\ &= \sum_{n\geq 1} \int \varphi P^n(\varphi h) dm = \sum_{n\geq 1} \int U^n\varphi\, \varphi d\mu.\end{aligned}$$

En sommant ces deux relations, on retrouve la formule de Green-Kubo pour σ^2. □

Preuve du théorème de la limite centrale. D'après le lemme précédent, $\lambda(\frac{it}{\sqrt{n}})^n = (1 - \frac{\sigma^2 t^2}{2n} + o(\frac{1}{n}))^n$ converge vers $e^{-\frac{\sigma^2 t^2}{2}}$, d'où il s'ensuit que

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) = \lambda(\frac{it}{\sqrt{n}})^n \langle m(\frac{it}{\sqrt{n}}), f\rangle \langle m, h(\frac{it}{\sqrt{n}})\rangle + \langle m, Q(\frac{it}{\sqrt{n}})f\rangle$$

converge lui aussi vers $e^{-\frac{\sigma^2 t^2}{2}}$, ce qui conclut la preuve, grâce au théorème de Paul Lévy. □

On peut préciser ce résultat en donnant une vitesse de convergence lorsque $\sigma^2 > 0$. On prouve tout d'abord le lemme suivant, qui donne la vitesse de convergence des fonctions caractéristiques :

Lemme 1.2.10. *Il existe $C > 0$ et $\rho < 1$ tel que pour tout $t \in \mathbb{R}$ et $n \geq 0$ avec $\frac{|t|}{\sqrt{n}}$ suffisamment petit, et tout $\nu \in \mathcal{M}_{\mathcal{B}}$, on a*

$$\left|\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{1}{2}\sigma^2 t^2}\right| \leq C\|\nu\| \left(e^{-\frac{\sigma^2 t^2}{2}} \left(\frac{|t| + |t|^3}{\sqrt{n}}\right) + \frac{|t|}{\sqrt{n}}\rho^n\right).$$

Démonstration. Cela découle essentiellement de la dérivabilité à l'ordre 3 de $\lambda(.)$ en 0 : pour $\frac{|t|}{\sqrt{n}}$ suffisamment petit, $\lambda(\frac{it}{\sqrt{n}})^n = \left(1 - \frac{\sigma^2 t^2}{2n} + \mathcal{O}\left(\frac{|t|^3}{n\sqrt{n}}\right)\right)^n = e^{-\frac{\sigma^2 t^2}{2}} + \mathcal{O}\left(e^{-\frac{\sigma^2 t^2}{2}}\frac{|t|^3}{\sqrt{n}}\right)$. En se rappelant que $f = \frac{d\nu}{dm}$ et $\|\nu\| = \|f\| \geq C$, où la constante C provient de l'injection continue $\mathcal{B} \subset L^1(m)$ et est indépendante de ν, on a

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) &= \lambda(\frac{it}{\sqrt{n}})^n \langle m(\frac{it}{\sqrt{n}}), f\rangle \langle m, h(\frac{it}{\sqrt{n}})\rangle + \langle m, Q(\frac{it}{\sqrt{n}})^n f\rangle \\
&= \left(e^{-\frac{\sigma^2 t^2}{2}} + \mathcal{O}\left(e^{-\frac{\sigma^2 t^2}{2}}\frac{|t|^3}{\sqrt{n}}\right)\right)\left(1 + \mathcal{O}\left(\frac{|t|}{\sqrt{n}}\|f\|\right)\right) + \mathcal{O}\left(\frac{|t|}{\sqrt{n}}\rho^n\|f\|\right),
\end{aligned}$$

ce qui démontre le lemme. □

En particulier, on a $\left|\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{1}{2}\sigma^2 t^2}\right| = \mathcal{O}\left(\frac{1+|t|^3}{\sqrt{n}}\right)$, ce qui nous sera à nouveau utile lorsque nous étudierons la version quenched du théorème de la limite centrale.

Théorème 1.2.11 (Théorème de la limite centrale avec vitesse de convergence)**.** *Si $\sigma^2 > 0$, il existe $C > 0$ telle que pour tout $\nu \in \mathcal{M}_{\mathcal{B}}$ et tout $n \geq 1$*

$$\sup_{t\in\mathbb{R}} \left|\tilde{\mathbb{P}} \otimes \nu\left(\frac{S_n}{\sqrt{n}} \leq t\right) - \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{t} e^{-\frac{u^2}{2\sigma^2}}du\right| \leq \frac{C\|\nu\|}{\sqrt{n}}.$$

Démonstration. On utilise l'inégalité classique de Berry-Esséen, dont l'énoncé donné ci-dessous peut être trouvé au lemme VI.3 de [76].

Lemme 1.2.12 (Inégalité de Berry-Esséen)**.** *Soient p_1 et p_2 deux mesures de probabilité sur $\mathbb{R}$, de fonctions de répartitions respectives F_1 et F_2, et de transformées de Fourier respectives $\hat{p}_1$ et $\hat{p}_2$. On suppose que*

$$\int |t|\, dp_1(t) < \infty \text{ et } \int |t|\, dp_2(t) < \infty,$$

et que F_2 possède une dérivée telle que $\sup_{t\in\mathbb{R}} F_2'(t) = a < \infty$.
Alors, pour tout $L > 0$, on a

$$\sup_{t\in\mathbb{R}} |F_1(t) - F_2(t)| \leq \frac{1}{\pi}\int_{-L}^{+L} \frac{|\hat{p}_1(t) - \hat{p}_2(t)|}{|t|} dt + \frac{24a}{\pi L}.$$

Soit $\delta > 0$ tel que le lemme 1.2.10 soit valable pour tout t et n tel que $\frac{|t|}{\sqrt{n}} \leq \delta$. On applique le lemme de Berry-Esséen avec $L = \delta\sqrt{n}$, p_1 la loi de $\frac{S_n}{\sqrt{n}}$ et $p_2 = \mathcal{N}(0, \sigma^2)$.
D'après le lemme 1.2.10 et l'inégalité de Berry-Esséen, on a

$$\begin{aligned}
\sup_{t\in\mathbb{R}} |F_1(t) - F_2(t)| &\leq C\int_{-\delta\sqrt{n}}^{+\delta\sqrt{n}} \|\nu\| \left[e^{-\frac{\sigma^2 t^2}{2}}\left(\frac{|t| + |t|^3}{\sqrt{n}}\right) + \frac{|t|}{\sqrt{n}}\rho^n\right]\frac{dt}{|t|} + \frac{C}{\sqrt{n}} \\
&\leq \frac{C\|\nu\|}{\sqrt{n}}\left(1 + \int_{-\delta\sqrt{n}}^{+\delta\sqrt{n}} \left[e^{-\frac{\sigma^2 t^2}{2}}(1 + |t|^2) + \rho^n\right] dt\right) \\
&\leq \frac{C\|\nu\|}{\sqrt{n}}\left(1 + 2\delta\sqrt{n}\rho^n + \int_{\mathbb{R}} e^{-\frac{\sigma^2 t^2}{2}}(1 + |t|^2) dt\right) \leq \frac{C\|\nu\|}{\sqrt{n}}.
\end{aligned}$$

□

On termine cette section par un lemme qui montre que la variance asymptotique σ^2 peut être calculée à partir de n'importe quelle loi initiale $\nu \in \mathcal{M}_{\mathcal{B}}$.

Lemme 1.2.13. *Pour tout $\nu \in \mathcal{M}_{\mathcal{B}}$, on a $\sigma^2 = \lim_{n\to\infty} \frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n^2)$.*

Démonstration. Pour $n \geq 1$ fixé, on écrit le développement en puissance de t au voisinage de 0 de $\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\frac{it}{\sqrt{n}}S_n})$. On a d'une part

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\frac{it}{\sqrt{n}}S_n}) = 1 + \frac{t}{\sqrt{n}}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n) - \frac{t^2}{2n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n^2) + \mathcal{O}(t^3),$$

et d'autre part

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\frac{it}{\sqrt{n}}S_n}) &= \lambda(\frac{it}{\sqrt{n}})^n\langle m(\frac{it}{\sqrt{n}}), f\rangle\langle m, h(\frac{it}{\sqrt{n}})\rangle + \langle m, Q(\frac{it}{\sqrt{n}})^n f\rangle \\
&= \left(1 - \frac{\sigma^2}{2}t^2 + \mathcal{O}_n(t^3)\right)\left(1 + \frac{ic_1}{\sqrt{n}}t - \frac{c_2}{2n}t^2 + \mathcal{O}_n(t^3)\right) + \left(\frac{ie_{1,n}}{\sqrt{n}}t - \frac{e_{2,n}}{2n}t^2 + \mathcal{O}_n(t^3)\right),
\end{aligned}$$

où l'indice n indique que les constantes impliquées dans les $\mathcal{O}$ dépendent de n (mais pas de t).
En identifiant les termes d'ordre t^2, on obtient

$$-\frac{1}{2n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n^2) = -\frac{\sigma^2}{2} - \frac{c_2}{2n} - \frac{e_{2,n}}{2n},$$

d'où

$$\frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n^2) = \sigma^2 + \frac{c_2}{n} + \frac{e_{2,n}}{n}.$$

On estime à présent $e_{2,n}$. On a $e_{2,n} = \langle m, \frac{d^2}{dz^2}(Q(z)^n)|_{z=0} f\rangle$. Or, d'après le point (iv) du Théorème III.8 de [76], on a $\left\|\frac{d^2}{dz^2}(Q(z)^n)|_{z=0}\right\| \leq C\lambda^n$ pour $C > 0$ et $0 < \lambda < 1$, d'où l'on tire $|e_{2,n}| \leq C\lambda^n$ et

$$\frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(S_n^2) \to \sigma^2.$$

□

1.2.3 Principe des grandes déviations

Soit à nouveau $\varphi \in \mathcal{B}$ un observable borné à valeurs réelles, avec $\int_X \varphi\, d\mu = 0$, et définissons $X_n(\underline{\omega}, x) = \varphi(T_{\underline{\omega}}x)$ et $S_n = \sum_{k=0}^{n-1} X_k$. La loi des grands nombres (Proposition 1.2.2) entraine que $\frac{S_n}{n}$ converge vers 0 en probabilité : pour tout $\epsilon > 0$, $\tilde{\mathbb{P}} \otimes \mu(S_n > n\epsilon)$ converge vers 0. Il se trouve que cette convergence vers 0 a lieu avec une vitesse exponentielle, et le principe des grandes déviations va nous permettre de préciser le taux exact de cette convergence exponentielle.

Théorème 1.2.14 (Principe des grandes déviations)**.** *Supposons que $\sigma^2 > 0$. Alors, il existe $\epsilon_0 > 0$ et une fonction taux $c : (-\epsilon_0, +\epsilon_0) \to [0, +\infty[$ strictement convexe, réelle-analytique, ne s'annulant qu'en 0 telle que pour tout $\nu \in \mathcal{M}_{\mathcal{B}}$ et tout $0 < \epsilon < \epsilon_0$, on a*

$$\lim_{n\to\infty} \frac{1}{n} \log \tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) = -c(\epsilon)$$

et

$$\lim_{n\to\infty} \frac{1}{n} \log \tilde{\mathbb{P}} \otimes \nu(S_n < -n\epsilon) = -c(-\epsilon).$$

Comme pour le théorème de la limite centrale, on remarque que le principe des grandes déviations est valable pour une large variété de mesures initiales, avec la même fonction taux c. Afin de démontrer le théorème, on va montrer la convergence de $\frac{1}{n}\log \mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n})$ pour $\theta \in \mathbb{R}$ assez petit et appliquer ensuite le théorème de Gartner-Ellis [48, 53]. Pour cela, on exploite l'étude des opérateurs P_z réalisée au paragraphe précédent.

Lemme 1.2.15. *Il existe $0 < \epsilon_2 < \epsilon_1$ tel que pour tout $\theta \in \mathbb{R}$, $|\theta| < \epsilon_2$, on a $\lambda(\theta) > 0$. De plus, les fonctions $h(.)$ et $m(.)$ peuvent être redéfinies de telle sorte qu'elles vérifient toujours les conclusions du lemme 1.2.8 mais aussi $h(\theta) \geq 0$, $m(\theta) \geq 0$ pour $\theta \in \mathbb{R}$, $|\theta| < \epsilon_2$.*

Démonstration. Comme P_θ est un opérateur réel, on a $P_\theta \overline{f} = \overline{P_\theta f}$ pour tout $f \in \mathcal{B}$. Ainsi, $P_\theta \overline{h(\theta)} = \overline{P_\theta h(\theta)} = \overline{\lambda(\theta)}\,\overline{h(\theta)}$. Comme $\lambda(\theta)$ est l'unique valeur propre de module maximal de P_θ, on obtient $\lambda(\theta) = \overline{\lambda(\theta)}$, et donc $\lambda(\theta) \in \mathbb{R}$. Comme $\lambda(0) = 1$, par un argument de continuité, on a $\lambda(\theta) > 0$ pour $\theta \in \mathbb{R}$ assez petit. Pour $z \in \mathbb{C}$ assez petit, $\langle m(z), \mathbb{1}\rangle \neq 0$. On définit $\tilde{h}(z) = \langle m(z), \mathbb{1}\rangle h(z)$ et $\tilde{m}(z) = \langle m(z), \mathbb{1}\rangle^{-1} m(z)$. Ces nouveaux éléments propres satisfont clairement les conclusions du lemme 1.2.8. Il suffit donc de prouver que $\tilde{h}(\theta)$ et $\tilde{m}(\theta)$ sont positifs pour $\theta \in \mathbb{R}$ assez petit. D'après la décomposition spectrale de P_θ, on observe que $\lambda(\theta)^{-n} P_\theta^n \mathbb{1}$ converge vers $\tilde{h}(\theta)$ dans $\mathcal{B}$, et donc dans $L^1(m)$. Cela entraine que $\tilde{h}(\theta) \geq 0$ puisque P_θ est un opérateur positif, et $\lambda(\theta)$ est strictement positif. Soit maintenant $\psi(\theta) \in \mathcal{B}^\star$ positif tel que $\langle \psi(\theta), \tilde{h}(\theta)\rangle = 1$ [1]. Alors $\lambda(\theta)^{-n}(P_\theta^\star)^n \psi(\theta)$ converge vers $\langle \psi(\theta), h(\theta)\rangle m(\theta) = \tilde{m}(\theta)$, ce qui prouve que $\tilde{m}(\theta)$ est une forme linéaire positive. □

On pose $\Lambda(\theta) = \log \lambda(\theta)$. On a alors

Proposition 1.2.16. *Il existe $0 < \epsilon_3 < \epsilon_2$ tel que pour tout $\nu \in \mathcal{M}_\mathcal{B}$ et tout $\theta \in \mathbb{R}$ avec $|\theta| < \epsilon_3$, on a*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n}) = \Lambda(\theta).$$

Démonstration. Soit $f \in \mathcal{B}$ la densité $\frac{d\nu}{dm}$. On a l'identité

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n}) = \langle m, P_\theta^n(f)\rangle &= \lambda(\theta)^n \langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \langle m, Q(\theta)^n f\rangle \\ &= \lambda(\theta)^n \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right)\end{aligned}$$

Comme toutes les quantités impliquées sont positives, on peut écrire

$$\frac{1}{n} \log \mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n}) = \log \lambda(\theta) + \frac{1}{n} \log \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right).$$

Pour θ réel assez petit, on a $\frac{1}{2} \leq \langle m(\theta), f\rangle \langle m, h(\theta)\rangle \leq \frac{3}{2}$, tandis que

$$\left|\lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right| \leq C \left(\frac{1-\eta_1-\eta_2}{1-\eta_1}\right)^n$$

d'après le lemme 1.2.8, ce qui entraine que

$$\lim_{n\to\infty} \frac{1}{n} \log \left(\langle m(\theta), f\rangle \langle m, h(\theta)\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^n f\rangle\right) = 0.$$

□

1. Il suffit de poser $\psi(\theta) = \alpha(\theta)^{-1} m$, où $\alpha(\theta) = \langle m, \tilde{h}(\theta)\rangle$ est strictement positif puisque $\tilde{h}(\theta)$ est un élément positif non nul de $L^1(m)$.

Démonstration du principe des grandes déviations. On applique une version locale du théorème de Gartner-Ellis, dont l'énoncé précis peut-être trouvé dans le lemme XIII.2 de [76]. D'après celui-ci et la proposition précédente, il suffit de prouver que $\Lambda(.)$ est différentiable et strictement convexe dans un voisinage de 0. C'est clairement le cas car $\lambda(.)$ est complexe-analytique et on a supposé $\Lambda''(0) = \lambda''(0) = \sigma^2 > 0$. D'après la proposition VIII.3 de [76], $c(\epsilon) = \epsilon\theta_\epsilon - \Lambda(\theta_\epsilon)$, où θ_ϵ est l'unique solution de l'équation $\Lambda'(\theta) = \epsilon$. On en déduit la réelle-analyticité de c en employant le théorème d'inversion locale, puisque $\Lambda(.)$ est réelle-analytique. □

On aura besoin plus tard du résultat suivant, qui précise la fonction taux c pour des valeurs petites de ϵ et permet d'obtenir une borne supérieure explicite pour la probabilité des grandes déviations. On s'inspire du lemme 4.3 de [11].

Lemme 1.2.17. *Il existe $C > 0$ tel que pour $\epsilon > 0$ assez petit,*

$$\tilde{\mathbb{P}} \otimes \nu(|S_n| > n\epsilon) \leq Ce^{-C\epsilon^2 n}.$$

Démonstration. On ne traite que le cas de $\tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon)$, le terme $\tilde{\mathbb{P}} \otimes \nu(S_n < -n\epsilon)$ se traitant de façon similaire.
On a pour tout $|\theta| < \epsilon_3$,

$$\begin{aligned}
\tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) &\leq e^{-n\theta\epsilon}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(e^{\theta S_n}) = e^{-n\theta\epsilon}\langle m, P_\theta^n f\rangle \\
&= e^{-n(\theta\epsilon-\Lambda(\theta))}\left[\langle m, h(\theta)\rangle\langle m(\theta), f\rangle + \lambda(\theta)^{-n}\langle m, Q(\theta)^{-n} f\rangle\right] \\
&= \mathcal{O}(1)e^{-n(\theta\epsilon-\Lambda(\theta))} \\
&\leq \mathcal{O}(1)e^{-nc(\epsilon)},
\end{aligned}$$

puisque $c(\epsilon) = \sup_{|\theta|<\epsilon_3}\{\theta\epsilon - \Lambda(\theta)\}$ par le lemme XIII.2 de [76].
Pour estimer inférieurement $c(\epsilon)$, on développe $\theta\epsilon - \Lambda(\theta) = \theta\epsilon - \frac{\sigma^2}{2}\theta^2 + \mathcal{O}(\theta^3)$. Une optimisation sur θ suggère de prendre $\theta = \frac{\epsilon}{\sigma^2}$, ce qui est toujours possible pour ϵ assez petit, et on obtient

$$c(\epsilon) \geq \frac{\epsilon^2}{2\sigma^2} + \mathcal{O}(\epsilon^3).$$

□

On peut déduire de cette version annealed du principe des grandes déviations une borne supérieure quenched, c'est-à-dire valable pour $\tilde{\mathbb{P}}$-presque toute réalisation $\underline{\omega}$, avec la même fonction taux c.

Proposition 1.2.18. *Pour tout $0 < \epsilon < \epsilon_0$, tout $\nu \in \mathcal{M}_\mathcal{B}$ et $\tilde{\mathbb{P}}$-presque tout $\underline{\omega} \in \Omega^{\mathbb{N}}$, on a*

$$\limsup_{n\to\infty} \frac{1}{n}\log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -c(\epsilon).$$

Démonstration. Soit $0 < \gamma < 1$. On pose

$$A_n = \{\underline{\omega} \in \tilde{\Omega} \,/\, \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \geq e^{-n(1-\gamma)c(\epsilon)}\}.$$

On a vu lors de la preuve du lemme 1.2.17 que $\tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) \leq Ce^{-nc(\epsilon)}$ pour $C > 0$. L'inégalité de Markov donne alors

$$\tilde{\mathbb{P}}(A_n) \leq e^{n(1-\gamma)c(\epsilon)} \tilde{\mathbb{P}} \otimes \nu(S_n > n\epsilon) \leq Ce^{-n\gamma c(\epsilon)}.$$

Par le lemme de Borel-Cantelli, on obtient que $\tilde{\mathbb{P}}$-presque tout $\underline{\omega}$ appartient à un nombre fini de A_n d'où $\limsup_{n\to\infty} \frac{1}{n} \log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -(1-\gamma)c(\epsilon)$ pour $\tilde{\mathbb{P}}$-presque tout $\underline{\omega}$. Comme γ peut être choisi comme un nombre rationnel arbitrairement proche de 0, on trouve que

$$\limsup_{n\to\infty} \frac{1}{n} \log \nu(\{x \in X \,/\, S_n(\underline{\omega}, x) > n\epsilon\}) \leq -c(\epsilon)$$

pour $\tilde{\mathbb{P}}$-presque tout $\underline{\omega}$. □

1.2.4 Théorème de la limite locale

Un raffinement de l'étude des opérateurs de Fourier P_{it} permet de démontrer un théorème de la limite locale. On a besoin d'imposer une condition supplémentaire sur l'observable φ. On suppose toujours que $\varphi \in \mathcal{B}_0$ est à valeurs réelles, borné, et d'intégrale nulle par rapport à μ.

Définition 1.2.19. *On dira que φ est apériodique si pour tout $t \in \mathbb{R}$, $t \neq 0$, le rayon spectral de P_{it} est strictement inférieur à 1.*

Théorème 1.2.20 (Théorème de la limite locale)**.** *Si $\sigma^2 > 0$ et si φ est apériodique, alors pour tout $\nu \in \mathcal{M}_{\mathcal{B}}$ et tout intervalle borné $I \subset \mathbb{R}$,*

$$\lim_{n\to\infty} \sup_{s\in\mathbb{R}} \left| \sigma\sqrt{n}\, \tilde{\mathbb{P}} \otimes \nu(s + S_n \in I) - \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} |I| \right| = 0.$$

Démonstration. On suit la preuve de Breiman [28] pour le cas i.i.d., et celle de Rousseau-Egele [127] pour les applications (déterministes) uniformément dilatantes de l'intervalle. Par un argument de densité, il suffit de prouver que

$$\left| \sigma\sqrt{n}\, \mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) - \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du \right|$$

converge vers 0 uniformément en $s \in \mathbb{R}$ lorsque $n \to \infty$, pour tout $g \in L^1(\mathbb{R})$ dont la transformée de Fourier $\hat{g}$ est continue à support compact. En utilisant la formule d'inversion de Fourier, on trouve

$$\sigma\sqrt{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) = \frac{\sigma\sqrt{n}}{2\pi} \int_{\mathbb{R}} e^{its} \hat{g}(t) \left(\int_X P_{it}^n(f)\, dm \right) dt.$$

Soit $\delta > 0$ tel que le support de $\hat{g}$ est inclus dans $[-\delta, +\delta]$, et en se rappelant que $\lambda(it) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$ et $\langle m(it), f\rangle\langle m, h(it)\rangle = 1 + \mathcal{O}(|t|)$, on choisit $0 < \tilde{\delta} < \delta$ assez petit de telle sorte que $|\lambda(it)| \leq 1 - \frac{\sigma^2 t^2}{4} \leq e^{-\frac{t^2\sigma^2}{4}}$ et $|\langle m(it), f\rangle\langle m, h(it)\rangle - 1| \leq C|t|$ pour $|t| < \tilde{\delta}$. Comme

$$\frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du = \frac{\hat{g}(0)\sigma}{2\pi} \int_{\mathbb{R}} e^{\frac{its}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} dt$$

et

$$\int_X P_{it}^n(f)\, dm = \lambda(it)^n \langle m, h(it)\rangle\langle m(it), f\rangle + \langle m, Q(it)^n f\rangle$$

pour tout $|t| < \tilde{\delta}$, on peut écrire

$$\begin{aligned}
\sigma\sqrt{n}\,\mathbb{E}_{\tilde{\mathbb{P}}\otimes\nu}(g(S_n + s)) \quad &- \quad \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2n\sigma^2}} \int_{\mathbb{R}} g(u)du \\
&= \quad \frac{\sigma}{2\pi}\left(\int_{|t|<\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} \left(\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}\right) dt\right. \\
&+ \quad \int_{|t|<\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} \hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n \left(\langle m(\frac{it}{\sqrt{n}}), f\rangle\langle m, h(\frac{it}{\sqrt{n}})\rangle - 1\right) dt \\
&+ \quad \sqrt{n}\int_{|t|<\tilde{\delta}} e^{its}\hat{g}(t)\langle m, Q(it)^n f\rangle dt \\
&+ \quad \left.\sqrt{n}\int_{\tilde{\delta}\leq|t|\leq\delta} e^{its}\hat{g}(t)\langle m, P_{it}^n f\rangle dt - \hat{g}(0)\int_{|t|\geq\tilde{\delta}\sqrt{n}} e^{\frac{its}{\sqrt{n}}} e^{-\frac{\sigma^2 t^2}{2}} dt\right) \\
&= \quad \frac{\sigma}{2\pi}\left(A_n^{(1)}(s) + A_n^{(2)}(s) + A_n^{(3)}(s) + A_n^{(4)}(s) + A_n^{(5)}(s)\right).
\end{aligned}$$

On a $\left|A_n^{(1)}(s)\right| \leq \int_{|t|<\tilde{\delta}\sqrt{n}} \left|\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}\right| dt$, d'où $\sup_{s\in\mathbb{R}} \left|A_n^{(1)}(s)\right| \to 0$ par convergence dominée, puisque $\left|\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n - \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}\right| \leq \|\hat{g}\|_{\sup}\left(e^{-\frac{\sigma^2 t^2}{4}} + e^{-\frac{\sigma^2 t^2}{2}}\right)$ est intégrable sur $\mathbb{R}$ et $\hat{g}(\frac{t}{\sqrt{n}})\lambda(\frac{it}{\sqrt{n}})^n \to \hat{g}(0)e^{-\frac{\sigma^2 t^2}{2}}$.
On peut majorer le second terme par $\frac{C\|\hat{g}\|_{\sup}}{\sqrt{n}} \int_{|t|<\tilde{\delta}\sqrt{n}} |t|e^{-\frac{\sigma^2 t^2}{4}} dt = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$ et donc $\sup_{s\in\mathbb{R}} \left|A_n^{(2)}(s)\right| \to 0$.
Le troisième terme est majoré par $C\sqrt{n}\|\hat{g}\|_{\sup}\rho^n$, et donc $\sup_{s\in\mathbb{R}} \left|A_n^{(3)}(s)\right| \to 0$.
Par convergence dominée, on a clairement $\sup_{s\in\mathbb{R}} \left|A_n^{(5)}(s)\right| \to 0$, il reste à traiter le quatrième terme. C'est ici que l'hypothèse d'apériodicité va jouer un rôle. Si l'on dénote par $r(P_{it})$ le rayon spectral de l'opérateur P_{it}, on sait que la fonction s.c.s. $t \mapsto r(P_{it})$ atteint son maximum sur le compact $\{\tilde{\delta} \leq |t| \leq \delta\}$ en un point t_0. Ce maximum vaut donc $r(P_{it_0})$ qui est strictement inférieur à 1 par l'hypothèse d'apériodicité, puisque $t_0 \neq 0$. Comme l'ensemble $\{P_{it}\}_{\tilde{\delta}\leq|t|\leq\delta}$ est borné, il existe donc C et $\theta < 1$ tel que $\|P_{it}^n\| \leq C\theta^n$ pour tout $\tilde{\delta} \leq |t| \leq \delta$ et $n \geq 0$. On a ainsi $\sup_{s\in\mathbb{R}} \left|A_n^{(4)}(s)\right| \leq C\sqrt{n}\theta^n\|\hat{g}\|_{\sup}\|m\|\|f\| \to 0$, ce qui termine la preuve. $\square$

On donne à présent un critère permettant de vérifier l'hypothèse d'apériodicité.

Proposition 1.2.21. *Supposons que la mesure stationnaire μ est équivalente à m et que le rayon spectral (resp. le rayon spectral essential) de P_{it} est plus petit (resp. strictement plus petit) que 1 pour tout $t \in \mathbb{R}$. Si φ n'est pas apériodique, alors il existe $t \neq 0$, $\lambda \in \mathbb{C}$ avec $|\lambda| = 1$ et une fonction mesurable $g : X \to \mathbb{C}$ telle que $gh \in \mathcal{B}$ et $\lambda g(T_\omega x) = e^{it\varphi(x)} g(x)$ pour m-p.t. $x \in X$ et $\mathbb{P}$-p.t. $\omega \in \Omega$.*

Démonstration. Si φ n'est pas apériodique, alors il existe par définition un certain $t \neq 0$ tel que le rayon spectral de P_{it} est plus grand ou égal à 1. On en déduit que le rayon spectral de P_{it} est exactement égal à 1. Comme le rayon spectral essentiel de P_{it} est strictement inférieur à 1, cela entraine qu'il existe une valeur propre λ de module 1 pour P_{it}, avec un vecteur propre associé $f \in \mathcal{B}$. On pose $g = \frac{f}{h}$. g est bien définie m-presque partout par hypothèse sur μ. On a alors $P(\phi g h) = gh$, où $\phi = \bar{\lambda} e^{it\varphi}$. On relève cette identité sur le produit-gauche : par le lemme 1.3.3, on a $P_F(\phi_p g_p h_p) = g_p h_p$, où P_F est l'opérateur de transfert du système produit-gauche F par rapport à la mesure $\tilde{\mathbb{P}} \otimes m$. Voir le paragraphe 1.3 pour les notations. D'après la proposition 1.1 de Morita [117], on en déduit que $g_p \circ F = \phi_p g_p$, $\tilde{\mathbb{P}} \otimes m$-p.p., ce qui termine la preuve. □

Remarque 1.2.22.

1. Les hypothèses sur le rayon spectral et le rayon spectral essentiel de la proposition précédente se démontrent généralement en prouvant une inégalité de Lasota-Yorke pour chaque opérateur P_{it}, en suivant la même stratégie que pour l'opérateur P. On pourra consulter Rousseau-Egele [127], Morita [117], Broise [29], ou Aaronson, Denker, Sarig et Zweimuller [2] pour de telles preuves dans le cadre déterministe et unidimensionnel.
2. La proposition précédente montre que si φ n'est pas apériodique pour le système aléatoire, alors il n'est pas apériodique pour presque tout système déterministe T_ω au sens usuel, et presque toutes les équations d'apériodicité partagent une même solution régulière. En particulier, si Ω est fini, et si φ est apériodique pour au moins une application T_ω, alors il est apériodique pour le système aléatoire. L'apériodicité dans le cadre déterministe peut être vérifiée par des méthodes classiques, on renvoie à [1, 2, 29] parmi d'autres pour plus de détails.

1.2.5 Principe d'invariance presque sure

Dans cette section, on démontre le principe d'invariance presque sure pour des observables à valeurs dans $\mathbb{R}^d$. Il s'agit d'un renforcement du théorème de la limite centrale qui entraine de nombreuses conséquences, comme la loi du logarithme itéré, la version fonctionnelle du théorème de la limite centrale, etc ... On renvoie à Melbourne et Nicol [112] et aux références présentes dans ce papier pour plus de détails.
On donne d'abord la définition :

Définition 1.2.23. *Pour $\lambda \in (0, \frac{1}{2}]$, et Σ^2 $d \times d$ matrice symétrique, semi-définie positive (éventuellement dégénérée), on dira qu'un processus $(X_n)_n$ à valeurs dans $\mathbb{R}^d$ satisfait le principe d'invariance presque sure avec l'exposant d'erreur λ et la covariance asymptotique Σ^2 s'il existe deux processus $(Y_n)_n$ et $(Z_n)_n$, définis sur un autre espace probabilisé, tels que*

1. *les processus $(X_n)_n$ et $(Y_n)_n$ ont la même distribution ;*
2. *les variables aléatoires Z_n sont indépendantes et distribuées selon la loi $\mathcal{N}(0, \Sigma^2)$;*
3. *presque surement, $\left|\sum_{k=0}^{n-1} Y_k - \sum_{k=0}^{n-1} Z_k\right| = o(n^\lambda)$.*

Soit $\varphi : X \to \mathbb{R}^d$ un observable borné, à valeurs dans $\mathbb{R}^d$ tel que chaque composante $\varphi_j : X \to \mathbb{R}$, $j = 1, \dots, d$, appartient à $\mathcal{B}_0$, avec $\int_X \varphi_j \, d\mu = 0$. On définit comme auparavant $X_k(\underline{\omega}, x) = \varphi(T^k_{\underline{\omega}} x)$.

Théorème 1.2.24 (Principe d'invariance presque sure)**.** *La matrice de covariance*

$$\frac{1}{n} \mathrm{cov}\left(\sum_{k=0}^{n-1} X_k\right)$$

converge vers une matrice Σ^2 et le processus $(X_n)_n$, défini sur l'espace probabilisé $(\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$ satisfait le principe d'invariance presque sure avec Σ^2 comme covariance asymptotique, pour tout exposant d'erreur $\lambda > \frac{1}{4}$.

Démonstration. On applique les résultats de Gouëzel [64]. Plus précisément, on va construire une famille d'opérateurs $(\mathcal{L}_t)_{t \in \mathbb{R}^d}$ sur $\mathcal{B}$ qui encode la fonction caractéristique du processus $(X_n)_n$ et on va vérifier les hypothèses (I1) et (I2) de [64]. Pour $k \geq 0$ et $j_1, \dots, j_k \in \{1, \dots, d\}$, on définit $C_{j_1,\dots,j_k}$ by $C_{j_1,\dots,j_k}(f) = P(\varphi_{j_1} \dots \varphi_{j_k} f)$. Les hypothèses faites sur $\mathcal{B}$ et $\mathcal{B}_0$ montrent que $C_{j_1,\dots,j_k}$ agit continument sur $\mathcal{B}$, avec une norme majorée par $C^k \|\varphi_{j_1}\|_0 \dots \|\varphi_{j_k}\|_0$. Pour $t = (t_1, \dots, t_d) \in \mathbb{R}^d$, on pose

$$\mathcal{L}_t = \sum_{k=0}^{\infty} \frac{i^k}{k!} \sum_{j_1,\dots,j_k=1}^{d} t_{j_1} \dots t_{j_k} C_{j_1,\dots,j_k}.$$

On définit ainsi sur $\mathbb{R}^d$ une famille réelle-analytique d'opérateurs bornés sur $\mathcal{B}$, puisque

$$\begin{aligned}\sum_{k=0}^{\infty} \left\| \frac{i^k}{k!} \sum_{j_1,\dots,j_k=1}^{d} t_{j_1} \dots t_{j_k} C_{j_1,\dots,j_k} \right\| &\leq \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{j_1,\dots,j_k=1}^{d} |t_{j_1}| \dots |t_{j_k}| \|C_{j_1,\dots,j_k}\| \\ &\leq e^{C \sum_{j=0}^{d} |t_j| \|\varphi_j\|_0} < \infty.\end{aligned}$$

Pour $t \in \mathbb{R}^d$ et $f \in \mathcal{B}$, on a $\mathcal{L}_t(f) = P(e^{i\langle t, \varphi \rangle} f)$, et donc la famille $\{\mathcal{L}_t\}_{t \in \mathbb{R}^d}$ encode la fonction caractéristique du processus (X_n) au sens de [64], d'après le lemme 1.2.7. Puisque $\mathcal{L}_0 = P$ a un trou spectral sur $\mathcal{B}$, l'hypothèse (I1) est vérifiée. Pour montrer l'hypothèse (I2), il suffit, d'après la proposition 2.3 de [64], que l'application $t \mapsto \mathcal{L}_t$ soit continue en $t = 0$, mais cela découle immédiatement de la réelle-analycité de cette application. □

1.3 Le théorème de la limite centrale par la méthode des martingales

L'objectif de cette section est de montrer que la méthode classique d'approximation par les martingales due à Gordin [59] (voir aussi Liverani [103]) permet de démontrer le théorème de la limite centrale annealed pour les systèmes dynamiques aléatoires. On donnera donc ici une preuve alternative du théorème 1.2.6 dans le cas où la mesure ν est la mesure stationnaire μ. La méthode utilisée sera réinvestie dans le paragraphe suivant pour démontrer un TCL pour les lemmes dynamiques de Borel-Cantelli, un contexte où le processus étudié n'est plus stationnaire, l'observable dépendant alors du temps n.

On fera tout au long de ce paragraphe l'hypothèse que la mesure stationnaire μ est équivalente à m. Les applications T_ω sont alors toutes non-singulières par rapport à μ, ce qui nous permet de définir l'opérateur de transfert moyenné P_μ du système aléatoire par rapport à μ. Il vérifie $P_\mu(f) = \frac{P(hf)}{h}$ pour tout $f \in L^1(\mu)$, où h est la densité de μ et P est l'opérateur de transfert moyenné défini par rapport à m.

On rappelle que l'opérateur de Koopman, ou opérateur de transition, agit sur les fonctions définies sur X par $Uf(x) = \int_\Omega f(T_\omega x)\, d\mathbb{P}(\omega)$. À U est associée une probabilité de transition sur X donnée par $U(x, A) = U(\mathbb{1}_A)(x) = \mathbb{P}(\{\omega \,/\, T_\omega x \in A\})$. On rappelle aussi que la mesure stationnaire μ vérifie $\mu U = \mu$ par définition. On peut alors associer à U et μ une chaîne de Markov canonique de la façon suivante.

Soit $\Omega^\star = X^{\mathbb{N}_0} = \{\underline{x} = (x_0, x_1, x_2, \ldots, x_n, \ldots)\}$, muni de la tribu $\mathcal{F}$ engendrée par les cylindres. Comme X est un espace polonais, $\mathcal{F}$ est la tribu borélienne associée à la topologie produit sur $\Omega^\star$. Le théorème de Ionescu-Tulcea [120, Proposition V-1-1] montre qu'il existe une unique mesure de probabilité sur $\Omega^\star$ telle que

$$\int_{\Omega^\star} f(\underline{x})\, d\mu_c(\underline{x}) = \int_X \mu(dx_0) \int_X U(x_0, dx_1) \ldots \int_X U(x_{n-1}, dx_n) f(x_0, \ldots, x_n)$$

pour tout n et toute fonction mesurable bornée $f : \Omega^\star \to \mathbb{R}$ qui ne dépend que de $x_0, \ldots, x_n$. Si on dénote par x_n l'application qui à chaque $\underline{x}$ associe sa n-ième coordonnée x_n, alors $\{x_n\}_{n\geq 0}$ est une chaîne de Markov définie sur l'espace probabilisé $(\Omega^\star, \mathcal{F}, \mu_c)$, avec distribution initiale μ_c et probabilité de transition U. Par stationnarité de la mesure μ, chaque x_n est distribué suivant la loi μ.

On peut définir le décalage à gauche τ sur $\Omega^\star$. Par stationnarité, celui-ci préserve la mesure μ_c. On rappelle aussi que l'on dispose d'un système produit gauche F sur $\Omega^{\mathbb{N}} \times X$, défini par $F(\underline{\omega}, x) = (\sigma\underline{\omega}, T_{\omega_1} x)$, où σ est le décalage à gauche sur $\Omega^{\mathbb{N}}$. F préserve la probabilité $\tilde{\mathbb{P}} \otimes \mu = \mathbb{P}^{\otimes \mathbb{N}} \otimes \mu$. On peut relier ce système produit-gauche au système symbolique τ sur $\Omega^\star$. En effet, si on définit $\Phi : \Omega^{\mathbb{N}} \times X \to \Omega^\star$ par $\Phi(\underline{\omega}, x) = (x, T_{\omega_1} x, T_{\omega_2} T_{\omega_1} x, \ldots, T_{\omega_n} \ldots T_{\omega_1} x, \ldots) = \{p(S^n(\underline{\omega}, x))\}_{n\geq 0}$, avec $p(\underline{\omega}, x) = x$, alors on a :

Lemme 1.3.1. *Φ est mesurable, envoie $\tilde{\mathbb{P}} \otimes \mu$ sur μ_c et vérifie $\Phi \circ F = \tau \circ \Phi$.*

En d'autres termes, $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ est une extension de $(\Omega^\star, \tau, \mu_c)$.

Démonstration. La seule chose non-évidente à prouver est le fait que Φ envoie $\tilde{\mathbb{P}} \otimes \mu$ sur μ_c. Il suffit de montrer que

$$\int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \dots f_n(T^n_{\underline{\omega}} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) \, d\mu(x)$$
$$= \int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \dots \int_X U(x_{n-1}, dx_n) f_n(x_n)$$

pour tout $n \geq 0$ et toutes les fonctions mesurables bornées $f_0, \dots, f_n : X \to \mathbb{R}$. On procède par récurrence sur n, le cas $n = 0$ étant trivial. On a

$$\begin{aligned}
&\int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \dots \int_X U(x_{n-1}, dx_n) f_n(x_n) \int_X U(x_n, dx_{n+1}) f_{n+1}(x_{n+1}) \\
&= \int_X \mu(dx_0) f_0(x_0) \int_X U(x_0, dx_1) f_1(x_1) \dots \int_X U(x_{n-1}, dx_n) f_n(x_n) U f_{n+1}(x_n) \\
&= \int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \dots f_n(T^n_{\underline{\omega}} x) U f_{n+1}(T^n_{\underline{\omega}} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) \, d\mu(x) \\
&= \int_X \mu(dx) \int_{\Omega^{\mathbb{N}}} f_0(x) f_1(T^1_{\underline{\omega}} x) \dots f_n(T^n_{\underline{\omega}} x) f_{n+1}(T^{n+1}_{\underline{\omega}} x) \, d\tilde{\mathbb{P}}(\underline{\omega}) \, d\mu(x)
\end{aligned}$$

ce qui termine la preuve. □

Soit $\pi_n : \Omega^\star \to X$ la projection $\pi_n(x_0, \dots, x_n, \dots) = x_n$ et $\pi = \pi_0$. On relève chaque fonction $\phi : X \to \mathbb{R}$ en une fonction ϕ_π sur $\Omega^\star$ par $\phi_\pi = \phi \circ \pi$. On a alors $\mathbb{E}_\mu(\phi) = \mathbb{E}_{\mu_c}(\phi_\pi)$. On remarque aussi que $\pi_n = \pi \circ \tau^n$ et $p \circ F^n = \pi_n \circ \Phi$.
Soit $\varphi : X \to \mathbb{R}$ un observable borné à valeurs réelles, appartenant à $\mathcal{B}_0$, tel que $\int_X \varphi \, d\mu = 0$. On définit $X_k = \varphi \circ p \circ F^k$ et $S_n = \sum_{k=0}^{n-1} X_k$. On a alors $X_k = \varphi \circ \pi_k \circ \Phi = \varphi_\pi \circ \tau^k \circ \Phi$. Ainsi, $S_n = (\sum_{k=0}^{n-1} \varphi_\pi \circ \tau^k) \circ \Phi$, et donc la loi de S_n sous la probabilité $\tilde{\mathbb{P}} \otimes \mu$ est la loi de la n-ième somme de Birkhoff de l'observable φ_π pour le système symbolique $(\Omega^\star, \tau, \mu_c)$. Pour démontrer un TCL pour S_n, il est donc suffisant de le démontrer pour l'observable φ_π et le système $(\Omega^\star, \tau, \mu_c)$.
Pour cela, on va utiliser l'opérateur de Koopman $\tilde{U}$ et l'opérateur de transfert $\tilde{P}$ associés à $(\Omega^\star, \tau, \mu_c)$. Comme τ préserve μ_c, ces opérateurs satisfont les relations $\tilde{P}^k \tilde{U}^k f = f$ et $\tilde{U}^k \tilde{P}^k f = \mathbb{E}_{\mu_c}(f | \mathcal{F}_k)$ pour tout $f \in L^1(\mu_c)$, où $\mathcal{F}_k = \tau^{-k} \mathcal{F} = \sigma(x_k, x_{k+1}, \dots)$.
L'opérateur $\tilde{P}$ est lié à l'opérateur de transfert moyenné P_μ de la façon suivante :

Lemme 1.3.2. *Pour tout $\phi \in L^1(\mu)$, on a $\tilde{P}(\phi_\pi) = (P\phi)_\pi$.*

On va déduire ce lemme du résultat correspondant pour l'opérateur P_F associé au système produit-gauche $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$. Pour $\phi \in L^1(\mu)$, on introduit la notation $\phi_p = \phi \circ p \in L^1(\tilde{\mathbb{P}} \otimes \mu)$. La même preuve que pour le lemme 1.1.2, mais avec la mesure μ au lieu de m donne :

Lemme 1.3.3. *Pour tout $\phi \in L^1(\mu)$, on a $P_F(\phi \circ p) = (P_\mu \phi) \circ p$.*

Démonstration du lemme 1.3.2. Soit $\psi : \Omega^\star \to \mathbb{R}$ un élément quelconque de $L^\infty(\mu_c)$. On doit montrer que $\int_{\Omega^\star} \phi_\pi(\psi \circ \tau)\, d\mu_c = \int_{\Omega^\star} (P_\mu \phi)_\pi \psi\, d\mu_c$. On a

$$\begin{aligned}
&\int_{\Omega^\star} \phi_\pi(\psi \circ \tau)\, d\mu_c = \int_{\Omega^{\mathbb{N}} \times X} (\phi \circ \pi \circ \Phi)(\psi \circ \tau \circ \Phi)\, d(\tilde{\mathbb{P}} \otimes \mu) \\
&= \int_{\Omega^{\mathbb{N}} \times X} (\phi \circ p)(\psi \circ \Phi \circ F)\, d(\tilde{\mathbb{P}} \otimes \mu) = \int_{\Omega^{\mathbb{N}} \times X} P_F(\phi \circ p)(\psi \circ \Phi)\, d(\tilde{\mathbb{P}} \otimes \mu) \\
&= \int_{\Omega^{\mathbb{N}} \times X} (P_\mu \phi \circ p)(\psi \circ \Phi)\, d(\tilde{\mathbb{P}} \otimes \mu) = \int_{\Omega^{\mathbb{N}} \times X} (P_\mu \phi \circ \pi \circ \Phi)(\psi \circ \Phi)\, d(\tilde{\mathbb{P}} \otimes \mu) \\
&= \int_{\Omega^\star} (P_\mu \phi)_\pi \psi\, d\mu_c
\end{aligned}$$

□

On est en mesure à présent de démontrer le TCL :

Théorème 1.3.4. *Soit $\varphi \in \mathcal{B}_0$ un observable borné à valeurs réelles, avec $\int_X \varphi\, d\mu = 0$. Soit $S_n(\underline{\omega}, x) = \sum_{k=0}^{n-1} \varphi(T^k_{\underline{\omega}} x)$. Alors il existe $\sigma^2 \geq 0$ tel que $\frac{S_n}{\sqrt{n}}$ converge en loi vers $\mathcal{N}(0, \sigma^2)$ sous la probabilité $\tilde{\mathbb{P}} \otimes \mu$, et $\sigma^2 = 0$ si et seulement si il existe $\psi : X \to \mathbb{R}$ mesurable avec $\psi h \in \mathcal{B}$ tel que $\varphi(x) = \psi(T_\omega x) - \psi(x)$ pour $\tilde{\mathbb{P}}$-p.t $\underline{\omega}$ et μ-p.t. x.*

Démonstration. D'après la décroissance des corrélations (point (1) de la proposition 1.2.1), on a $\|P^n_\mu \varphi\|_{L^1_\mu} \leq C\lambda^n \|\varphi\|_0$, puisque $\|g\|_{L^\infty_m} = \|g\|_{L^\infty_\mu}$, μ étant équivalente à m par hypothèse. Pour tout $p > 1$, on a $\|P^n_\mu \varphi\|_{L^p_\mu} \leq \|P^n_\mu \varphi\|^{p-1}_{L^\infty_\mu} \|P^n_\mu \varphi\|_{L^1_\mu} \leq C\|\varphi\|^{p-1}_{L^\infty_\mu} \|\varphi\|_0 \lambda^n$, d'après l'inégalité de Hölder. Ainsi, la série $\psi = \sum_{n \geq 1} P^n_\mu \varphi$ est sommable dans $L^p(\mu)$ pour tout $1 \leq p < \infty$, et en particulier, $\psi \in L^2(\mu)$.
On définit $\chi = \varphi_\pi + \psi_\pi - \psi_\pi \circ \tau$ sur $\Omega^\star$. D'après le lemme 1.3.2, χ vérifie $\tilde{P}\chi = \tilde{P}(\varphi_\pi) + \tilde{P}(\psi_\pi) - \tilde{P}\tilde{U}(\psi_\pi) = (P_\mu \varphi)_\pi + (P_\mu \psi - \psi)_\pi = 0$, puisque $P_\mu \psi = \psi - P_\mu \varphi$.
On a que $\{\chi \circ \tau^k\}_{k \geq 0}$ est une suite de différences de martingales inversées par rapport à la filtration $\{\mathcal{F}_k\}_{k \geq 0}$: en effet, on a $\mathbb{E}_{\mu_c}(\chi \circ \tau^k | \mathcal{F}_{k+1}) = \tilde{U}^{k+1} \tilde{P}^{k+1} \tilde{U}^k \chi = \tilde{U}^{k+1} \tilde{P} \chi = 0$. Cette suite est clairement stationnaire et dans L^2, puisque $\chi \in L^2$. Elle est ergodique car le système $(\Omega^\star, \tau, \mu_c)$ est ergodique : c'est en effet un facteur du système $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ qui est ergodique d'après la proposition 1.1.4. Ainsi $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \chi \circ \tau^k$ converge en loi vers $\mathcal{N}(0, \sigma^2)$, avec $\sigma^2 = \mathbb{E}_{\mu_c}(\chi^2)$, d'après Billingsley [23].
Puisque $\sum_{k=0}^{n-1} \varphi_\pi \circ \tau^k = \sum_{k=0}^{n-1} \chi \circ \tau^k + \psi_\pi \circ \tau^n - \psi_\pi$, et $\frac{\psi_\pi \circ \tau^n - \psi_\pi}{\sqrt{n}}$ converge vers 0 en probabilité (car il converge vers 0 en norme L^1), on obtient que $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \varphi_\pi \circ \tau^k$ converge en loi vers $\mathcal{N}(0, \sigma^2)$, et comme on l'a vu plus haut, ceci entraine le TCL pour S_n.
Si $\sigma^2 = 0$, alors $\chi = 0$, μ_c-p.p., ce qui entraine que $\varphi_\pi = \psi_\pi \circ \tau - \psi_\pi$, μ_c-p.p., et donc $\varphi(x) = \psi(T_\omega x) - \psi(x)$ pour $\tilde{\mathbb{P}}$-p.t $\underline{\omega}$ et μ-p.t. x. Pour obtenir la régularité de φ, on observe que la décroissance des corrélations montre que $\|hP^n_\mu(\varphi)\|_{\mathcal{B}} = \|P^n(\varphi h)\|_{\mathcal{B}}$ converge exponentiellement vite vers 0. Ainsi, la série $\psi h = \sum_{n \geq 1} hP^n_\mu(\varphi)$ est sommable dans $\mathcal{B}$, d'où le résultat. □

En suivant la méthode de Melbourne et Török [113], on peut montrer que l'on a aussi convergence des moments. Soit W une variable aléatoire distribuée suivant la loi $\mathcal{N}(0,\sigma^2)$. Si $\sigma^2 = 0$, alors W est constante égale à 0 presque surement.

Proposition 1.3.5. *Pour $p \in [1,+\infty)$, on a*

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}\left|\frac{S_n}{\sqrt{n}}\right|^p \to \mathbb{E}|W|^p.$$

Démonstration. Si $\sigma^2 = 0$, alors la caractérisation par les cobords donne que S_n est bornée dans chaque L^p (on a vu au cours de la preuve du théorème 1.3.4 que le cobord ψ appartient à tous les L^p), d'où le résultat. On traite à présent le cas $\sigma^2 > 0$. On reprend les notations de la preuve du théorème 1.3.4. On a $\sum_{k=0}^{n-1}\varphi_\pi\circ\tau^k = \sum_{k=0}^{n-1}\chi\circ\tau^k + \psi_\pi\circ\tau^n - \psi_\pi$. Comme $\{\sum_{k=0}^{n-1}\chi\circ\tau^k\}_k$ est une martingale inversée, on obtient par l'inégalité de Burkholder [33] que $\|\sum_{k=0}^{n-1}\chi\circ\tau^k\|_{L^p_{\mu_c}} = \mathcal{O}(n^{\frac{1}{2}})$, ce qui entraine que $\|\sum_{k=0}^{n-1}\varphi_\pi\circ\tau^k\|_{L^p_{\mu_c}} = \mathcal{O}(n^{\frac{1}{2}})$. D'autre part, on a $\|\sum_{k=0}^{n-1}\varphi_\pi\circ\tau^k\|_{L^p_{\mu_c}} = \|S_n\|_{L^p(\tilde{\mathbb{P}}\otimes\mu)}$. Ainsi, pour chaque $p > 1$, les moments d'ordre p de $\frac{S_n}{\sqrt{n}}$ sont bornés. La convergence des moments découle alors immédiatement d'un résultat connu en probabilité, voir par exemple l'exercice 2.5 page 86 de Durrett [51]. □

1.4 Lemmes dynamiques de Borel-Cantelli

Dans toute cette section, on fait l'hypothèse que la décroissance des corrélations est exponentielle pour les observables dans $\mathcal{B}_0$ contre les observables dans $L^1(\mu)$, c'est-à-dire que

$$\left|\int_X f\,U^n g\,d\mu - \int_X f\,d\mu\int_X g\,d\mu\right| \le C\lambda^n\|f\|_0\|g\|_{L^1_\mu}$$

pour tout $f \in \mathcal{B}_0$ et $g \in L^1(\mu)$. Comme on l'a vu à la proposition 1.2.1, c'est le cas lorsque $\mathcal{B}$ s'injecte continument dans $L^\infty(m)$ et que la densité de μ est bornée inférieurement.

1.4.1 Convergence presque sure

On va utiliser le résultat suivant, du à Sprindzuk [132] :

Proposition 1.4.1. *Soit f_k une suite de fonctions mesurables positives sur un espace mesuré (Y,ν) et soit $\bar{f}_k$, φ_k deux suites de nombres réels satisfaisant $0 \le \bar{f}_k \le \varphi_k \le M$ pour tout $k \ge 1$ et un certain $M > 0$. On suppose que*

$$\int_Y\left(\sum_{m<k\le n}\left(f_k(y)-\bar{f}_k\right)\right)^2 d\nu(y) \le C\sum_{m<k\le n}\varphi_k$$

pour tous les entiers $m < n$ et un certain $C > 0$. Alors

$$\sum_{1\le k\le n} f_k(y) = \sum_{1\le k\le n}\bar{f}_k + O(\Phi^{1/2}(n)\log^{3/2+\epsilon}\Phi(n))$$

pour ν-p.t. $y \in Y$, tout $\epsilon > 0$, et $\Phi(n) = \sum_{1 \le k \le n} \varphi_k$.

On applique ce résultat à l'espace $(\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$. En adaptant au cas aléatoire la démarche de Kim [95], et en utilisant notre hypothèse sur la décroissance des corrélations, on obtient

Théorème 1.4.2. *Soit ϕ_n une suite de fonctions positives, avec $\phi_n \in \mathcal{B}_0$ et $\sup_n \|\phi_n\|_0 < \infty$, et telle que $E_n \to \infty$, où $E_n = \sum_{k=0}^{n-1} \int \phi_k d\mu$. Alors*

$$\lim_{n \to \infty} \frac{1}{E_n} \sum_{k=0}^{n-1} \phi_k(T_{\underline{\omega}}^k x) = 1,$$

pour $\tilde{\mathbb{P}} \otimes \mu$-p.t. $(\underline{\omega}, x) \in \tilde{\Omega} \times X$.

Démonstration. On applique le théorème de Sprindzuk avec $(Y, \nu) = (\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$, $f_k = (\phi_k)_p \circ F^k$, $\bar{f}_k = \varphi_k = \int \phi_k d\mu$. On a

$$\begin{aligned}
&\int_Y \left(\sum_{m<k\le n} \left(f_k(y) - \bar{f}_k \right) \right)^2 d\nu(y) \\
&\le 2 \sum_{m<k\le l\le n} \left| \int_{\tilde{\Omega}\times X} (\phi_k)_p (\phi_l)_p \circ F^{l-k} d\tilde{\mathbb{P}} \otimes \mu - \int \phi_k d\mu \int \phi_l d\mu \right| \\
&= 2 \sum_{m<k\le l\le n} \left| \int_X \phi_k U^{l-k} \phi_l d\mu - \int \phi_k d\mu \int \phi_l d\mu \right| \le C \sum_{m<k\le l\le n} \lambda^{l-k} \|\phi_k\|_0 \|\phi_l\|_{L^1_\mu} \\
&\le C \sum_{m<l\le n} \varphi_l.
\end{aligned}$$

□

1.4.2 Étude des fluctuations

On peut préciser la nature des fluctuations dans la convergence presque sure donnée par le théorème 1.4.2. On suppose toujours que ϕ_n est une suite de fonctions positives, bornée dans $\mathcal{B}_0$ avec $E_n = \sum_{k=0}^{n-1} \int \phi_k \, d\mu$ divergente. On fera aussi l'hypothèse que $\mathcal{B} = \mathcal{B}_0$ est une algèbre de Banach qui s'injecte continument dans $L^\infty(m)$, ce qui est le cas lorsque $\mathcal{B} = \mathrm{BV}$ en dimension 1, ou lorsque $\mathcal{B} = V_\alpha$ est l'espace Quasi-Hölder. On supposera aussi que la densité h de μ ne s'annule pas et vérifie $\frac{1}{h} \in \mathcal{B}$: c'est le cas lorsque $\mathcal{B}$ est BV en dimension 1 ou Quasi-Hölder et que la densité est bornée inférieurement.
On pose $\widetilde{\phi}_n = \phi_n - \int \phi_n \, d\mu$, $S_n(\underline{x}) = \sum_{k=0}^{n-1} \widetilde{\phi}_k(x_k)$, et on introduit une approximation par des martingales similaire à celle du paragraphe 1.3.
On définit $\psi_0 = 0$ et, pour $n \ge 1$, on définit

$$\psi_n = \sum_{k=1}^{n} P_\mu^k \widetilde{\phi}_{n-k}.$$

Un calcul simple montre que $P_\mu \psi_n = \psi_{n+1} - P_\mu \widetilde{\phi}_n$. On définit alors sur $\Omega^\star$,

$$\chi_n = (\widetilde{\phi}_n)_\pi - (\psi_{n+1})_\pi \circ \tau + (\psi_n)_\pi.$$

On a alors

$$\sum_{k=0}^{n-1} \chi_k \circ \tau^k = \sum_{k=0}^{n-1} (\widetilde{\phi}_k)_\pi \circ \tau^k - (\psi_n)_\pi \circ \tau^n.$$

D'après le lemme 1.3.2, on a $\tilde{P}\chi_n = (P_\mu \widetilde{\phi}_n - \psi_{n+1} + P_\mu \psi_n)_\pi = 0$. Ceci montre que $\{\chi_k \circ \tau^k\}_{k\geq 0}$ est une suite de différences de martingales inversées par rapport à la filtration $\{\mathcal{F}_k\}_{k\geq 0}$: en effet, $\mathbb{E}_{\mu_c}(\chi_k \circ \tau^k | \mathcal{F}_{k+1}) = \tilde{U}^{k+1}\tilde{P}^{k+1}\tilde{U}^k \chi_k = \tilde{U}^{k+1}\tilde{P}\chi_k = 0$.
La décroissance des corrélations, et les hypothèses faites au début de cette sous-section, entrainent que ψ_n est une suite bornée dans $\mathcal{B}$.
On pose $a_n^2 = \mathbb{E}_{\mu_c}(S_n^2)$.

Théorème 1.4.3. *On a l'alternative suivante :*

1. *Soit (a_n^2) est bornée, et dans ce cas S_n est bornée presque surement ;*
2. *Soit (a_n^2) converge vers $+\infty$, et dans ce cas $\frac{S_n}{a_n}$ converge en loi vers $\mathcal{N}(0,1)$.*

Pour démontrer ce théorème, on va utiliser le résultat suivant, du à Brown [30]. (voir aussi [46, Théorème 5.8] pour la formulation pour les martingales inversées)

Théorème 1.4.4. *Soit $(U_n, \mathcal{F}_n)_{n\geq 0}$ une suite de différences de martingales inversées, de carré intégrable. Pour $n \geq 0$, on pose*

$$T_n = \sum_{k=0}^{n-1} U_k, \;\; \sigma_n^2 = \sum_{k=0}^{n-1} \mathbb{E}(U_k^2), \;\text{ et }\; V_n = \sum_{k=0}^{n-1} \mathbb{E}(U_k^2 | \mathcal{F}_{k+1}).$$

On suppose que les deux conditions suivantes sont vérifiées :

(i) la suite de variables aléatoires $(\frac{V_n}{\sigma_n^2})$ converge en probabilité vers 1 ;

(ii) pour tout $\epsilon > 0$, $\lim_{n\to\infty} \frac{1}{\sigma_n^2} \sum_{k=0}^{n-1} \mathbb{E}(U_k^2 \mathbb{1}_{|U_k| > \epsilon\sigma_n}) = 0$.

Alors

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}} \left| \mathbb{P}(\frac{T_n}{\sigma_n} \leq t) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{s^2}{2}} ds \right| = 0.$$

Démonstration du théorème 1.4.3. On applique le théorème de Brown, dont on conserve les mêmes notations, avec $U_n = \chi_n \circ \tau^n$ définie sur $(\Omega^\star, \mu_c)$. On remarque que $|a_n^2 - \sigma_n^2| \leq \|S_n - T_n\|_2 = \|(\psi_n)_\pi\|_2$ est borné, donc (a_n^2) est soit bornée, soit divergente vers $+\infty$.
On suppose que a_n^2, et donc σ_n^2 tend vers $+\infty$ et on vérifie la condition (ii) : on a

$$\frac{1}{\sigma_n^2} \sum_{k=0}^{n-1} \mathbb{E}(U_k^2 \mathbb{1}_{|U_k| > \epsilon\sigma_n}) \leq \frac{1}{\sigma_n^2} \sum_{k=0}^{n-1} \|U_k^2\|_{L^\infty_{\mu_c}} \mu_c(|U_k| > \epsilon\sigma_n) \leq \frac{C}{\sigma_n^2} \sum_{k=0}^{n-1} \frac{\mathbb{E}_{\mu_c}(U_k^2)}{\epsilon^2 \sigma_n^2} = \frac{C}{\epsilon^2 \sigma_n^2},$$

qui converge vers 0 pour tout $\epsilon > 0$.
Pour la condition (i), on a

$$V_n = \sum_{k=0}^{n-1} \mathbb{E}(U_k^2|\mathcal{F}_{k+1}) = \sum_{k=0}^{n-1} \mathbb{E}_{\mu_c}(\chi_k^2 \circ \tau^k|\tau^{-(k+1)}\mathcal{F}) = \sum_{k=0}^{n-1} \tilde{P}^{k+1}(\chi_k^2 \circ \tau^k) \circ \tau^{k+1}$$
$$= \sum_{k=0}^{n-1} \tilde{P}(\chi_k^2) \circ \tau^{k+1}.$$

Un calcul simple utilisant les propriétés de l'opérateur de transfert donne

$$\tilde{P}(\chi_k^2) = \left[P_\mu(\widetilde{\phi}_k^2) + \psi_{n+1}^2 + P_\mu(\psi_n^2) - 2P_\mu(\widetilde{\phi}_k)\psi_{n+1} + 2P_\mu(\widetilde{\phi}_k\psi_k) - 2P_\mu(\psi_k)\psi_{k+1}\right]_\pi$$
$$=: (\varphi_k)_\pi.$$

Comme $\widetilde{\phi}_k$ et ψ_k sont bornées dans $\mathcal{B}$, on obtient par nos hypothèses sur $\mathcal{B}$ que φ_k est une suite de fonctions positives bornées dans $\mathcal{B}_0$. Comme $\sum_{k=0}^{n-1} \int \varphi_k \, d\mu = \sigma_n^2$ diverge, on obtient par le théorème 1.4.2 que $\frac{V_n}{\sigma_n^2} \circ \Phi$ converge $\tilde{\mathbb{P}} \otimes \mu$-p.s. vers 1, d'où la validité de la condition (i). Ceci démontre la deuxième partie du théorème 1.4.3.
Si a_n^2, et donc σ_n^2, est bornée, alors d'après le théorème de convergence des martingales, on a que T_n est convergente p.s., d'où l'on obtient que $S_n = T_n + (\psi_n)_\pi \circ \tau^n$ est bornée presque surement. □

1.4.3 Application au problème des cibles rétrécissantes

Dans cette sous-section, on se restreindra au cas où le système aléatoire est un des systèmes décrits aux paragraphes 1.1.2, 1.1.3 et 1.1.4, à l'exception de l'exemple 1.1.11. En particulier, X est soit l'intervalle $[0, 1]$, soit un compact régulier de $\mathbb{R}^d$, et $\mathcal{B} = \mathcal{B}_0$ est l'espace BV en dimension 1, ou l'espace Quasi-Hölder. Ces deux espaces vérifient les conditions requises dans les sous sections précédentes, tandis que la condition de recouvrement remplie par les systèmes considérés entraine que $\frac{1}{h} \in \mathcal{B}$ et l'hypothèse de décroissance des corrélations contre $L^1(\mu)$.
On fixe un point $p \in X$, et on considère une suite décroissantes de boules $B_n(p)$ centrées en p telles que $E_n = \sum_{k=0}^{n-1} \mu(B_k(p)) \to +\infty$. Une application directe du théorème 1.4.2 avec $\phi_k = \mathbb{1}_{B_k(p)}$, qui définit bien une suite bornée dans $\mathcal{B}_0$, donne :

Proposition 1.4.5. *Presque toutes les orbites aléatoires $T_{\underline{\omega}}^k x$ passent infiniment souvent dans $B_k(p)$: pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$ et μ-p.t. $x \in X$, on a $T_{\underline{\omega}}^k x \in B_k(p)$ pour une infinité de $k \geq 0$. On a de plus une information quantitative sur le nombre de visites :*

$$\frac{1}{E_n} \sharp\{k \geq 0 \,/\, T_{\underline{\omega}}^k x \in B_k(p)\} \to 1$$

pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$ et μ-p.t. $x \in X$.

On pose $a_n^2 = \mathbb{E}_{\mu_c}\left(\left(\sum_{k=0}^{n-1}(\widetilde{\phi}_k)_\pi \circ \tau^k\right)^2\right)$. On a vu au théorème 1.4.3 que si a_n^2 est non bornée, alors la somme $\sum_{k=0}^{n-1}(\widetilde{\phi}_k)_\pi \circ \tau^k$ satisfait un théorème de la limite centrale. On donne ici un critère simple permettant de vérifier cette condition :

Proposition 1.4.6. *Si* $(\log k)\mu(B_k(p)) \to 0$, *alors* $\liminf \frac{a_n^2}{E_n} \geq 1$. *En particulier,* $a_n^2 \to \infty$.

Démonstration. On choisit $a > 0$ suffisamment grand pour que $a \log \lambda < -1$, où $\lambda < 1$ est le taux de décroissance exponentielle pour les corrélations.
On a

$$\begin{aligned} a_n^2 = \mathbb{E}_{\mu_c}\left(\left(\sum_{k=0}^{n-1}(\widetilde{\phi}_k)_\pi \circ \tau^k\right)^2\right) &= \sum_{k=0}^{n-1}\mathbb{E}_{\mu_c}((\widetilde{\phi}_k)_\pi^2) + 2\sum_{0\leq k<l\leq n-1}\mathbb{E}_{\mu_c}((\widetilde{\phi}_k)_\pi(\widetilde{\phi}_l)_\pi \circ \tau^{l-k}) \\ &= \sum_{k=0}^{n-1}\mathbb{E}_\mu(\widetilde{\phi}_k^2) + 2\sum_{0\leq k<l\leq n-1}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l). \end{aligned}$$

On découpe le deuxième terme de cette somme en deux parties :

$$\sum_{0\leq k<l\leq n-1}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l) = \sum_{k=0}^{n-1}\left[\sum_{l=k+1}^{k+a\log(k+1)}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l) + \sum_{l=k+a\log(k+1)+1}^{n-1}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l)\right].$$

On a

$$\sum_{l=k+1}^{k+a\log(k+1)}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l) = \sum_{l=k+1}^{k+a\log(k+1)}\left(\mathbb{E}_\mu(\phi_k U^{l-k}\phi_l) - \int \phi_k d\mu \int \phi_l d\mu\right).$$

Par positivité et décroissance des ϕ_k, on obtient que

$$\sum_{l=k+1}^{k+a\log(k+1)}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l) \geq -\mathcal{O}\left((\log k)\mu(B_k(p))^2\right).$$

D'autre part, un calcul direct en utilisant la décroissance des corrélations donne

$$\left|\sum_{l=k+a\log(k+1)+1}^{n-1}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l)\right| = \mathcal{O}\left(k^{a\log\lambda}\right).$$

Ainsi, on obtient

$$\sum_{0\leq k<l\leq n-1}\mathbb{E}_\mu(\widetilde{\phi}_k U^{l-k}\widetilde{\phi}_l) \geq -\mathcal{O}(1) - \mathcal{O}\left(\sum_{k=0}^{n-1}(\log k)\mu(B_k(p))^2\right)$$

Comme $(\log k)\mu(B_k(p))$ tend vers 0 et $E_n = \sum_{k=0}^{n-1}\mu(B_k(p))$ est divergente, on trouve que $\mathcal{O}\left(\sum_{k=0}^{n-1}(\log k)\mu(B_k(p))^2\right) = o(E_n)$. Or,

$$\sum_{k=0}^{n-1}\mathbb{E}_\mu(\widetilde{\phi}_k^2) = \sum_{k=0}^{n-1}\left(\mu(B_k(p)) - \mu(B_k(p))^2\right) \sim E_n$$

car $\mu(B_k(p)) \to 0$ et $E_n \to \infty$, d'où l'on déduit que

$$a_n^2 \geq E_n + o(E_n),$$

ce qui termine la preuve. $\square$

1.5 Version quenched du théorème de la limite centrale

Dans cette section, on s'intéresse à la version quenched du TCL, c'est à dire à un TCL valable pour presque tout choix de $\underline{\omega}$. Plus précisément, on étudie les lois limites pour la somme $\sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$, pour $\underline{\omega}$ fixé, mais générique.

1.5.1 Un résultat général

On décrit ici un résultat abstrait, démontré par Ayyer, Liverani et Stenlund [11], qui sera notre outil principal pour étudier le TCL quenched. On considère ici un système dynamique aléatoire i.i.d. $\{T_\omega\}_{\omega\in\Omega}$ agissant sur X, avec une mesure stationnaire μ. Soit $\varphi : X \to \mathbb{R}$ un observable avec $\int_X \varphi \, d\mu = 0$, et on définit comme auparavant $X_k(\underline{\omega}, x) = \varphi(T_{\underline{\omega}}^k x)$ et $S_n = \sum_{k=0}^{n-1} X_k$. On aura besoin d'introduire un système auxiliaire, défini de la façon suivante. L'espace probabilisé est toujours $(\Omega, \mathbb{P})$. Les applications, que l'on dénotera par $\hat{T}_\omega$, agissent sur X^2 par $\hat{T}_\omega(x, y) = (T_\omega x, T_\omega y)$. On définit un nouvel observable $\hat{\varphi} : X^2 \to \mathbb{R}$ pour ce système auxiliaire par $\hat{\varphi}(x, y) = \varphi(x) - \varphi(y)$, et on dénote par $\hat{S}_n$ sa somme de Birkhoff associée.

Théorème 1.5.1. *On suppose qu'il existe $\sigma^2 > 0$ et $C > 0$ tels que pour tout $t \in \mathbb{R}$ et $n \geq 1$ avec $\frac{t}{\sqrt{n}}$ assez petit,*

(1) $\left| \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n}) - e^{-\frac{t^2\sigma^2}{2}} \right| \leq C\frac{1+|t|^3}{\sqrt{n}}$,

(2) $\left| \mathbb{E}_{\tilde{\mathbb{P}}\otimes(\mu\otimes\mu)}(e^{i\frac{t}{\sqrt{n}}\hat{S}_n}) - e^{-t^2\sigma^2} \right| \leq C\frac{1+|t|^3}{\sqrt{n}}$.

On suppose de plus que pour $n \geq 1$, et tout $\epsilon > 0$ assez petit,

(3) $\tilde{\mathbb{P}} \otimes \mu \left(\left| \frac{S_n}{n} \right| \geq \epsilon \right) \leq Ce^{-C\epsilon^2 n}$.

Alors, le théorème de la limite centrale quenched est vérifié : pour $\tilde{\mathbb{P}}$-presque toute suite $\underline{\omega} \in \Omega^{\mathbb{N}}$, on a

$$\frac{\sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k}{\sqrt{n}} \Longrightarrow_\mu \mathcal{N}(0, \sigma^2).$$

On peut vérifier les hypothèses de ce théorème en procédant de la façon suivante.

Étape 1 On démontre que le système aléatoire qui nous intéresse, c'est-à-dire $\{T_\omega\}_{\omega\in\Omega}$, vérifie les hypothèses (H1) à (H6) décrites au paragraphe 1.1.1, avec un espace fonctionnel $\mathcal{B}$. On vérifie que φ appartient à $\mathcal{B}_0$, où $\mathcal{B}_0$ satisfait la condition donnée au début du paragraphe 1.2. Sous ces hypothèses, on sait que le système aléatoire $\{T_\omega\}_{\omega\in\Omega}$ admet une unique mesure stationnaire μ absolument continue par rapport à m. En faisant l'hypothèse que φ n'est pas un cobord au sens de la proposition 1.2.5, par les résultats du paragraphe 1.2, on sait qu'il existe $\sigma^2 > 0$ tel que les hypothèses 1 et 3 du théorème 1.5.1 sont vérifiées.

Étape 2 On démontre que le système auxiliaire $\{\hat{T}_\omega\}_{\omega\in\Omega}$ vérifie les hypothèses (H1) à (H6), avec un espace fonctionnel $\hat{\mathcal{B}}$ de fonctions définies sur X^2, avec comme mesure de référence $m_2 = m \otimes m$. Ainsi, ce système admet alors une unique mesure stationnaire μ_2 absolument continue par rapport à m_2. L'observable $\hat{\varphi}$ doit appartenir à un espace $\hat{\mathcal{B}}_0$ qui vérifie la condition de compatibilité avec $\hat{\mathcal{B}}$ du paragraphe 1.2, et la mesure $\mu \otimes \mu$ doit appartenir à $\mathcal{M}_{\hat{\mathcal{B}}}$, c'est-à-dire $h \otimes h \in \hat{\mathcal{B}}$. Il est facile de voir que les marginales de μ_2 sont absolument continues par rapport à m et sont stationnaires pour le système $\{T_\omega\}_{\omega\in\Omega}$. Comme μ est l'unique mesure stationnaire absolument continue pour le système considéré, on en déduit que les marginales de μ_2 sont toutes deux égales à μ, et donc $\int_{X^2} \hat{\varphi}\, d\mu_2 = 0$. Sous ces hypothèses, on sait qu'il existe $\hat{\sigma}^2 > 0$ (la stricte positivité découle en fait de l'étape 3) tel que $\left|\mathbb{E}_{\tilde{\mathbb{P}}\otimes(\mu\otimes\mu)}(e^{i\frac{t}{\sqrt{n}}\hat{S}_n}) - e^{-\frac{t^2\hat{\sigma}^2}{2}}\right| \leq C\frac{1+|t|^3}{\sqrt{n}}$ pour $t \in \mathbb{R}$ et $n \geq 1$ avec $\frac{t}{\sqrt{n}}$ assez petit.

Étape 3 On démontre que $\hat{\sigma}^2 = 2\sigma^2$. Ceci entraine la validité de l'hypothèse 2 du théorème 1.5.1.

Les systèmes aléatoires que l'on considère dans tout ce chapitre vérifient les hypothèses (H1) à (H6), donc l'étape 1 est ici inexistante. L'étape 2 ne présente en général que des difficultés d'ordre technique : le système aléatoire auxiliaire agit sur un espace de dimension deux fois plus grande que celle du système d'origine, ce qui nécessite d'utiliser un espace fonctionnel $\hat{\mathcal{B}}$ plus compliqué que $\mathcal{B}$. Ceci est particulièrement vrai lorsque $X = [0, 1]$: en effet, dans ce cas, $\mathcal{B}$ est l'espace BV en dimension 1, tandis que l'on devra choisir pour $\hat{\mathcal{B}}$ l'espace Quasi-Hölder (ou n'importe quelle autre alternative comme BV en dimension quelconque). On devra alors tenir compte de l'accumulation des discontinuités, un phénomène qui n'apparait qu'en dimension supérieure ou égale à 2. Néanmoins, la structure particulièrement simple des applications $\hat{T}_\omega$ (des produits directs) et des partitions de régularité (composées de rectangles) entraine que les hypothèses classiques où la dilatation l'emporte sur la complexité sont souvent vraies. Toute la difficulté réside en réalité dans l'étape 3. Pour démontrer que $\hat{\sigma}^2 = 2\sigma^2$, on utilisera les formules de Green-Kubo qui nous donnent une expression explicite pour les variances asymptotiques.

On démontrera dans les sous-sections suivantes que cette hypothèses sur les variances est vérifiée dans le cas où toutes les applications T_ω préserve la mesure μ. On discutera aussi du cas général (et ouvert) où la mesure μ n'est pas nécessairement préservée par toutes les

applications. Avant de terminer cette sous-section, on démontre le lemme suivant, qui montre que bien qu'elle soit surprenante, la condition $\hat{\sigma}^2 = 2\sigma^2$ est nécessaire pour obtenir le TCL quenched tel que nous l'avons formulé.

Lemme 1.5.2. *En conservant les notations précédentes, si on suppose qu'il existe $\sigma^2 > 0$ et $\hat{\sigma}^2 > 0$ tel que*

1. $\frac{S_n}{\sqrt{n}}$ *converge en loi vers $\mathcal{N}(0, \sigma^2)$ sous la probabilité $\tilde{\mathbb{P}} \otimes \mu$,*
2. $\frac{\hat{S}_n}{\sqrt{n}}$ *converge en loi vers $\mathcal{N}(0, \hat{\sigma}^2)$ sous la probabilité $\tilde{\mathbb{P}} \otimes (\mu \otimes \mu)$,*
3. *pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$, $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$ converge en loi vers $\mathcal{N}(0, \sigma^2)$ sous la probabilité μ.*

Alors $\hat{\sigma}^2 = 2\sigma^2$.

Démonstration. On pose $S_{n,\underline{\omega}} = S_n(\underline{\omega}, .) = \sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$. Procédant comme dans [11], on écrit pour $t \in \mathbb{R}$ et $n \geq 1$:

$$\begin{aligned}
\mathbb{E}_{\tilde{\mathbb{P}}}\left(\left|\mu(e^{i\frac{t}{\sqrt{n}}S_{n,\underline{\omega}}}) - e^{-\frac{\sigma^2 t^2}{2}}\right|^2\right) &= \mathbb{E}_{\tilde{\mathbb{P}}}\left(\left|\mu(e^{i\frac{t}{\sqrt{n}}S_n})\right|^2\right) - e^{-t^2\sigma^2} + 2e^{-\frac{\sigma^2 t^2}{2}}\Re\left(e^{-\frac{\sigma^2 t^2}{2}} - \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n})\right) \\
&= \mathbb{E}_{\tilde{\mathbb{P}}}\left(\mu \otimes \mu(e^{i\frac{t}{\sqrt{n}}\hat{S}_n})\right) - e^{-\frac{t^2\hat{\sigma}^2}{2}} + \left(e^{-\frac{t^2\hat{\sigma}^2}{2}} - e^{-t^2\sigma^2}\right) + 2e^{-\frac{\sigma^2 t^2}{2}}\Re\left(e^{-\frac{\sigma^2 t^2}{2}} - \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(e^{i\frac{t}{\sqrt{n}}S_n})\right).
\end{aligned}$$

Par les deux premières hypothèses, ce terme converge vers $e^{-\frac{t^2\hat{\sigma}^2}{2}} - e^{-t^2\sigma^2}$ lorsque $n \to \infty$. Mais $\mathbb{E}_{\tilde{\mathbb{P}}}\left(\left|\mu(e^{i\frac{t}{\sqrt{n}}S_{n,\underline{\omega}}}) - e^{-\frac{\sigma^2 t^2}{2}}\right|^2\right)$ converge vers 0 par la troisième hypothèse et le théorème de convergence dominée. Ceci montre que $e^{-\frac{t^2\hat{\sigma}^2}{2}} = e^{-t^2\sigma^2}$ et donc $\hat{\sigma}^2 = 2\sigma^2$. □

1.5.2 Le cas où la mesure μ est préservée par toutes les applications T_ω

On suppose ici que toutes les applications T_ω préservent une même mesure μ. Dans ce cas, la mesure μ est clairement stationaire, tandis que chaque $\hat{T}_\omega$ préserve $\mu \otimes \mu$, et donc $\mu \otimes \mu$ est stationaire pour le système auxiliaire.
On suppose que les deux premières étapes décrites au paragraphe 1.5.1 ont été vérifiées. On a donc que l'opérateur P du système aléatoire admet un trou spectral sur $\mathcal{B}$, et l'observable $\varphi \in \mathcal{B}_0$ avec $\int \varphi \, d\mu = 0$ satisfait tous les théorèmes limites décrits à la section 1.2 pour l'unique mesure stationnaire a.c. μ. De même, l'opérateur $\hat{P}$ du système auxiliaire admet un trou spectral sur $\hat{\mathcal{B}}$ et l'observable $\hat{\varphi} \in \hat{\mathcal{B}}_0$ satisfait les mêmes théorèmes limites pour l'unique mesure stationaire a.c. $\mu_2 = \mu \otimes \mu$.

Proposition 1.5.3. *On a $\hat{\sigma}^2 = 2\sigma$. En particulier, le TCL quenched est vérifié pour l'observable φ, s'il n'est pas un cobord aléatoire : pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$, $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} \varphi \circ T_{\underline{\omega}}^k$ converge en loi vers $\mathcal{N}(0, \sigma^2)$ sous la probabilité μ.*

Démonstration. D'après la proposition 1.2.3, on a

$$\hat{\sigma}^2 = \int_{X^2} \hat{\varphi}^2 \, d\mu_2 + 2 \sum_{n=1}^{\infty} \int_{X^2} \hat{\varphi} \, \hat{U}^n \hat{\varphi} \, d\mu_2,$$

où $\hat{U}$ est l'opérateur de Koopman en moyenne pour le système auxiliaire : $\hat{U} f(x,y) = \int_\Omega f(\hat{T}_\omega(x,y)) \, d\mathbb{P}(\omega)$.
On a

$$\begin{aligned}
\int_{X^2} \hat{\varphi}^2 \, d\mu_2 &= \int_X \int_X (\varphi(x) - \varphi(y))^2 d\mu(x) d\mu(y) \\
&= \int_X \int_X \varphi(x)^2 d\mu(x) d\mu(y) - 2 \int_X \int_X \varphi(x)\varphi(y) d\mu(x) d\mu(y) \\
&+ \int_X \int_X \varphi(y)^2 d\mu(x) d\mu(y) \\
&= 2 \int_X \varphi^2 d\mu - 2 \left(\int_X \varphi \, d\mu \right)^2 = 2 \int_X \varphi^2 \, d\mu.
\end{aligned}$$

D'autre part, $\hat{U}\varphi(x,y) = \int_\Omega (\varphi(T_\omega x) - \varphi(T_\omega y)) d\mathbb{P}(\omega) = U\varphi(x) - U\varphi(y)$, et donc

$$\begin{aligned}
\int_{X^2} \hat{\varphi} \, \hat{U} \hat{\varphi} \, d\mu_2 &= \int_X \int_X (\varphi(x) - \varphi(y))(U\varphi(x) - U\varphi(y)) d\mu(x) d\mu(y) \\
&= \int_X \int_X \varphi(x) U\varphi(x) d\mu(x) d\mu(y) - \int_X \int_X \varphi(x) U\varphi(y) d\mu(x) d\mu(y) \\
&- \int_X \int_X \varphi(y) U\varphi(x) d\mu(x) d\mu(y) + \int_X \int_X \varphi(y) U\varphi(y) d\mu(x) d\mu(y) \\
&= 2 \int_X \varphi \, U\varphi \, d\mu - 2 \left(\int_X \varphi \, d\mu \right) \left(\int_X U\varphi \, d\mu \right) = 2 \int_X \varphi \, U\varphi \, d\mu
\end{aligned}$$

On obtient donc $\hat{\sigma}^2 = 2 \int_X \varphi^2 \, d\mu + 4 \sum_{n=1}^{\infty} \int_X \varphi \, U^n \varphi \, d\mu = 2\sigma^2$ par la proposition 1.2.3. □

Sous les mêmes hypothèses, c'est-à-dire après avoir vérifié les deux premières étapes dans le cas où la mesure μ est préservée par chaque application T_ω, on peut adapter la preuve du théorème 4.1 de Furman et Shalom [58] et obtenir une décroissance quenched des corrélations. On supposera que pour tout $\psi \in \mathcal{B}_0$, on a $\psi \otimes \psi \in \hat{\mathcal{B}}_0$, ce qui est le cas lorsque $\mathcal{B}_0 = \mathrm{BV}([0,1])$ et $\hat{\mathcal{B}}_0 = V_1([0,1]^2)$, ou $\mathcal{B}_0 = V_\alpha(X)$ et $\hat{\mathcal{B}}_0 = V_\alpha(X^2)$.

Théorème 1.5.4. *Il existe $0 < \kappa < 1$ tel que pour tout $\phi \in L^\infty(m)$ et tout $\psi \in \mathcal{B}_0$, on a*

$$\left| \int_X \phi(T_{\underline{\omega}}^n x) \psi(x) \, d\mu - \int_X \phi \, d\mu \int_X \psi \, d\mu \right| \leq C_{\underline{\omega},\phi,\psi} \kappa^n,$$

pour $\tilde{\mathbb{P}}$-presque tout $\underline{\omega}$.

Remarque 1.5.5. IL est à noter que ce résultat est plus faible que le type de résultat obtenu par Stenlund et Sulku [135] dans le cas d'applications uniformément dilatantes et lisses du cercle, car la dépendance de la constante $C_{\underline{\omega},\phi,\psi}$ vis-à-vis des normes de ϕ et ψ n'est pas précisée, et surtout car l'ensemble exceptionnel des $\underline{\omega}$ pour lesquels l'inégalité n'est pas vérifiée dépend a priori de ϕ et ψ.

Démonstration. On peut supposer sans perdre de généralité que $\int_X \phi \, d\mu = 0$. On a

$$\begin{aligned}
\int_{\tilde{\Omega}} \left| \int_X \phi(T^n_{\underline{\omega}} x)\psi(x)\, d\mu \right| d\tilde{\mathbb{P}}(\underline{\omega}) &\leq \left[\int_{\tilde{\Omega}} \left(\int_X \phi(T^n_{\underline{\omega}} x)\psi(x)\, d\mu \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) \right]^{\frac{1}{2}} \\
&= \left[\int_{\tilde{\Omega}} \int_X \int_X \phi(T^n_{\underline{\omega}} x)\phi(T^n_{\underline{\omega}} y)\psi(x)\psi(y) d\mu(x) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \right]^{\frac{1}{2}} \\
&= \left[\int_{\tilde{\Omega}} \int_{X^2} \phi \otimes \phi(\hat{T}_{\underline{\omega}}(x,y))\psi \otimes \psi(x,y) d(\mu \otimes \mu)(x,y) d\tilde{\mathbb{P}}(\underline{\omega}) \right]^{\frac{1}{2}} \\
&= \left[\int_{X^2} \hat{U}^n(\phi \otimes \phi)\psi \otimes \psi d(\mu \otimes \mu) \right]^{\frac{1}{2}} \\
&\leq C^{\frac{1}{2}} (\lambda^{\frac{1}{2}})^n \|\phi \otimes \phi\|^{\frac{1}{2}}_{L^\infty_{m \otimes m}} \|\psi \otimes \psi\|^{\frac{1}{2}}_{\hat{\mathcal{B}}_0},
\end{aligned}$$

d'après le point 1 de la proposition 1.2.1, car $\mu \otimes \mu$ est la mesure stationnaire du système auxiliaire.
Pour $\lambda^{\frac{1}{2}} < \kappa < 1$, on a par conséquent que la fonction

$$S(\underline{\omega}) = \sum_{n=1}^{\infty} \frac{1}{\kappa^n} \left| \int_X \phi(T^n_{\underline{\omega}} x)\psi(x)\, d\mu \right|$$

est d'espérance finie. Ainsi, $S(\underline{\omega}) < \infty$ pour presque tout $\underline{\omega}$, ce qui entraine que le terme général de la série est borné pour presque tout $\underline{\omega}$, et termine la preuve. □

Remarque 1.5.6. En utilisant au cours de la preuve le point 2 de la proposition 1.2.1 plutôt que le point 1, on peut énoncer le théorème 1.5.4 pour des observables $\phi \in L^1(m)$ si on suppose que $\hat{\mathcal{B}}$ s'injecte continument dans $L^\infty(m \otimes m)$, ce qui est le cas là encore de BV ou Quasi-Hölder.

On décrit à présent deux exemples. Pour simplifier, on ne considèrere que le cas où l'espace Ω est fini, $X = [0,1]$ et toutes les applications T_ω préservent la mesure de Lebesgue m.

Exemple 1.5.7. On suppose que toutes les applications T_ω sont données par $T_\omega x = \beta_\omega x \mathrm{mod} 1$, où $\beta_\omega > 1$ est un entier. L'opérateur de transfert en moyenne P satisfait clairement une inégalité de Lasota-Yorke sur $\mathcal{B} = \mathrm{BV}$ et le système possède la propriété du recouvrement, donc l'étape 1 est vérifiée pour les observables $\varphi \in \mathrm{BV}$. D'autre part, l'opérateur de transfert du système auxiliaire sur le carré $X^2 = [0,1]^2$ admet aussi un trou spectral sur l'espace Quasi-Hölder $\hat{\mathcal{B}} = V_1(X^2)$. Comme $\hat{\varphi} \in V_1(X^2)$ pour tout $\varphi \in \mathrm{BV}$, on obtient l'étape 2. D'après la proposition 1.5.3, le quenched CLT est vrai pour les observables $\varphi \in \mathrm{BV}$ qui ne sont pas des cobords aléatoires.

Exemple 1.5.8. Il existe des applications non linéaires qui préservent la mesure de Lebesgue. De tels exemples sont données par les applications de type Lorenz étudiées dans [47] : ces applications possèdent à la fois un point fixe neutre et un point où la dérivée devient infinie.

C'est la coexistence de ces deux phénomènes qui entraine que la mesure de Lebesgue peut etre préservée par ces applications. Dans cet exemple, on considère le cas où $\Omega = \{0, 1\}$, T_0 est donnée par $T_0 x = 2x \bmod 1$ et T_1 est l'une des applications considérées dans [47], dont on rappelle la définition : pour $\gamma > 0$, l'application T_1 est définie implicitement sur l'intervalle $[-1, 1]$ par les équations

$$x = \begin{cases} \dfrac{1}{2\gamma}(1 + T_1(x))^\gamma & \text{si } 0 \leq x \leq \dfrac{1}{2\gamma} \\ T_1(x) + \dfrac{1}{2\gamma}(1 - T_1(x))^\gamma & \text{si } \dfrac{1}{2\gamma} < x \leq 1. \end{cases}$$

Pour les valeurs négatives de x, on pose $T(-x) = -T(x)$. Par homothétie-translation, on ramène T_1 en une application définie de $[0, 1]$ dans lui-même.
On va montrer qu'il existe $0 \leq p^\star < 1$ tel que si T_0 est itéré avec probabilité $p > p^\star$, alors le quenched CLT est vrai pour les observables φ Lipschitz qui ne sont pas des cobords aléatoires. Comme l'opérateur de transfert en moyenne peut s'écrire $P = pP_0 + (1-p)P_1$, où P_0, resp. P_1, est l'opérateur de transfert de T_0, resp. T_1 (et de même, $\hat{P} = p\hat{P}_0 + (1-p)\hat{P}_1$ pour le système auxiliaire), il suffit de prouver que P_0 et $\hat{P}_0$ admettent un trou spectral sur des espaces de Banach $\mathcal{B}$ et $\hat{\mathcal{B}}$, que P_1 et $\hat{P}_1$ agissent continument sur ces espaces et que $\varphi \in \mathcal{B}_0$ et $\hat{\varphi} \in \hat{\mathcal{B}}_0$. On utilisera les espaces Quasi-Hölder, avec $\mathcal{B} = \mathcal{B}_0 = V_\alpha(X)$ et $\hat{\mathcal{B}} = \hat{\mathcal{B}} = V_\alpha(X^2)$ pour un choix convenable de α. Clairement, les opérateurs de transfert de T_0 et $\hat{T}_0$ admettent un trou spectral sur ces espaces, et φ (resp. $\hat{\varphi}$) appartient à $\mathcal{B}$ (resp. $\hat{\mathcal{B}}$) lorsque φ est Lipschitz. Pour démontrer que P_1 et $\hat{P}_1$ agissent continument, on utilisera le résultat général suivant :

Proposition 1.5.9. *Soit M un compact de $\mathbb{R}^d$ avec $m_d(M) = 1$, où m_d est la mesure de Lebesgue sur $\mathbb{R}^d$. Soit $T : M \to M$ une application non-singulière. On définit $g(x) = \frac{1}{|\det DT(x)|}$, et on suppose qu'il existe une famille finie d'ouverts disjoints $\{U_i\}_i$ inclus dans M, une constante $C > 0$ et $0 < \alpha \leq 1$ tels que*

1. *$m_d(\cup_i U_i) = 1$,*
2. *$T : U_i \to TU_i$ est un C^1-difféomorphisme,*
3. *$d(Tx, Ty) \geq d(x, y)$ pour tout i et tout $x, y \in U_i$,*
4. *$|g(x) - g(y)| \leq Cd(x, y)^\alpha$, pour tout i et tout $x, y \in U_i$,*
5. *$m_d(B_\epsilon(\partial TU_i)) \leq C\epsilon^\alpha$ pour tout i et tout $\epsilon > 0$.*

Alors l'opérateur de transfert de T agit continument sur $V_\alpha(M)$.

Pour un paramètre $\gamma > 1$, les applications T_1 et $\hat{T}_1$ satisfont ces hypothèses pour $\alpha = \min\{1, \gamma - 1\}$. Le quenched CLT est donc valable pour des observables φ Lipschitz qui ne sont pas des cobords aléatoires dès que $p^\star$ est suffisamment proche de 1.

Démonstration de la proposition 1.5.9. On note par $T_i^{-1} : TU_i \to U_i$ la branche inverse de T restreinte à U_i.

L'opérateur de transfert P de T est donné par

$$Pf(x) = \sum_i (gf)(T_i^{-1}x)\mathbb{1}_{TU_i}(x).$$

Suivant Saussol [129], on a pour tout $\epsilon > 0$ et $x \in \mathbb{R}^d$:

$$\operatorname{osc}(Pf, B_\epsilon(x)) \leq \sum_i R_i^{(1)}(x)\mathbb{1}_{TU_i}(x) + 2\sum_i R_i^{(2)}(x),$$

où $R_i^{(1)}(x) = \operatorname{osc}(gf, T^{-1}B_\epsilon(x) \cap U_i)$ et $R_i^{(2)}(x) = \left(\operatorname*{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} |gf|\right)\mathbb{1}_{B_\epsilon(\partial TU_i)}(x)$.
En utilisant la proposition 3.2 (iii) de [129], on obtient

$$R_i^{(1)}(x) \leq \operatorname{osc}(f, T^{-1}B_\epsilon(x) \cap U_i) \operatorname*{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} g + \operatorname{osc}(g, T^{-1}B_\epsilon(x) \cap U_i) \operatorname*{ess\,inf}_{T^{-1}B_\epsilon(x)\cap U_i} |f|.$$

Par l'hypothèse (3), on a $T^{-1}B_\epsilon(x) \cap U_i \subset B_\epsilon(T_i^{-1}x)$, tandis que l'hypothèse 4 donne $\operatorname{osc}(g, T^{-1}B_\epsilon(x) \cap U_i) \leq C\epsilon^\alpha$ et $\operatorname*{ess\,sup}_{T^{-1}B_\epsilon(x)\cap U_i} g \leq g(T_i^{-1}x) + C\epsilon^\alpha$.
Cela montre que $R_i^{(1)}(x) \leq g(T_i^{-1}x)\operatorname{osc}(f, B_\epsilon(T_i^{-1}x)) + C\epsilon^\alpha\|f\|_{\sup}$, d'où

$$\int \sum_i R_i^{(1)}(x)\mathbb{1}_{TU_i}(x)dx \leq \int P(\operatorname{osc}(f, B_\epsilon(.))(x)dx + C\|f\|_\infty\epsilon^\alpha \sum_i m_d(TU_i).$$

Comme la somme est finie, on a $\int \sum_i R_i^{(1)}(x)\mathbb{1}_{TU_i}(x)dx \leq \epsilon^\alpha\left(|f|_\alpha + C\|f\|_{\sup}\right) \leq C\epsilon^\alpha\|f\|_\alpha$.
On estime à présent $R_i^{(2)}$: on a $R_i^{(2)}(x) \leq \|g\|_{\sup}\|f\|_{\sup}\mathbb{1}_{B_\epsilon(\partial TU_i)}(x)$, donc par l'hypothèse (5), $\int \sum_i R_i^{(2)}dx \leq C\|f\|_{\sup}\sum_i m_d(B_\epsilon(\partial TU_i)) \leq C\epsilon^\alpha\|f\|_{\sup}$. Cela prouve que $|Pf|_\alpha \leq C\|f\|_\alpha$ et termine la démonstration. □

1.5.3 Le cas où la mesure μ n'est pas nécessairement préservée par les applications T_ω

Soit Ω un ensemble fini, $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$ un vecteur de probabilité avec $p_\omega > 0$ pour chaque $\omega \in \Omega$, et $T = \{T_\omega\}_{\omega\in\Omega}$ une famille finie d'applications de Lasota-Yorke sur l'intervalle $[0, 1]$. On suppose que $\lambda(T_\omega) > 1$ pour chaque ω, c'est-à-dire que toutes les applications sont uniformément dilatantes. On suppose aussi que le système aléatoire $(\Omega, \mathbb{P}, T)$ possède la propriété de recouvrement. Par les résultats du paragraphe 1.1.2, on sait qu'il existe une unique mesure stationnaire μ absolument continue.
Soit $\varphi \in \mathrm{BV}$ avec $\int_X \varphi\, d\mu = 0$, qui n'est pas un cobord aléatoire au sens de la proposition 1.2.5.

Théorème 1.5.10. *Le quenched CLT est vérifié si et seulement si*

$$\lim_{n\to\infty} \frac{1}{n}\int_{\tilde{\Omega}} \left(\sum_{k=0}^{n-1}\int_X \varphi \circ T_{\underline{\omega}}^k\, d\mu\right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) = 0.$$

Cette condition, bien que difficile à vérifier en pratique, a l'avantage de ne faire intervenir que des quantités intrinsèques au système considéré, et non des quantités associées au système auxiliaire, qui est plus ou moins inconnu. On voit que cette condition est trivialement satisfaite si toutes les applications T_ω préservent μ.

Démonstration. On suit la démarche décrite à la section 1.5.1.

Étape 1. Les résultats du paragraphe 1.1.2 montrent que l'opérateur de transfert P du système $(\Omega, \mathbb{P}, T)$ admet un trou spectral sur $\mathcal{B} = \mathrm{BV}$. On a donc que $\varphi \in \mathcal{B}_0$ satisfait les théorèmes limites du paragraphe 1.2 avec $\sigma^2 > 0$.

Étape 2. On va démontrer que l'opérateur de transfert en moyenne $\hat{P}$ du système auxiliaire admet un trou spectral sur l'espace Quasi-Hölder $\hat{\mathcal{B}} = V_1(X^2)$. Comme $\hat{\varphi} \in \hat{\mathcal{B}}_0 = \hat{\mathcal{B}}$, ceci entraine la validité de l'étape 2. On commence par prouver une inégalité de Lasota-Yorke. On rappelle quelques définitions et résultats de [129]. On ne considère que le cas $\alpha = 1$. Soit $S : M \to M$ une application de classe C^2 par morceaux sur un compact régulier $M \subset \mathbb{R}^d$, avec une partition de régularité et injectivité $\{U_i\}_i$, qui satisfait les hypothèses (PEi), $i = 1, 2, 3, 4$ de [129], avec les constantes $\epsilon_0(S) > 0$, $c(S)$ et $s(S) < 1$ correspondantes. Il est facile de voir que si T est une application de type Lasota-Yorke sur l'intervalle [0,1], alors $S = \hat{T} = T \times T$ satisfait ces hypothèses sur $M = [0, 1]^2$, avec $s(S) = \lambda(T)^{-1}$.
On définit, pour tout $0 \leq \delta \leq \epsilon_0(S)$,

$$G_S(\epsilon, \delta) = \sup_x \sum_i \frac{m_N(S_i^{-1} B_\epsilon(\partial S U_i) \cap B_{(1-s(S))\delta}(x))}{m_N(B_{(1-s(S))\delta}(x))},$$

et $\eta_S(\delta) = s(S) + 2\left(\sup_{0<\epsilon\leq\delta} \frac{G_S(\epsilon,\delta)}{\epsilon}\right)\delta$. On rappelle que $B_\epsilon(A)$ désigne le ϵ-voisinage de A pour la distance euclidienne.
Soit $V_1(M)$ l'espace Quasi-Hölder sur M. On rappelle que la définition de cet espace fait intervenir un paramètre ϵ_0. L'espace $V_1(M)$ ne dépend pas du choix de ϵ_0, mais la semi-norme elle en dépend, bien que deux choix de ϵ_0 donneront deux semi-normes équivalentes. On notera la semi-norme $|.|_{\epsilon_0}$ pour souligner la dépendance en ϵ_0. L'opérateur de transfert P_S de S satisfait l'inégalité de Lasota-Yorke suivante sur $V_1(M)$:

Proposition 1.5.11 ([129, Lemme 4.1])**.** *Sous les conditions précédentes, pour tout* $0 < \epsilon_0 \leq \epsilon_0(S)$ *et* $f \in V_1(M)$*, on a*

$$|P_S f|_{\epsilon_0} \leq (1 + c(S)s(S)\epsilon_0)\eta_S(\epsilon_0)\,|f|_{\epsilon_0} + A(S, \epsilon_0)\|f\|_{L^1_{m_d}},$$

où $A(S, \epsilon_0)$ *est une constante finie qui ne dépend que de* S *et* ϵ_0*.*

On estime la fonction G_S dans le cas où $S = T \times T$, avec T une application de type Lasota-Yorke sur $[0, 1]$. Si $\{I_i\}$ est la partition associée à T, alors la partition de S est donnée par $\{I_i \times I_j\}_{i,j}$.

Lemme 1.5.12. *Soit $T : [0,1] \to [0,1]$ une application de type Lasota-Yorke et $S = T \times T : [0,1]^2 \to [0,1]^2$. Alors il existe $\delta(T) > 0$ tel que pour tout $0 < \epsilon \leq \delta \leq \delta(T)$, on a $G_S(\epsilon, \delta) \leq \frac{64 s(S)}{\pi(1-s(S))} \frac{\epsilon}{\delta}$.*

Démonstration. On désigne par $B_\epsilon^\infty(A)$ le ϵ-voisinage de A pour la distance ℓ^∞. On a $B_\epsilon(A) \subset B_\epsilon^\infty(A)$. Pour tout i et j, on a $SU_{i,j} = TI_i \times TI_j$, qui est un rectangle dont les côtés sont parallèles aux axes de coordonnées. Ainsi, $B_\epsilon(\partial SU_{i,j}) \subset B_\epsilon^\infty(\partial SU_{i,j})$, ce qui donne $S_{i,j}^{-1} B_\epsilon(\partial SU_{i,j}) \subset B_{s(S)\epsilon}^\infty(\partial U_{i,j})$. L'inclusion provient de la structure produit de S. On a alors $m_2(S_{i,j}^{-1} B_\epsilon(\partial SU_{i,j}) \cap B_{(1-s(S))\delta}(x)) \leq m_2(B_{s(S)\epsilon}^\infty(\partial U_{i,j}) \cap B_{(1-s(S))\delta}(x)) \leq 16 s(S)\epsilon(1-s(S))\delta$ [2]. Pour tout δ assez petit, en fonction de T, toute boule $B_{(1-s(S))\delta}(x)$ ne rencontrera qu'au plus 4 éléments $S_{i,j}^{-1} B_\epsilon(\partial SU_{i,j})$, d'où le lemme. □

On peut à présent prouver une inégalité de Lasota-Yorke sur $V_1([0,1]^2)$ pour l'opérateur $\hat{P}$ associé au système auxiliaire $(\Omega, \mathbb{P}, S)$, avec $S = \{S_\omega\}_{\omega \in \Omega} = \{\hat{T}_\omega\}_{\omega \in \Omega}$. On observe d'abord que $S_{\underline{\omega}}^n = T_{\underline{\omega}}^n \oplus T_{\underline{\omega}}^n$. On sait aussi qu'il existe $\lambda > 1$ tel que $\lambda(T_\omega) > \lambda$ pour tout ω, et donc $s(S_\omega) = \lambda(T_\omega)^{-1} \leq \lambda^{-1}$.

Lemme 1.5.13. *Il existe $n \geq 1$, $\epsilon_0 > 0$, $\theta < 1$ et $K > 0$ tel que pour tout $f \in V_1(X^2)$,*

$$|\hat{P}^n f|_{\epsilon_0} \leq \theta |f|_{\epsilon_0} + K \|f\|_{L^1_{m_2}}.$$

Démonstration. Soit $\epsilon_0^{(n)} > 0$ plus petit que $\epsilon_0(S_{\underline{\omega}}^n)$, $\frac{1}{c(S_{\underline{\omega}}^n)}$ et $\delta(T_{\underline{\omega}}^n)$ pour tout $\underline{\omega} \in \Omega^n$. En sommant les inégalité données par le lemme 1.5.11 combiné au lemme 1.5.12 pour chaque $S_{\underline{\omega}}^n$, on obtient

$$|\hat{P}^n f|_{\epsilon_0^{(n)}} \leq (1 + \lambda^{-n}) \left(\lambda^{-n} + \frac{128 \lambda^{-n}}{\pi(1 - \lambda^{-n})} \right) |f|_{\epsilon_0^{(n)}} + A(\epsilon_0^{(n)}) \|f\|_{L^1_{m_2}},$$

où $A(\epsilon_0^{(n)})$ est une constante finie qui ne dépend que de $\epsilon_0^{(n)}$. Comme le terme en face de $|f|_{\epsilon_0^{(n)}}$ tend vers 0, il suffit de choisir n assez grand pour obtenir le lemme. □

Ceci montre que l'opérateur $\hat{P}$ vérifie les hypothèses (H1) à (H5) du paragraphe 1.1.1. Pour vérifier (H6), il suffit de montrer que le système auxiliaire possède la propriété de recouvrement et d'appliquer la proposition 1.1.14.

Lemme 1.5.14. *Le système $(\Omega, \mathbb{P}, S)$ possède la propriété de recouvrement : pour toute boule $B \subset [0,1]^2$, il existe $n \geq 1$ et $\underline{\omega} \in \Omega^n$ tel que $S_{\underline{\omega}}^n(B) = [0,1]^2$.*

2. Pour tout disque D de rayon R et tout rectangle $\mathcal{R} = [a,b] \times [c,d]$ dans le plan, on a $m_2(D \cap B_r^\infty(\partial \mathcal{R})) \leq 16 r R$ pour tout $r > 0$.

Démonstration. Soit $B \subset [0,1]^2$ une boule. Il existe deux intervalles non triviaux I et J inclus dans $[0,1]$ tels que $I \times J \subset B$. Comme le système $(\Omega, \mathbb{P}, T)$ possède la propriété de recouvrement, il existe $n_1 \geq 1$ et $\underline{\omega}^{(1)} \in \Omega^{n_1}$ tel que $T^{n_1}_{\underline{\omega}^{(1)}}(I) = [0,1]$. Soit $K \subset T^{n_1}_{\underline{\omega}^{(1)}}(J)$ un intervalle non trivial. Par la propriété de recouvrement, il existe $n_2 \geq 1$ et $\underline{\omega}^{(2)} \in \Omega^{n_2}$ tel que $T^{n_2}_{\underline{\omega}^{(2)}}(K) = [0,1]$, et donc $T^{n_2}_{\underline{\omega}^{(2)}}(T^{n_1}_{\underline{\omega}^{(1)}}(J)) = [0,1]$. En particulier, $T^{n_2}_{\underline{\omega}^{(2)}}$ est surjective. On pose $n = n_1 + n_2$ et $\underline{\omega} = \underline{\omega}^{(2)}\underline{\omega}^{(1)}$. On a

$$\begin{aligned} S^n_{\underline{\omega}}(I \times J) &= T^n_{\underline{\omega}}(I) \times T^n_{\underline{\omega}}(J) \\ &= T^{n_2}_{\underline{\omega}^{(2)}}(T^{n_1}_{\underline{\omega}^{(1)}}(I)) \times T^{n_2}_{\underline{\omega}^{(2)}}(T^{n_1}_{\underline{\omega}^{(1)}}(J)) \\ &= T^{n_2}_{\underline{\omega}^{(2)}}([0,1]) \times [0,1] \\ &= [0,1]^2, \end{aligned}$$

puisque $T^{n_2}_{\underline{\omega}^{(2)}}$ est surjective. □

Ainsi, l'opérateur de transfert $\hat{P}$ admet un trou spectral sur $V_1(X^2)$, ce qui termine l'étape 2.

Étape 3. D'après le lemme 1.5.2, le quenched CLT est vrai si et seulement si $\hat{\sigma}^2 = 2\sigma^2$. Le lemme suivant montre que cette condition est équivalente à la condition donnée par le théorème 1.5.10, ce qui en termine la preuve. □

Lemme 1.5.15. *On a $\hat{\sigma}^2 = 2\sigma^2$ si et seulement si*

$$\lim_{n\to\infty} \frac{1}{n} \int_{\tilde{\Omega}} \left(\sum_{k=0}^{n-1} \int_X \varphi \circ T^k_{\underline{\omega}} \, d\mu \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) = 0.$$

Démonstration. Comme la densité $h \otimes h$ de la mesure $\mu \otimes \mu$ appartient à $\hat{\mathcal{B}} = V_1(X^2)$, on a d'après le lemme 1.2.13

$$\hat{\sigma}^2 = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu\otimes\mu}(\hat{S}^2_n).$$

On a

$$\begin{aligned} \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu\otimes\mu}(\hat{S}^2_n) &= \sum_{k,l=0}^{n-1} \int_{\tilde{\Omega}} \int_X \int_X \hat{\varphi}(\hat{T}^k_{\underline{\omega}}(x,y))\hat{\varphi}(\hat{T}^l_{\underline{\omega}}(x,y)) d\mu(x) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \\ &= \sum_{k,l=0}^{n-1} \int_{\tilde{\Omega}} \int_X \int_X \left(\varphi(T^k_{\underline{\omega}}x) - \varphi(T^k_{\underline{\omega}}y)\right) \left(\varphi(T^l_{\underline{\omega}}x) - \varphi(T^l_{\underline{\omega}}y)\right) d\mu(x) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}). \end{aligned}$$

Or,

$$\begin{aligned} &\int_{\tilde{\Omega}} \int_X \int_X \left(\varphi(T^k_{\underline{\omega}}x) - \varphi(T^k_{\underline{\omega}}y)\right) \left(\varphi(T^l_{\underline{\omega}}x) - \varphi(T^l_{\underline{\omega}}y)\right) d\mu(x) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \\ &= \int_{\tilde{\Omega}} \int_X \varphi(T^k_{\underline{\omega}}x)\varphi(T^l_{\underline{\omega}}x) d\mu(x) d\tilde{\mathbb{P}}(\underline{\omega}) - \int_{\tilde{\Omega}} \left(\int_X \varphi(T^k_{\underline{\omega}}x) d\mu(x)\right) \left(\int_X \varphi(T^l_{\underline{\omega}}y) d\mu(y)\right) d\tilde{\mathbb{P}}(\underline{\omega}) \\ &- \int_{\tilde{\Omega}} \left(\int_X \varphi(T^k_{\underline{\omega}}y) d\mu(x)\right) \left(\int_X \varphi(T^l_{\underline{\omega}}x) d\mu(y)\right) d\tilde{\mathbb{P}}(\underline{\omega}) + \int_{\tilde{\Omega}} \int_X \varphi(T^k_{\underline{\omega}}y)\varphi(T^l_{\underline{\omega}}y) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}). \end{aligned}$$

On obtient donc

$$\begin{aligned}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu\otimes\mu}(\hat{S}_n^2) &= 2\int_{\tilde{\Omega}}\int_X \left(\sum_{k=0}^{n-1}\varphi(T_{\underline{\omega}}^k x)\right)^2 d\mu(x)d\tilde{\mathbb{P}}(\underline{\omega}) - 2\int_{\tilde{\Omega}}\left(\sum_{k=0}^{n-1}\int_X \varphi(T_{\underline{\omega}}^k x)d\mu(x)\right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) \\ &= 2\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2) - 2\int_{\tilde{\Omega}}\left(\sum_{k=0}^{n-1}\int_X \varphi\circ T_{\underline{\omega}}^k d\mu\right)^2 d\tilde{\mathbb{P}}(\underline{\omega}),\end{aligned}$$

d'où le résultat, puisque $\sigma^2 = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}(S_n^2)$ par la proposition 1.2.3. □

Remarque 1.5.16. La démonstration du lemme 1.5.15 est tout à fait générale, et ce lemme peut donc être formulé dans le même cadre que la section 1.5.2 : si les étapes 1 et 2 ont été vérifiées, alors le quenched CLT est vrai pour φ si et seulement si

$$\lim_{n\to\infty}\frac{1}{n}\int_{\tilde{\Omega}}\left(\sum_{k=0}^{n-1}\int_X \varphi\circ T_{\underline{\omega}}^k\, d\mu\right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) = 0.$$

Exemple 1.5.17.[3] On montre par un contre-exemple que le quenched CLT n'est pas toujours vérifié dès qu'on ne suppose plus que la mesure stationnaire est préservée par toutes les applications T_ω. On considèrera deux applications T_0 et T_1 de type Lasota-Yorke, uniformément dilatantes, et possédant toutes deux la propriété de recouvrement. Par conséquent, elles admettent chacune une unique mesure absolument continue et invariante, que l'on notera respectivement μ_0 et μ_1. On supposera que $\mu_0 \neq \mu_1$. On suppose aussi qu'il existe $C > 0$ et $\lambda < 1$ tels que pour toute suite (ω_n) dans $\{0,1\}$, tout $n \geq 1$ et toute fonction $f \in \mathrm{BV}$ d'intégrale nulle par rapport à Lebesgue, on a

$$\|P_{\omega_n}\dots P_{\omega_1} f\|_{\mathrm{BV}} \leq C\lambda^n \|f\|_{\mathrm{BV}}.$$

Cette hypothèse, dite de perte de mémoire exponentielle, sera étudiée au chapitre 2, où on verra qu'elle est satisfaite dans plusieurs cas, comme par exemple si T_0 et T_1 sont suffisamment proches dans une topologie convenable, ou si T_0 et T_1 sont deux β-transformations, avec des β assez proches.
On considère alors le système aléatoire de Lasota-Yorke donné par $\Omega = \{0,1\}$, et $p_0 = p_1 = \frac{1}{2}$. Il admet lui aussi la propriété de recouvrement, et donc admet une unique mesure stationnaire absolument continue μ. On considère un observable ψ de classe C^∞ tel que $\int_X \psi\, d\mu_0 \neq \int_X \psi\, d\mu_1$, et on pose $\varphi = \psi - \int_X \psi\, d\mu$. On va montrer que φ ne vérifie pas le quenched CLT. D'après le théorème 1.5.10, il s'agit de montrer que

$$\int_{\tilde{\Omega}}\left(\frac{\sum_{k=1}^{n-1}\int_X \varphi\circ T_{\omega_k}\circ\dots T_{\omega_1}\, d\mu}{\sqrt{n}}\right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) \tag{1.1}$$

ne tend pas vers 0.

3. Cet exemple m'a été communiqué par Sébastien Gouëzel.

En changeant la direction du temps, c'est-à-dire en remplaçant $(\omega_1, \dots, \omega_n)$ par $(\omega_n, \dots, \omega_1)$, et en appliquant l'opérateur de transfert, (1.1) se réécrit

$$\int_{\tilde{\Omega}} \left(\frac{\sum_{k=1}^{n-1} \int_X \varphi P_{\omega_k} \dots P_{\omega_n} h \, dm}{\sqrt{n}} \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}), \tag{1.2}$$

où h est la densité de μ.
Pour toute suite $\underline{\omega}$, on a grâce la propriété de perte de mémoire exponentielle

$$\begin{aligned} \|P_{\omega_1} \dots P_{\omega_n} h - P_{\omega_1} \dots P_{\omega_n} P_{\omega_{n+1}} h\|_{\mathrm{BV}} = \|P_{\omega_1} \dots P_{\omega_n}(h - P_{\omega_{n+1}} h)\|_{\mathrm{BV}} &\leq C\lambda^n \|h - P_{\omega_{n+1}} h\|_{\mathrm{BV}} \\ &\leq C\lambda^n. \end{aligned}$$

Par conséquent, $P_{\omega_1} \dots P_{\omega_n} h$ converge exponentiellement vite vers une fonction $h_{\underline{\omega}} \in \mathrm{BV}$: il existe $C > 0$ et $\lambda < 1$ telles que pour tout $n \geq 1$ et tout $\underline{\omega}$,

$$\|P_{\omega_1} \dots P_{\omega_n} h - h_{\underline{\omega}}\|_{\mathrm{BV}} \leq C\lambda^n.$$

Soit σ le décalage à gauche sur l'espace des suites infinies $\Omega^{\mathbb{N}}$. L'équation (1.2) se réécrit alors

$$\int_{\tilde{\Omega}} \left(\frac{\sum_{k=1}^{n-1} \left[\int_X \varphi h_{\sigma^{k-1}\underline{\omega}} \, dm + \mathcal{O}(\lambda^{n-k}) \right]}{\sqrt{n}} \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}).$$

On pose $G(\underline{\omega}) = \int_X \varphi \, h_{\underline{\omega}} \, dm$. Cette fonction est Lipschitz sur $\Omega^{\mathbb{N}}$ pour la distance symbolique $d_\lambda(\underline{\omega}, \underline{\omega}') = \lambda^{s(\underline{\omega}, \underline{\omega}')}$ avec

$$s(\underline{\omega}, \underline{\omega}') = \inf\{n \geq 1 \,:\, \omega_n \neq \omega'_n\}.$$

En effet, si $\omega_1 = \omega'_1, \dots, \omega_n = \omega'_n$, alors

$$\begin{aligned} |G(\underline{\omega}) - G(\underline{\omega}')| &\leq \|\varphi\|_{\sup} \|h_{\underline{\omega}} - h_{\underline{\omega}'}\|_{\mathrm{BV}} \\ &\leq \|h_{\underline{\omega}} - P_{\omega_1} \dots P_{\omega_n} h\|_{\mathrm{BV}} + \|(P_{\omega_1} \dots P_{\omega_n} - P_{\omega'_1} \dots P_{\omega'_n}) h\|_{\mathrm{BV}} + \|P_{\omega'_1} \dots P_{\omega'_n} h - h_{\underline{\omega}'}\|_{\mathrm{BV}} \\ &\leq C\lambda^n. \end{aligned}$$

Cette fonction G permet de réécrire (1.2) comme

$$\int_{\tilde{\Omega}} \left(\frac{\mathcal{O}(1) + \sum_{k=0}^{n-1} G(\sigma^k \underline{\omega})}{\sqrt{n}} \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}).$$

Si la fonction G était de moyenne non nulle, alors le théorème de Birkhoff pour σ (et $\tilde{\mathbb{P}}$) assurerait que l'équation précédente exploserait, et on aurait gagné. En fait, G est de moyenne nulle (sa moyenne est l'intégrale de φ pour la mesure stationnaire), donc il faut travailler un peu plus. Comme G est Hölder, elle satisfait un théorème de la limite centrale pour σ. Si ce théorème limite est non dégénéré, la variance asymptotique est non nulle, et on obtient (1.1). Reste à exclure le cas dégénéré. En ce cas, la fonction G serait un cobord L^2,

et même un cobord Hölder (résultat classique de régularité automatique). En particulier, G devrait être nulle sur les points fixes de F, et donc $G(0,0,0\dots) = G(1,1,1,\dots) = 0$. Comme $h(0,0,0,\dots) = \lim_n P_0^n h$ est la densité de la mesure stationnaire μ_0 pour T_0, on a $G(0,0,\dots) = \int \varphi\, d\mu_0$. De même, $G(1,1,\dots) = \int \varphi\, d\mu_1$. Ainsi,

$$\int \varphi\, d\mu_0 = \int \varphi\, d\mu_1 (= 0).$$

C'est absurde, car cela contredit le choix de notre fonction φ.

1.6 Remarques et questions ouvertes

1. La plupart des résultats du paragraphe 1.2 restent vrais lorsque le système est ergodique, mais plus nécessairement mélangeant, c'est-à-dire que 1 est une valeur propre simple de P, mais qu'il peut exister d'autres valeurs propres sur le cercle unité. On renvoie à [76] pour les démonstrations dans un cadre similaire. On peut aussi étendre ces résultats (hormis le principe des grandes déviations) au cas non-ergodique, mais dans ce cas, la loi limite n'est plus une loi gaussienne, mais une mixture de lois gaussiennes, voir la seconde partie de [76] ou [84].
2. Dans le cas des systèmes unidimensionnels, on a vu qu'il y a existence d'une mesure stationnaire absolument continue et que les théorèmes limites sont vérifies pour les systèmes dits "dilatants en moyenne". Néanmoins, Morita [115] a démontré l'existence d'une mesure stationnaire absolument continue pour une classe plus grande de systèmes. Lorsque Ω est fini, cette classe correspond aux systèmes $\{T_\omega\}_{\omega\in\Omega}$ qui vérifient $\sum_{\omega\in\Omega} p_\omega \log \lambda(T_\omega) > 0$. Pour la famille d'exemples de Pelikan étudiée à la section 1.1.11, cette condition est équivalente à $p > \frac{1}{2}$, qui est exactement la condition sous laquelle on a pu démontrer les théorèmes limites pour ces systèmes. Il est donc possible de se demander si ces résultats concernent l'ensemble des systèmes considérés par Morita, comme c'est le cas pour la famille précédente, ou s'il existe des systèmes satisfaisant la condition de Morita, mais pour lesquels certains théorèmes limites ne sont pas vérifiés. En particulier, est-ce que ces systèmes ont toujours une décroissance exponentielle des corrélations pour une classe raisonnable d'observables, ou cette décroissance peut-elle parfois être plus lente, par exemple polynomiale ?
3. Le théorème 1.5.4 donne un résultat sur la décroissance quenched des corrélations. Néanmoins, ce résultat est faible car il ne concerne que le cas où la mesure μ est préservée par toutes les applications T_ω et car il ne donne pas la dépendance vis-à-vis des normes des observables. Est-il possible de démontrer que pour tout système de Lasota-Yorke aléatoire mélangeant, il existe $\kappa < 1$ tel que pour tout $n \geq 1$, tout $\psi \in \mathrm{BV}$ et tout $\phi \in L^\infty(m)$,

$$\left| \int_X \psi\, \phi \circ T_{\underline{\omega}}^n\, d\mu - \int_X \psi\, d\mu \int_X \phi \circ T_{\underline{\omega}}^n\, d\mu \right| \leq C_{\underline{\omega}} \kappa^n \|\phi\|_{\mathrm{BV}} \|\phi\|_{L^\infty_m},$$

pour $\tilde{\mathbb{P}}$-presque tout $\underline{\omega}$? On pourrait par exemple adapter la technique de couplage de Stenlund et Sulku [135], qu'ils ont utilisée pour étudier les applications lisses et uniformément dilatantes du cercle.

4. On a vu à l'exemple 1.5.17 que la condition $\lim_{n\to\infty} \frac{1}{n} \int_{\tilde{\Omega}} \left(\sum_{k=0}^{n-1} \int_X \varphi \circ T_{\underline{\omega}}^k \, d\mu \right)^2 d\tilde{\mathbb{P}}(\underline{\omega}) = 0$ du théorème 1.5.10 n'est pas toujours vérifiée dès que que la mesure stationnaire n'est pas préservée par toutes les applications. Est-il possible néanmoins de prouver un quenched CLT avec recentrage aléatoire, c'est-à-dire

$$\frac{\sum_{k=0}^{n-1} \left(\varphi \circ T_{\underline{\omega}}^k - \int_X \varphi \circ T_{\underline{\omega}}^k \, d\mu \right)}{\sqrt{n}} \Longrightarrow_\mu \mathcal{N}(0, \sigma^2),$$

pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$, comme c'est le cas dans Kifer [93]. On remarquera que la condition du théorème 1.5.10 dit exactement que le terme rajouté pour la version avec recentrage tend vers 0 en norme L^2. Il est donc naturel de penser que si cette condition n'est pas vérifiée, on puisse néanmoins prouver un quenched CLT avec recentrage aléatoire.

5. Est-il possible de démontrer des versions quenched de théorèmes limites plus fins, comme le théorème de la limite locale ou le principe d'invariance presque sure, en se restreignant éventuellement au cas plus simple où la même mesure est préservée par toutes les applications ? La même question se pose pour le principe des grandes déviations quenched, puisque l'on a seulement démontré une borne supérieure quenched à la proposition 1.2.18

Chapitre 2

Perte de mémoire et théorèmes limites pour les systèmes dynamiques non-autonomes

Dans ce chapitre, on introduira la notion de système dynamique non-autonome, et on décrira le cadre développé par Conze et Raugi [46] pour les systèmes unidimensionnels. On étudiera ensuite un système constitué d'applications intermittentes de l'intervalle $[0, 1]$ ayant 0 comme point fixe neutre commun, et on montrera que la perte de mémoire pour ce type de système non-autonome est polynomiale. La première section de ce chapitre est un exposé de la théorie de Conze et Raugi tiré de l'article [6] coécrit avec Jérôme Rousseau, en cours de soumission, tandis que la deuxième section sur les systèmes avec points fixes neutres présente les résultats de l'article [5], à paraître sur *Discrete and Continuous Dynamical Systems*.

2.1 Systèmes dynamiques non-autonomes

2.1.1 Quelques généralités

Soit $(X, \mathcal{F}, m)$ un espace probabilisé. Un système dynamique non-autonome, aussi appelé système dynamique séquentiel, est donné par une suite $(T_n)_{n\geq 1}$ de transformations non singulières sur $(X, \mathcal{F}, m)$.
Pour $n \leq m$, on note par T_n^m la composition $T_n^m = T_m \circ \ldots \circ T_n$. Pour une application non singulière T, on notera par P_T son opérateur de transfert, défini par rapport à la mesure m. Il agit sur $L^1(m)$ et satisfait pour tout $f \in L^1(m)$ et $g \in L^\infty(m)$:

$$\int P_T f.g \, dm = \int f.g \circ T \, dm.$$

On a clairement $P_{T_n^m} = P_{T_m} \circ \ldots \circ P_{T_n}$. Pour alléger les notations, on oubliera l'indice T lorsque la suite de transformations (T_n) est fixée, et on écrira P_n^m pour désigner l'opérateur

$P_{T_n^m}$.
On peut généraliser au cas des systèmes séquentiels les notions d'ergodicité, mélange et exactitude. Ces définitions ont été introduites par Berend et Bergelson [21] (sous une forme légérement plus forte que celle présentée ci-après) dans le cas où toutes les applications T_n préservent m, et par Conze et Raugi [46] pour le cas général.

Définition 2.1.1. *On dira que le système séquentiel $(T_n)_{n\geq 1}$ est :*

1. *ergodique si pour tout $f \in L^1(m)$ et tout $g \in L^\infty$,*
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\left[\int_X f.g\circ T_1^k\,dm - \int_X f\,dm\int_X g\circ T_1^k\,dm\right] = 0;$$
2. *faiblement mélangeant si pour tout $f \in L^1(m)$ et tout $g \in L^\infty$,*
$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}\left|\int_X f.g\circ T_1^k\,dm - \int_X f\,dm\int_X g\circ T_1^k\,dm\right| = 0;$$
3. *mélangeant si si pour tout $f \in L^1(m)$ et tout $g \in L^\infty$,*
$$\lim_{n\to\infty} \int_X f.g\circ T_1^n\,dm - \int_X f\,dm\int_X g\circ T_1^n\;dm = 0.$$

Pour définir la notion de système exact, on a besoin d'introduire la tribu asymptotique $\mathcal{F}_\infty$. Pour $n \geq 1$, on définit la tribu
$$\mathcal{F}_n = (T_1^n)^{-1}(\mathcal{F}),$$
et la tribu asymptotique
$$\mathcal{F}_\infty = \bigcap_{n\geq 1}\mathcal{F}_n.$$

Définition 2.1.2. *On dira que le système séquentiel $(T_n)_{n\geq 1}$ est exact si tout événement $A \in \mathcal{F}_\infty$ vérifie $m(A) = 0$.*

Soit $f \in L^1(m)$. Comme $P_1^n f(x) = 0$ p.p. sur l'ensemble $\{P_1^n\mathbb{1} = 0\}$, on définit le quotient $\frac{P_1^n f}{P_1^n\mathbb{1}}$ comme valant 0 sur $\{P_1^n\mathbb{1} = 0\}$. On a alors la relation
$$\mathbb{E}_m(f|\mathcal{F}_n) = \left(\frac{P_1^n f}{P_1^n\mathbb{1}}\right)\circ T_1^n.$$

D'après le théorème de convergence de Doob pour les martingales, la suite des espérances conditionnelles $\mathbb{E}_m(f|\mathcal{F}_n)$ converge p.s. et dans $L^1(m)$ vers $\mathbb{E}_m(f|\mathcal{F}_\infty)$. Comme $\|\mathbb{E}_m(f|\mathcal{F}_n)\|_{L^1_m} = \|P_1^n f\|_{L^1_m}$, on en déduit :

Lemme 2.1.3. *Le système séquentiel $(T_n)_{n\geq 1}$ est exact si et seulement si $\lim_{n\to\infty}\|P_1^n f\|_{L^1_m} = 0$ pour tout $f \in L^1(m)$ avec $\int_X f\,dm = 0$.*

Il est donc clair que l'exactitude entraine le mélange et l'ergodicité au sens de la définition 2.1.1.

2.1.2 Un cadre fonctionnel abstrait

Conze et Raugi [46] ont étendu la théorie spectrale pour les itérés d'un seul opérateur au cas de la compositions de différents opérateurs, avec pour objectif d'appliquer cette théorie aux opérateurs de transfert des systèmes séquentiels. On décrit brièvement le cadre dans lequel ils se placent et leurs résultats.
Soit $(\mathcal{B}, \|.\|)$ un espace de Banach, $\mathcal{V}$ un sous-espace de $\mathcal{B}$ équipé d'une norme $|.|_v$ telle que $\|.\| \leq |.|_v$. Soit $\mathcal{P}$ un ensemble de contractions linéaires de $(\mathcal{B}, \|.\|)$ [1] laissant $\mathcal{V}$ invariant et satisfaisant

(H1) La boule unité de $(\mathcal{V}, |.|_v)$ est relativement compacte dans $(\mathcal{B}, \|.\|)$.

(H2) Il existe une famille dénombrable dans $\mathcal{V}$ qui est dense dans $(\mathcal{B}, \|.\|)$.

(H3) Il existe un entier $r \geq 1$ et des constantes $0 < \rho_r < 1$, $M_0, C_r > 0$ telles que :

$$\forall P \in \mathcal{P},\ |Pf|_v \leq M_0 |f|_v,\ \forall f \in \mathcal{V};$$

et pour tout r-uplet $P_1, \ldots, P_r$ d'opérateurs dans $\mathcal{P}$:

$$\forall f \in \mathcal{V},\ |P_r \ldots P_1 f|_v \leq \rho_r |f|_v + C_r \|f\|.$$

Ces hypothèses entrainent qu'il existe $M > 0$ tel que $|P_n \ldots P_1 f|_v \leq M|f|_v$ pour tout $n \geq 1$, tout $f \in \mathcal{V}$ et tout choix d'opérateurs $P_1, \ldots, P_n$ dans $\mathcal{P}$.
Conze et Raugi ont principalement considéré le cas de systèmes unidimensionnels, où $(X, \mathcal{F}, m)$ est l'intervale $[0, 1]$ muni de sa tribu borélienne et de la mesure de Lebesgue, et où les applications T sont de classe C^2 par morceaux et uniformément dilatantes. Dans ce contexte, la classe $\mathcal{P}$ est constituée de l'ensemble des opérateurs de transferts des applications considérées, et le choix naturel pour les espaces de Banach est de prendre $\mathcal{B} = L^1(m)$ et $\mathcal{V} = \mathrm{BV}$ l'espace des fonctions à variation bornée.
Soit $\mathcal{V}_0$ un sous-espace de $\mathcal{V}$ invariant par tous les opérateurs de $\mathcal{P}$. Dans la pratique, $\mathcal{V}_0$ sera constitué des fonctions de $\mathcal{V}$ d'intégrale nulle par rapport à m.

Définition 2.1.4. *On dit qu'une suite d'opérateurs $(P_n)_{n\geq 1}$ dans $\mathcal{P}$ est exacte sur $\mathcal{V}_0$ si pour tout $f \in \mathcal{V}_0$,*

$$\lim_{n\to\infty} \|P_n \ldots P_1 f\| = 0.$$

Un opérateur $P \in \mathcal{P}$ est exact sur $\mathcal{V}_0$ si la suite donnée par $P_n = P$ est exacte sur $\mathcal{V}_0$. Lorsque $(T_n)_{n\geq 1}$ est un système séquentiel sur $(X, \mathcal{F}, m)$, si P_n est l'opérateur de transfert de T_n, si $\mathcal{B} = L^1(m)$ et si $\mathcal{V}_0$ est l'ensemble des fonctions de $\mathcal{V}$ d'intégrale nulle par rapport à m, alors l'exactitude de la suite $(P_n)_{n\geq 1}$ sur $\mathcal{V}_0$ est équivalente à l'exactitude du système $(T_n)_{n\geq 1}$. Ceci se déduit facilement du fait que $\mathcal{V}$ est dense dans $\mathcal{B}$ et que $\mathcal{P}$ est équicontinue sur $\mathcal{B}$.

1. c-à-d un ensemble d'opérateurs linéaires $P : \mathcal{B} \to \mathcal{B}$ vérifiant $\|Pf\| \leq \|f\|$ pour tout $f \in \mathcal{B}$.

2.1.3 Perte de mémoire exponentielle

Il est bien connu depuis le travail de Ionescu-Tulcea et Marinescu [83] que chaque opérateur de $\mathcal{P}$ est quasi-compact sur $\mathcal{V}$ et bénéficie de bonnes propriétés spectrales ; on pourra regarder la proposition 2.8 de [46] pour une preuve auto-contenue, et [19, 27, 76] pour plus d'informations. Comme conséquence directe de la quasi-compacité, on obtient que pour tout opérateur $P \in \mathcal{P}$ exact sur $\mathcal{V}_0$, la norme de P^n vu comme opérateur sur $\mathcal{V}_0$ décroit exponentiellement vite : il existe $K > 0$ et $\theta < 1$ tels que $|P^n f|_v \leq K\theta^n |f|_v$ pour tout $n \geq 0$ et tout $f \in \mathcal{V}_0$.
Pour la composition d'opérateurs différents, la situation est plus délicate, et on a besoin de plus d'hypothèses afin d'assurer une décroissance exponentielle. Premièrement, on énonce la propriété de décorrélation :

(Dec) un sous-ensemble $\mathcal{P}_0 \subset \mathcal{P}$ satisfait la propriété de décorrélation (exponentielle) dans $\mathcal{V}_0$ s'il existe $\theta < 1$ et $K > 0$ tel que, pour tous entiers $l \geq 1$, tout l-uplet d'opérateurs $P_1, \ldots, P_l$ dans $\mathcal{P}_0$:
$$\forall f \in \mathcal{V}_0,\ |P_l \ldots P_1 f|_v \leq K\theta^l |f|_v.$$

Conze et Raugi [46] ont donné un critère permettant de vérifier que **(Dec)** est satisfaite. Plutôt que d'énoncer leur condition, on donne deux corollaires de ce critère.
Le premier, qui est de nature locale, montre que tout opérateur $P \in \mathcal{P}$ exact possède un voisinage convenable pour lequel **(Dec)** est vérifiée. Plus précisément , pour deux opérateurs $P, P' \in \mathcal{P}$, on définit
$$d(P, P') = \sup_{\{f \in \mathcal{V} : |f|_v \leq 1\}} \|Pf - P'f\|,$$
et pour $\delta > 0$, on pose $B(P, \delta) = \{P' \in \mathcal{P} \,:\, d(P, P') < \delta\}$.

Proposition 2.1.5 ([46, Proposition 2.10])**.** *Pour tout opérateur $P \in \mathcal{P}$ exact sur $\mathcal{V}_0$, il existe $\delta_0 > 0$ tel que $\mathcal{P}_0 = \mathcal{P} \cap B(P, \delta_0)$ vérifie* **(Dec)** *dans $\mathcal{V}_0$.*

L'autre corollaire, qui ne nécessite pas que les opérateurs soient très proches, requiert en revanche une hypothèse de compacité :

(Comp) un sous-ensemble $\mathcal{P}_0 \subset \mathcal{P}$ satisfait la propriété de compacité si pour toute suite (P_n) dans $\mathcal{P}_0$, il existe une sous-suite (P_{n_j}) et un opérateur $P \in \mathcal{P}_0$ tel que
$$\forall f \in \mathcal{B},\ \lim_j \|P_{n_j} f - Pf\| = 0.$$

Proposition 2.1.6 ([46, Proposition 2.11])**.** *Si $\mathcal{P}_0$ satisfait la condition* **(Comp)** *et est tel que toutes les suite (P_n) de $\mathcal{P}_0$ sont exactes sur $\mathcal{V}_0$, alors il satisfait la condition* **(Dec)** *dans $\mathcal{V}_0$.*

2.1.4 Application aux systèmes unidimensionnels

On décrit comment cette théorie peut s'appliquer dans le cas de systèmes séquentiels, pour des applications dilatantes de $X = [0,1]$. Plus précisément, on considère des applications $T : X \to X$ uniformément dilatantes, c-à-d $\lambda(T) := \inf_x |T'(x)| > 1$, pour lesquelles il existe une partition finie $\mathcal{A}_T$ constituée d'un nombre fini d'intervalles avec intérieurs disjoints, et telle que T peut être prolongée en une application C^2 sur un voisinage de chaque élément de cette partition.
L'opérateur de transfert $P_T : L^1(m) \to L^1(m)$ vérifie

$$P_T f(x) = \sum_{Ty=x} \frac{f(y)}{|T'(y)|},$$

pour tout $f \in L^1(m)$, et est quasi-compact sur l'espace BV des fonctions à variation bornée, comme rappelé au paragraphe 1.1.2.
Le couple $(\mathcal{B}, \|.\|) = (L^1(m), \|.\|_{L^1_m})$ et $(\mathcal{V}, |.|_v) = (\mathrm{BV}, \|.\|_{\mathrm{BV}})$ vérifie les hypothèses (H1) et (H2). Pour s'assurer que (H3) est satisfaite, on utilisera l'inégalité de Lasota-Yorke [99] : pour toute application T dilatante par morceaux et toute fonction $f \in \mathrm{BV}$, on a

$$\mathrm{Var}(P_T f) \leq \frac{2}{\lambda(T)} \mathrm{Var} f + C(T) \|f\|_{L^1_m},$$

avec $C(T) = \sup_x \frac{|T''|}{|T'|^2} + 2 \sup_{I \in \mathcal{A}_T} \frac{\sup_{x \in I} |T'(x)|^{-1}}{m(I)}$.
Ainsi, toute classe $\mathcal{C}$ d'applications dilatantes par morceaux T telle que $\sup_{T \in \mathcal{C}} C(T) < \infty$ et pour laquelle il existe $r \geq 1$ tel que

$$\inf_{T_1, \ldots, T_r \in \mathcal{C}} \lambda(T_1^r) > 2, \text{ et } \sup_{T_1, \ldots, T_r \in \mathcal{C}} C(T_1^r) < \infty,$$

satisfait (H3) lorsque $\mathcal{P}$ est la classe des opérateurs de transfert des applications de $\mathcal{C}$. A partir de maintenant, on dira que

(D_r) une classe $\mathcal{C}$ satisfait cette condition si l'ensemble $\mathcal{P}$ des opérateurs de transfert des applications de $\mathcal{C}$ satisfait (H3) pour ce $r \geq 1$.

La classe décrite juste auparavant satisfait donc **(D_r)**.
Il est démontré dans Conze et Raugi [46, Théorème 3.4] que la classe des β-transformations, c-à-d les applications de la forme $x \mapsto \beta x \bmod 1$, avec $\beta \geq 1 + \alpha$, $\alpha > 0$ satisfait **(D_r)** pour un certain $r \geq 1$ ne dépendant que de α. De plus, les auteurs démontrent que toute suite de cette classe de transformations est exacte (Théorème 3.6), et que $d(P_1, P_2) \leq C|\beta_1 - \beta_2|$, pour une constante universelle $C > 0$, où les $\beta_i > 1$ sont arbitraires et P_i désigne l'opérateur de transfert de la β_i-transformation, pour $i = 1, 2$ (Lemme 3.9).
Par conséquent, la classe des β-transformations, avec β compris entre $\beta_{\min} > 1$ et $\beta_{\max} < \infty$ satisfait la condition **(Comp)** et vérifie donc la condition de décorrélation **(Dec)**.

Il est à noter que la condition **(Dec)** donne la décroissance exponentielle de la norme BV de $P_1^n f - P_1^n g$ pour toutes densités de probabilité $f, g \in \mathrm{BV}$, à la différence du résultat de Ott, Stenlund et Young [121], où seule la décroissance exponentielle de la norme $L^1(m)$ est démontrée.

2.1.5 Théorèmes limites

Soit $\mathcal{C}$ une classe de transformations de type Lasota-Yorke sur $[0,1]$ qui satisfait $(\boldsymbol{D_r})$ et **(Dec)**. Soit $(T_n)_{n\geq 1}$ une suite d'applications de $\mathcal{C}$ et $(f_n)_{n\geq 0}$ une suite bornée dans BV. On pose

$$S_n = \sum_{k=0}^{n-1} \left[f_k \circ T_1^k - \int f_k \circ T_1^k \, dm \right].$$

On a alors la loi des grands nombres :

Théorème 2.1.7 ([46, Théorème 3.7])**.** Si $(T_n)_{n\geq 1}$ est une suite de $\mathcal{C}$ et $(f_n)_{n\geq 0}$ est bornée dans BV, alors on a $\lim_{n\to\infty} \frac{S_n}{n} = 0$ *presque surement.*

Afin d'obtenir le théorème de la limite centrale correspondant, on a besoin d'une hypothèse supplémentaire :

(Min) la suite d'applications $(T_n)_{n\geq 1}$ satisfait cette condition s'il existe $\delta > 0$ tel que $P_1^n \mathbb{1} \geq \delta$ pour tout n.

On remarquera que cette condition ne concerne que la suite (T_n), et non toute la classe $\mathcal{C}$. Sous cette hypothèse additionnelle, le TCL est vérifié :

Théorème 2.1.8 ([46, Théorème 5.1])**.** *Si $(T_n)_{n\geq 1}$ est une suite d'applications de $\mathcal{C}$ qui vérifie* **(Min)** *et si $(f_n)_{n\geq 0}$ est bornée dans BV, alors on a l'alternative suivante :*

1. *Soit $\|S_n\|_{L^2_m}$ est bornée, et dans ce cas S_n est bornée presque surement ;*
2. *Soit $\|S_n\|_{L^2_m}$ converge vers $+\infty$, et dans ce cas $\frac{S_n}{\|S_n\|_{L^2_m}}$ converge en loi vers $\mathcal{N}(0,1)$.*

Sous une hypothèse de croissance suffisante de la variance $\|S_n\|^2_{L^2_m}$, Haydn, Nicol, Török et Vaienti [72] ont démontré récemment un principe d'invariance presque sure :

Théorème 2.1.9 ([72, Théorème 3.1])**.** *Si (T_n) est une suite d'applications de $\mathcal{C}$ qui vérifie* **(Min)** *et si $(f_n)_{n\geq 0}$ est bornée dans BV et vérifie $\|S_n\|_{L^2_m} \geq C n^{\frac{1}{4}+\epsilon}$, pour $C > 0$ et $0 < \epsilon < \frac{1}{4}$, alors, quitte à élargir l'espace de probabilité, on peut trouver une suite $(Z_k)_{k\geq 0}$ de variables aléatoires gaussiennes centrées indépendantes telles que pour tout $\beta < \epsilon$:*

$$\sup_{1\leq k\leq n} \left| S_k - \sum_{i=0}^{k-1} Z_i \right| = o(\|S_n\|^{1-\beta}_{L^2_m}), \quad m - p.s.$$

On peut aussi démontrer une inégalité de grandes déviations :

Proposition 2.1.10. *Si (T_n) est une suite d'applications de $\mathcal{C}$ qui vérifie* **(Min)** *et si $(f_n)_{n\geq 0}$ est bornée dans BV, alors pour tout $t > 0$, il existe $\tau = \tau((f_n)_n)$ et $C = C((f_n)_n, t) > 0$ tels que pour tout $n \geq 1$:*

$$m(S_n > nt) \leq Ce^{-\tau n t^2}.$$

Démonstration. On suit la preuve de la proposition 2.5 de [7], en utilisant l'approximation par des martingales décrite à la section 5 de [46]. On pose $\tilde{f}_n = f_n - \int f_n \circ T_1^n \, dm$ et on définit les opérateurs Q_n, $n \geq 1$, par

$$Q_n g = \frac{P_n(gP_1^{n-1}\mathbb{1})}{P_1^n \mathbb{1}}.$$

Soit h_n la suite définie par les relations $h_0 = 0$ et $h_{n+1} = Q_{n+1}\tilde{f}_n + Q_{n+1}h_n$. on a

$$h_n = \frac{1}{P_1^n \mathbb{1}} \sum_{k=0}^{n-1} P_{k+1}^n(\tilde{f}_k P_1^k \mathbb{1}).$$

Comme BV est une algèbre de Banach et $(P_1^k\mathbb{1})_k$ est bornée dans BV, on en déduit que $\tilde{f}_k P_1^k \mathbb{1}$ est une suite bornée dans BV de fonctions d'intégrale nulle. Par conséquent, la condition **(Dec)** entraine que la norme BV de $P_{k+1}^n(\tilde{f}_k P_1^k \mathbb{1})$ décroit exponentiellement vite avec $n-k$, d'où l'on obtient, grâce à la condition **(Min)**, que la suite (h_n) est bornée pour la norme L^∞. On pose $\varphi_n = \tilde{f}_n + h_n - h_{n+1} \circ T_{n+1}$ et $U_n = \varphi_n \circ T_1^n$. On a alors que (U_n) est une martingale renversée pour la filtration $(\mathcal{F}_n)$, et est bornée pour la norme L^∞. De plus, on a

$$S_n = \sum_{k=0}^{n-1} U_k + h_n \circ T_1^n.$$

Par conséquent, $m(S_n > nt) = m(\sum_{k=0}^{n-1} U_k + h_n \circ T_1^n > nt) \leq m(\sum_{k=0}^{n-1} U_k > \frac{nt}{2})$, puisque (h_n) est bornée pour la norme L^∞, et donc $h_n \circ T_1^n < \frac{nt}{2}$ pour tout n assez grand. D'autre part, en utilisant l'inégalité de Azuma-Hoeffding, il existe $\tau > 0$ tel que $m(\sum_{k=0}^{n-1} U_k > \frac{nt}{2}) \leq e^{-\tau n t^2}$, puisque (U_n) est bornée pour la norme L^∞, ce qui termine la démonstration. □

2.1.6 La condition (Min)

Dans cette section, on discute de la validité de la condition **(Min)**, nécessaire pour les théorèmes limites de la section précédente. Pour le cas des β-transformations, Conze et Raugi [46] ont démontré que pour tout $\beta > 1$, il existe un voisinage $[\beta - a, \beta + a]$ de β tel que si T_n est la β_n-transformation, avec $\beta_n \in [\beta - a, \beta + a]$ pour tout n, alors la suite (T_n) vérifie la condition **(Min)**.

On donne une condition qui entraine automatiquement la validité de **(Min)**. Si (T_n) est un système séquentiel sur $[0, 1]$ constitué d'applications de type Lasota-Yorke, on désigne par $\mathcal{A}_n$ la partition de monotonicité et régularité de T_n, et par $\mathcal{A}_n^m$ la partition de T_n^m, $n \leq m$.

Définition 2.1.11. *Le système dynamique séquentiel $(T_n)_{n\geq 1}$ possède la propriété de recouvrement si*

$$\forall n,\ \exists N(n)\ :\ \forall m,\ \forall I \in \mathcal{A}_{m+1}^{m+n},\ T_{m+1}^{m+N(n)}(I) = [0,1].$$

Soit $\mathcal{C}$ une classe d'applications surjectives de type Lasota-Yorke sur l'intervalle $[0,1]$ qui satisfait $(\boldsymbol{D_r})$ pour un certain $r \geq 1$ et pour laquelle il existe $\lambda > 1$ et $C > 0$ tels que $\inf |T'| \geq \lambda$ et $\sup |T'| \leq C$ pour tout $T \in \mathcal{C}$.

Proposition 2.1.12. *Soit $(T_n)_{n\geq 1}$ un système dynamique séquentiel appartenant à une classe $\mathcal{C}$ qui satisfait les conditions ci-dessus. On suppose que $(T_n)_{n\geq 1}$ possède la propriété de recouvrement. Alors la condition* **(Min)** *est vérifiée.*

Démonstration. Pour $a > 0$, on pose

$$\mathcal{E}_a = \left\{ f \in \mathrm{BV}\ :\ f \geq 0,\ f \neq 0,\ \|f\|_{\mathrm{BV}} \leq a \int f \right\}.$$

Rappelons que d'après le lemme 3.2 de Liverani [102], pour toute partition $\mathcal{A}$ constituée d'intervalles de longueur au plus $\frac{1}{2a}$, et pour tout $f \in \mathcal{E}_a$, il existe $I \in \mathcal{A}$ tel que $f(x) \geq \frac{1}{2}\int f$ pour tout $x \in I$. Comme $\mathcal{C}$ satisfait (D_r), pour tout choix $P_1, \ldots, P_r$ d'opérateurs de transfert associés à des applications de $\mathcal{C}$ et tout $f \in \mathcal{E}_a$, on a

$$\|P_r \ldots P_1 f\|_{\mathrm{BV}} \leq \rho_r \|f\|_{\mathrm{BV}} + C_r \|f\|_{L^1_m} \leq (a\rho_r + C_r) \int f.$$

Ainsi, si on choisit $a \geq \frac{C_r}{1-\rho_r}$, on a $(P_r \ldots P_1)(\mathcal{E}_a) \subset \mathcal{E}_a$ pour tout choix de $P_1, \ldots, P_r$. Afin d'avoir $\mathbb{1} \in \mathcal{E}_a$, on prendra $a = \max\{1, \frac{C_r}{1-\rho_r}\}$.
Soit $(T_n)_{n\geq 1}$ une suite d'applications dans $\mathcal{C}$ qui possède la propriété de recouvrement. Pour tout $n \leq m$, le diamètre de $\mathcal{A}_n^m$ est inférieur à $\lambda^{-(m-n+1)}$, on peut donc choisir $n_0 \geq 1$ tel que le diamètre de $\mathcal{A}_{n+1}^{n+n_0}$ est inférieur à $\frac{1}{2a}$ pour tout n.
Soit $m \geq 0$. On a $P_1^{m+N(n_0)}\mathbb{1} = P_{m+1}^{m+N(n_0)} P_1^m \mathbb{1}$. On écrit $m = p_m r + q_m$ avec $0 \leq q_m < r$. On a :

$$P_1^{q_m}\mathbb{1}(x) = \sum_{T_1^{q_m} y = x} \frac{1}{|(T_1^{q_m})'(y)|} \geq C^{-q_m} \geq C^{-r}, \tag{2.1}$$

car toutes les applications sont surjectives et ont une dérivée uniformément bornée par C. Par conséquent, on a $P_1^{m+N(n_0)}\mathbb{1} \geq C^{-r} P_{m+1}^{m+N(n_0)} g_m$, avec $g_m = P_{q_m+1}^m \mathbb{1}$. Comme $P_{q_m+1}^m \mathbb{1}$ est une composition de p_m blocs de r opérateurs appliqués à une fonction de $\mathcal{E}_a$, on obtient que g_m appartient à $\mathcal{E}_a$. Il existe donc un intervalle $I_m \in \mathcal{A}_{m+1}^{m+n_0}$ sur lequel $g_m \geq \frac{1}{2}$. Ceci entraine

$$P_1^{m+N(n_0)}\mathbb{1}(x) \geq \frac{C^{-r}}{2} P_{m+1}^{m+N(n_0)} \mathbb{1}_{I_m}(x) = \frac{C^{-r}}{2} \sum_{T_{m+1}^{m+N(n_0)} y = x} \frac{\mathbb{1}_{I_m}(y)}{|(T_{m+1}^{m+N(n_0)})'(y)|} \geq \frac{C^{-r}}{2} C^{-N(n_0)},$$

puisque pour tout $x \in [0,1]$, il existe $y \in I_m$ tel que $T_{m+1}^{m+N(n_0)} y = x$ par l'hypothèse de recouvrement. En utilisant le même argument que pour (2.1), on trouve $P_1^k \mathbb{1} \geq C^{-N(n_0)}$ pour tout $k < N(n_0)$, ce qui montre que **(Min)** est satisfaite pour $\delta = \min\{C^{-N(n_0)}, \frac{C^{-(r+N(n_0))}}{2}\}$. □

2.2 Applications de l'intervalle avec un point fixe neutre

Dans la section précédente, les systèmes non-autonomes considérés présentaient tous une vitesse exponentielle de perte de mémoire. Si l'on recherche des systèmes avec une vitesse de perte de mémoire polynomiale, il est naturel de considérer des applications intermittentes de l'intervalle, avec un point fixe neutre en commum. En effet, dans un cadre purement autonome, ces applications ont une décroissance polynomiale des corrélations, voir par exemple Liverani, Saussol et Vaienti [104], ou Young [144] par exemple.

2.2.1 Le modèle et le résultat principal

Pour $0 < \alpha < 1$, on définit $T_\alpha : [0,1] \to [0,1]$ par

$$T_\alpha(x) = \begin{cases} x + \dfrac{3^\alpha}{2^{1+\alpha}} x^{1+\alpha} & \text{si } 0 \leq x < \dfrac{2}{3} \\ 3x - 2 & \text{si } \dfrac{2}{3} \leq x \leq 1. \end{cases}$$

On identifie l'intervalle unité avec le cercle, de façon à rendre T_α continue. Nous allons considérer le système non-autonome (T_n) donné par $T_n = T_{\beta_n}$, où la suite (β_n) vérifie $0 < \beta_n \leq \alpha < 1$ pour tout $n \geq 1$. On note par P_β l'opérateur de transfert de T_β par rapport à la mesure de Lebesgue m. Comme dans la section précédente, on notera $T_n^m = T_m \circ \ldots \circ T_n$ et $P_n^m = P_m \circ \ldots \circ P_n$ l'opérateur de transfert de T_n^m.

Théorème 2.2.1. *Soit $0 < \alpha < 1$. Alors pour toute suite (β_n) avec $\beta^\star \leq \beta_n \leq \alpha$, et tout ϕ, ψ de classe C^1 et de même intégrale $\int \phi dm = \int \psi dm = 1$, on a*

$$\|P_1^n \phi - P_1^n \psi\|_{L_m^1} \leq \mathcal{F}(\|\phi\|_{C^1} + \|\psi\|_{C^1}) \frac{(\log n)^{\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}-1}},$$

où $\mathcal{F} : \mathbb{R} \to \mathbb{R}$ est affine.

Remarque 2.2.2. Le facteur logarithmique est un artefact de la preuve, et le résultat est certainement vrai sans ce terme, bien que la preuve que nous utilisons ne permet pas de le montrer.

Afin de démontrer ce théorème, on suivra la méthode employée par Liverani, Saussol et Vaienti [104].

2.2.2 Un cône invariant

Par les résultats de Thaler [136] et de Hu [82], on sait que la densité de la mesure absolument continue et invariante par T_α satisfait $f(x) \leq Cx^{-\alpha}$ dans un voisinage de 0. On va construire un cône de fonctions, vérifiant entre autre cette propriété, qui est invariant par l'opérateur de

transfert de T_α. Comme les opérateurs de transfert de T_β, pour $\beta \leq \alpha$ possèdent de meilleures propriétés régularisantes que l'opérateur de T_α, on peut s'attendre à ce qu'ils préservent aussi le cône.
On définit le cône de fonctions

$$\mathcal{C}_1 = \left\{ f \in C^0(]0,1]) \,:\, f \geq 0,\ f \text{ décroissante},\ X^{\alpha+1} f \text{ croissante} \right\},$$

où $X(x) = x$ est la fonction identité.

Lemme 2.2.3. *Le cône $\mathcal{C}_1$ est invariant par les opérateurs P_β, $0 < \beta \leq \alpha$.*

Démonstration. Clairement, seule l'invariance de la propriété $X^{\alpha+1} f$ croissante est à vérifier. Soit $y_1 < y_2$ donnés par $T_\beta^{-1}(x) = \{y_1, y_2\}$. On pose $\chi_\beta = \frac{3^\beta y_1^\beta}{2^{1+\beta}}$. Un calcul direct donne

$$X^{\alpha+1} P_\beta f(x) = \frac{f(y_1) y_1^{\alpha+1} (1+\chi_\beta)^{\alpha+1}}{1+(1+\beta)\chi_\beta} + f(y_2) \left(\frac{3y_2 - 2}{y_2} \right)^{\alpha+1} \frac{y_2^{\alpha+1}}{3}.$$

Le résultat en découle directement, puisque les applications $x \mapsto x^{\alpha+1} f(x)$, $x \mapsto \chi_\beta$, $x \mapsto y_1$ et $x \mapsto y_2$ sont croissantes. Le fait que $\alpha \geq \beta$ entraine la croissance de $\chi \mapsto \frac{(1+\chi)^{\alpha+1}}{1+(1+\beta)\chi}$. □

Lemme 2.2.4. *Pour $0 < \alpha < 1$, le cône*

$$\mathcal{C}_2 = \left\{ f \in \mathcal{C}_1 \cap L^1(m) \,:\, f(x) \leq a x^{-\alpha} \int f dm \right\}$$

est préservé par tous les opérateurs P_β, $0 < \beta \leq \alpha$, pourvu que a soit choisi assez grand.

Démonstration. On peut supposer que $\int f dm = 1$. En utilisant les notations du lemme précédent et en remarquant que $x^{\alpha+1} f(x) \leq f(1) \leq \int f dm = 1$, on trouve

$$P_\beta f(x) = \frac{f(y_1)}{T'_\beta(y_1)} + \frac{f(y_2)}{T'_\beta(y_2)} \leq \frac{a y_1^{-\alpha}}{T'_\beta(y_1)} + \frac{y_2^{-\alpha-1}}{T'_\beta(y_2)} = \left\{ \left(\frac{x}{y_1} \right)^\alpha \frac{1}{T'_\beta(y_1)} + \frac{1}{a} \frac{x^\alpha}{y_2^{\alpha+1} T'_\beta(y_2)} \right\} a x^\alpha.$$

On a $\left(\frac{x}{y_1}\right)^\alpha \frac{1}{T'_\beta(y_1)} = \frac{(1+\chi_\beta)^\alpha}{1+(1+\beta)\chi_\beta}$ et

$$\frac{1}{a} \frac{x^\alpha}{y_2^{\alpha+1} T'_\beta(y_2)} = \frac{1}{a} \left(\frac{3y_1}{2} \right)^{\alpha-\beta} \left(\frac{2}{x+2} \right)^{\alpha+1} \chi_\beta (1+\chi_\beta)^\alpha \leq \frac{1}{a} \left(\frac{3}{2} \right)^\alpha \chi_\beta,$$

où la dernière étape est justifiée par le fait que $y_1 \leq \frac{2}{3}$ et $0 \leq \chi_\beta \leq \frac{1}{2}$.
En réduisant au même dénominateur, on trouve

$$P_\beta f(x) \leq a x^{-\alpha} \frac{1 + \{\beta + [(\alpha-\beta) + 2^\alpha a^{-1}(\beta+2)]\}}{1+(1+\beta)\chi_\beta}$$

puisque $(1+\chi_\beta)^\alpha \leq 1 + \alpha\chi_\beta$ par concavité. On obtient le résultat voulu si $(\alpha-\beta) + 2^\alpha a^{-1}(\beta+2) \leq 1$, ce qui est le cas dès que

$$a \geq \frac{2^\alpha (2+\alpha)}{1-\alpha}.$$

□

On fixe de façon définitive un $a > 0$ tel que le lemme 2.2.4 est vérifié. On supposera aussi $a \geq 1$, de façon à avoir $\mathbb{1} \in \mathcal{C}_2$.

Corollaire 2.2.5. *Il existe $\delta = \delta(\alpha) > 0$ tel que pour toute suite (β_n) avec $0 < \beta_n \leq \alpha$, tout $n \geq 1$ et tout $x \in [0,1]$,*

$$P_1^n \mathbb{1}(x) \geq \delta.$$

Démonstration. D'après le lemme 2.4 de [104], toute fonction $f \in \mathcal{C}_2$ vérifie

$$f(x) \geq \min\left\{a, \left(\frac{\alpha(1+\alpha)}{a^\alpha}\right)^{\frac{1}{1-\alpha}}\right\} \int f dm$$

pour tout x. Le corollaire suit immédiatement avec $\delta(\alpha) = \min\left\{a, \left(\frac{\alpha(1+\alpha)}{a^\alpha}\right)^{\frac{1}{1-\alpha}}\right\}$ puisque $P_1^n \mathbb{1} \in \mathcal{C}_2$ pour tout n, d'après le lemme 2.2.4. □

Remarque 2.2.6. Par le même argument que pour le lemme 2.3 de [104], on peut montrer que l'application T_α admet une mesure de probabilité invariante absolument continue dont la densité appartient à $\mathcal{C}_2$.

2.2.3 Une perturbation aléatoire et son noyau associé

On introduit à présent une perturbation aléatoire qui consiste à remplacer un observable par son espérance conditionnelle sur une boule de rayon ϵ autour de chaque point. Ce nouveau système aléatoire, pour lequel le point fixe neutre est en quelque sorte inexistant, admet une vitesse de perte de mémoire exponentielle, et le résultat principal s'en déduit en estimant l'erreur commise lorsqu'on remplace le système (T_n) par le système aléatoire.
Pour $\epsilon > 0$, on définit l'opérateur $\mathbb{A}_\epsilon$ par

$$\mathbb{A}_\epsilon f(x) = \frac{1}{2\epsilon} \int_{B_\epsilon(x)} f\, dm,$$

pour tout $f \in L^1(m)$. On rappelle qu'on a identifié l'intervalle unité avec le cercle, les boules $B_\epsilon(x)$ sont donc à comprendre au sens de la distance géodésique sur le cercle.
On définit à présent un opérateur de transfert perturbé par

$$\mathbb{P}_{\epsilon,m} := P_{m+1}^{m+n_\epsilon} \circ \mathbb{A}_\epsilon = P_{\beta_{m+n_\epsilon}} \circ \ldots \circ P_{\beta_{m+1}} \circ \mathbb{A}_\epsilon,$$

où n_ϵ sera défini plus loin.

Lemme 2.2.7. *Pour toute suite (β_n) avec $\beta_n \leq \alpha$, tous $\epsilon > 0$, $m \geq 0$ et $f \in \mathcal{C}_2$, on a*

$$\|\mathbb{P}_{\epsilon,m} f - P_{m+1}^{m+n_\epsilon} f\|_{L_m^1} \leq c\|f\|_{L_m^1} \epsilon^{1-\alpha},$$

où c est une constante indépendante de (β_n).

Démonstration. Comme les opérateurs $P_{m+1}^{m+n_\epsilon}$ sont des contractions sur $L^1(m)$, on a $\|\mathbb{P}_{\epsilon,m}f - P_{m+1}^{m+n_\epsilon}f\|_{L^1_m} \leq \|\mathbb{A}_\epsilon f - f\|_{L^1_m} \leq c\|f\|_{L^1_m}\epsilon^{1-\alpha}$, d'après le lemme 3.1 de [104], ce qui termine la preuve. □

On a la représentation intégrale suivante de l'opérateur $\mathbb{P}_{\epsilon,m}$:

$$\mathbb{P}_{\epsilon,m}f(x) = \int_0^1 K_{\epsilon,m}(x,z)f(z)dz,$$

où le noyau $K_{\epsilon,m}$ est donné par

$$K_{\epsilon,m}(x,z) = \frac{1}{2\epsilon}P_{m+1}^{m+n_\epsilon}\mathbb{1}_{B_\epsilon(z)}(x).$$

Le point crucial réside dans le fait que ce noyau est borné uniformément loin de 0 pour un choix judicieux de n_ϵ.

Lemme 2.2.8. *Il existe $n_\epsilon = \mathcal{O}(\epsilon^{-\alpha})$ et $\gamma > 0$ tel que pour tout $\epsilon > 0$, tous $x, z \in [0,1]$ et toute suite (β_n) avec $0 \leq \beta_n \leq \alpha$, on a*

$$K_{\epsilon,m}(x,z) \geq \gamma.$$

On démontrera ce lemme à la section suivante. Voyons pour l'instant comment on peut en déduire le résultat principal sur la vitesse de perte de mémoire.

Proposition 2.2.9. *Soit $0 < \alpha < 1$. Alors pour toute suite (β_n) avec $0 \leq \beta_n \leq \alpha$, tout $n \geq 1$ et toute paire d'observables ϕ, ψ dans $\mathcal{C}_2$, de même intégrale $\int \phi dm = \int \psi dm$, on a*

$$\|P_1^n\phi - P_1^n\psi\|_{L^1_m} \leq C\left(\|\phi\|_{L^1_m} + \|\psi\|_{L^1_m}\right)\frac{(\log n)^{\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}-1}},$$

où C est une constante ne dépendant que de α.

Démonstration. On écrit $n = kn_\epsilon + m$, avec $0 \leq m < n_\epsilon$. On compare alors l'opérateur P_1^n avec l'opérateur $\mathbb{P}_{\epsilon,m+(k-1)n_\epsilon} \circ \ldots \circ \mathbb{P}_{\epsilon,m} \circ P_1^m$:

$$\begin{aligned}\|P_1^n\phi - P_1^n\psi\|_{L^1_m} &\leq \|\mathbb{P}_{\epsilon,m+(k-1)n_\epsilon} \circ \ldots \circ \mathbb{P}_{\epsilon,m}P_1^m(\phi-\psi)\|_{L^1_m} \\ &\quad + \|\left(\mathbb{P}_{\epsilon,m+(k-1)n_\epsilon} \circ \ldots \circ \mathbb{P}_{\epsilon,m} - P_{m+(k-1)n_\epsilon+1}^{m+kn_\epsilon} \circ \ldots \circ P_{m+1}^{m+n_\epsilon}\right)P_1^m(\phi-\psi)\|_{L^1_m} \\ &= (\mathrm{I}) + (\mathrm{II}).\end{aligned}$$

Afin d'alléger les notations, on posera $\mathbb{R}_j = \mathbb{P}_{\epsilon,m+(j-1)n_\epsilon}$ et $\mathbb{Q}_j = P_{m+(j-1)n_\epsilon+1}^{m+jn_\epsilon}$, pour $j = 1, \ldots, k$.
Pour contrôler le terme (I), on utilise le lemme suivant :

Lemme 2.2.10. *Soit (X, m) un espace de probabilité et $P : L^1(m) \to L^1(m)$ un opérateur donné par un noyau positif et borné $K : X^2 \to \mathbb{R}^+$:*

$$Pf(x) = \int_X K(x,y)f(y)dm(y),$$

pour tout $f \in L^1(m)$. Si P est markovien, c'est-à-dire que $\int_X K(x,y)dm(x) = 1$ pour tout $y \in X$ et si il existe $\gamma > 0$ tel que $K(x,y) \geq \gamma$ pour tout $x, y \in X$, alors

$$\|Pf\|_{L^1_m} \leq (1-\gamma)\|f\|_{L^1_m},$$

pour tout $f \in L^1(m)$ avec $\int_X f(x)dm(x) = 0$.

Démonstration. On pose $X^+ = \{f \geq 0\}$ et $\hat{X}^+ = \{Pf \geq 0\}$. Comme P est markovien, on a $\int_X Pf\, dm = \int_X f\, dm = 0$, et donc

$$\begin{aligned}
\|Pf\|_{L^1_m} = 2\int_{\hat{X}^+} Pf(x)dm(x) &= 2\int_{\hat{X}^+}\int_X (K(x,y)-\gamma)f(y)dm(y)dm(x) \\
&\leq 2\int_{X^+}\int_X (K(x,y)-\gamma)dm(x)f(y)dm(y) \\
&= 2(1-\gamma)\int_{X^+} f(y)dm(y) = (1-\gamma)\|f\|_{L^1_m},
\end{aligned}$$

où l'inégalité de la seconde ligne est justifiée par le fait que, après interversion des intégrales, $\int_{\hat{X}^+}(K(x,y)-\gamma)dm(x) \leq \int_X (K(x,y)-\gamma)dm(x)$ et $\int_X (K(x,y)-\gamma)dm(x)f(y)$ est négatif pour $y \in X \setminus X^+$, puisque $K(x,y)-\gamma$ est positif. □

Comme tous les opérateurs $\mathbb{R}_1, \ldots, \mathbb{R}_k$ satisfont les hypothèses du lemme précédent avec le même γ d'après le lemme 2.2.8, on obtient

$$\text{(I)} \leq (1-\gamma)^k \|P_1^m(\phi-\psi)\|_{L^1_m} \leq (1-\gamma)^k\|\phi-\psi\|_{L^1_m} \leq e^{-\gamma k}(\|\phi\|_{L^1_m} + \|\psi\|_{L^1_m}).$$

Pour le terme (II), on utilise une somme télescopique :

$$\begin{aligned}
\text{(II)} &\leq \sum_{j=1}^k \|\mathbb{R}_k \ldots \mathbb{R}_{j+1}(\mathbb{R}_j - \mathbb{Q}_j)\mathbb{Q}_{j-1}\ldots\mathbb{Q}_1 P_1^m(\phi-\psi)\|_{L^1_m} \\
&\leq \sum_{j=1}^k \|(\mathbb{R}_j - \mathbb{Q}_j)\mathbb{Q}_{j-1}\ldots\mathbb{Q}_1 P_1^m(\phi-\psi)\|_{L^1_m},
\end{aligned}$$

car les opérateurs $\mathbb{R}_j$ sont des contractions sur $L^1(m)$. Comme $\mathbb{Q}_{j-1}\ldots\mathbb{Q}_1 P_1^m\phi$ et $\mathbb{Q}_{j-1}\ldots\mathbb{Q}_1 P_1^m\psi$ appartiennent tous deux à $\mathcal{C}_2$ d'après le lemme 2.2.4, et que les opérateurs $\mathbb{Q}_{j-1}\ldots\mathbb{Q}_1 P_1^m$ sont aussi des contractions sur $L^1(m)$, on trouve que

$$\text{(II)} \leq c\epsilon^{1-\alpha}k(\|\phi\|_{L^1_m} + \|\psi\|_{L^1_m}),$$

d'après le lemme 2.2.7.
En conclusion, on obtient

$$\begin{aligned}\|P_1^n\phi - P_1^n\psi\|_{L_m^1} &\leq (e^{-\gamma k} + ck\epsilon^{1-\alpha})(\|\phi\|_{L_m^1} + \|\psi\|_{L_m^1}) \\ &\leq \left(e^{-\gamma\frac{n}{n_\epsilon}} + c\left(\frac{n}{n_\epsilon}+1\right)\epsilon^{1-\alpha}\right)(\|\phi\|_{L_m^1} + \|\psi\|_{L_m^1}) \\ &\leq Cn^{1-\frac{1}{\alpha}}(\log n)^{\frac{1}{\alpha}}(\|\phi\|_{L_m^1} + \|\psi\|_{L_m^1}),\end{aligned}$$

où l'on a pris $\epsilon = n^{-\frac{1}{\alpha}}\left(\log n^{(\frac{1}{\alpha}-1)\kappa}\right)^{\frac{1}{\alpha}}$, avec κ convenablement choisi en fonction de α et γ. □

On en déduit le théorème principal :

Démonstration du théorème 2.2.1. On utilise le même argument qu'à la fin de la preuve du théorème 4.1 de [104]. En effet, pour tout $\phi, \psi \in C^1$, on peut trouver $\lambda, \nu \in \mathbb{R}$ tels que $\phi_{\lambda,\nu}(x) = \phi(x) + \lambda x + \nu$ et $\psi_{\lambda,\nu}(x) = \psi(x) + \lambda x + \nu$ appartiennent à $\mathcal{C}_2$, la dépendance des paramètres λ et ν vis-à-vis de la somme des normes C^1 de ϕ et ψ étant affine. Par exemple, on peut choisir λ et ν de façon à satisfaire $\lambda \leq -\|\phi'\|_\infty$ et $\nu \geq \max\left\{\frac{(1+\alpha)\|\phi\|_\infty + \|\phi'\|_\infty - \lambda(2+\alpha)}{1+\alpha}, \frac{1+a}{a-1}\|\phi\|_\infty - \frac{a\lambda}{2(a-1)}\right\}$, et similairement pour ψ. Comme $\phi - \psi = \phi_{\lambda,\nu} - \psi_{\lambda,\nu}$ est d'intégrale nulle, par la proposition 2.2.9, on a

$$\begin{aligned}\|P_1^n\phi - P_1^n\psi\|_{L_m^1} = \|P_1^n\phi_{\lambda,\nu} - P_1^n\psi_{\lambda,\nu}\|_{L_m^1} &\leq C\frac{(\log n)^{\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}-1}}(\|\phi_{\lambda,\nu}\|_{L_m^1} + \|\psi_{\lambda,\nu}\|_{L_m^1}) \\ &\leq C\frac{(\log n)^{\frac{1}{\alpha}}}{n^{\frac{1}{\alpha}-1}}\mathcal{F}(\|\phi\|_{C^1} + \|\psi\|_{C^1}),\end{aligned}$$

pour une certaine fonction affine $\mathcal{F}$. □

2.2.4 Distorsion : démonstration du lemme 2.2.8

On aura besoin d'estimer la quantité suivante :

$$n'_\epsilon = \inf\left\{n \geq 1 \,:\, T_1^n(J) = [0,1], \forall J \text{ avec } |J| \geq 2\epsilon, \forall(\beta_n) \text{ avec } \beta_n \leq \alpha\right\}.$$

n'_ϵ est donc le temps minimal nécessaire pour un intervalle de longueur au moins 2ϵ pour recouvrir $[0,1]$, indépendamment de la composition d'applications choisie. La démonstration du lemme suivant est reportée à la fin de la section.

Lemme 2.2.11. *On a $n'_\epsilon = \mathcal{O}(\epsilon^{-\alpha})$.*

Démonstration du lemme 2.2.8. La principale difficulté de la démonstration du théorème 2.2.1 consiste à vérifier la positivité du noyau $K_{\epsilon,m}$. Pour cela, on va suivre la démarche de la preuve de la proposition 3.3 de [104].

On rappelle que

$$\begin{aligned}2\epsilon K_{\epsilon,m}(x,z) &= P_{m+1}^{m+n_\epsilon}\mathbb{1}_{B_\epsilon(z)}(x)\\ &= \sum_{T_{m+1}^{m+n_\epsilon}y=x}\frac{\mathbb{1}_{B_\epsilon(z)}(y)}{|(T_{m+1}^{m+n_\epsilon})'(y)|}\\ &\geq \mathbb{1}_{T_{m+1}^{m+n_\epsilon}(B_\epsilon(z))}(x)\inf_{y\in B_\epsilon(z)}\frac{1}{|(T_{m+1}^{m+n_\epsilon})'(y)|}.\end{aligned}$$

On doit donc estimer le temps n_ϵ mis par un intervalle J quelconque de longueur 2ϵ pour recouvrir $[0,1]$, tout en bornant inférieurement le ratio $\inf_{y\in J}\frac{1}{|(T_{m+1}^{m+n_\epsilon})'(y)|}$, et ce indépendamment de la suite (β_n). On peut donc sans perdre de généralité supposer que $m=0$.
On comme par introduire quelques notations supplémentaires. On pose $a_n^\alpha = T_{\alpha,1}^{-n}(1)$, où $T_{\alpha,1}^{-1}$ l'inverse de T_α restreinte à $[0,\frac{2}{3}]$.
Des calculs standards, voir par exemple le lemme 3.2 dans [104] ou la discussion avant le lemme 5 dans [144], montrent que a_n vérifie $a_n \approx \frac{1}{n^{\frac{1}{\alpha}}}$. En particulier, il existe une constante C telle que $a_n \leq C\frac{1}{n^{\frac{1}{\alpha}}}$. On définit aussi a_n^0 comme la suite des préimages de 1 par la branche gauche de l'application

$$T_0(x) = \begin{cases}\frac{3x}{2},\ 0\leq x\leq 2/3\\ 3x-2,\ 2/3\leq x\leq 1.\end{cases}$$

Pour $k\geq 1$, on définit la suite a_n^k par $a_0^k=1$, et $a_n^k = T_{k+1,1}^{-1}\circ\dots T_{k+n,1}^{-1}(1)$, c'est-à-dire que a_n^k est la préimage de 1 par T_{k+1}^{k+n} la plus proche de 0. On vérifie facilement que a_n^k est une suite strictement décroissante en n, et que $a_n^0\leq a_n^k\leq a_n^\alpha$.
On définit les intervalles $I_n^k=[a_{n+1}^k,a_n^k]$, qui vérifient $T_{k+1}^{k+n}I_n^k=[\frac{2}{3},1]$. On pose aussi $I_{n,+}^k = I_{n+1}^k\cup I_n^k = [a_{n+2}^k,a_n^k]$.
On définit la région intermittente $I=[0,a_2^0]$ et la région hyperbolique $H=[a_2^0,1]$.
Soit J un intervalle de longueur au moins 2ϵ. On va considérer les itérés successifs T_1^kJ jusqu'à ce que l'un d'entre eux recouvre l'espace entier, et on contrôle la distorsion durant ce temps.
Au temps $k\geq 1$, l'itéré $K=T_1^kJ$ vérifie l'une des conditions suivantes :

1. $K\cap I=\emptyset$;
2. $K\cap I\neq\emptyset$, et K contient au plus un a_l^k, $l>2$;
3. $K\cap I\neq\emptyset$, et K contient plus d'un a_l^k, $l>2$.

Cas 1. On suppose que l'on est dans la situation 1. Alors soit l'un des itérés de K va rencontrer le point $\frac{2}{3}$ où les applications ne sont plus dérivables, soit il vérifie la condition 2 ou 3.
Soit $n\geq 1$ le temps nécessaire pour rencontrer un des cas de figures. Par des arguments standards, on a pour tous $a,b\in K$:

$$\frac{(T_{k+1}^{k+n})'(a)}{(T_{k+1}^{k+n})'(b)} \leq \exp\left\{\sum_{l=0}^{n-1} \frac{\sup_\xi |T''_{k+n-l}\xi|}{\inf_\xi |T'_{k+n-l}\xi|} |T_{k+1}^{n+k-l-1}a - T_{k+1}^{n+k-l-1}b|\right\}.$$

Comme $0 < \beta \leq \alpha < 1$, le rapport $\frac{|T_\beta^{''}x|}{|T_\beta^{'}x|}$ et la quantité $|T_\beta^{'}x|$ sont majorés uniformément en β et $x \in H$ respectivement par $D > 0$ et $0 < r < 1$. On a donc

$$\frac{(T_{k+1}^{k+n})'(a)}{(T_{k+1}^{k+n})'(b)} \leq \exp\left\{c_1 |T_{k+1}^{k+n}(K)|\right\},$$

où $c_1 = \frac{D}{1-r}$ ne dépend que de α, et non de la concaténation particulière considérée. Après intégration par rapport à b, on trouve $|(T_{k+1}^{k+n})'(a)| \leq \frac{|T_{k+1}^{k+n}(K)|}{|K|} \exp\left\{-c_1 |T_{k+1}^{k+n}(K)|\right\}$, d'où

$$P_{k+1}^{k+n} \mathbb{1}_K \geq \mathbb{1}_{T_{k+1}^{k+n}(K)} \frac{|K|}{|T_{k+1}^{k+n}(K)|} \exp\left\{-c_1 |T_{k+1}^{k+n}(K)|\right\}.$$

Si le nouvel itéré de K intersecte la région intermittente, alors on considère la situation 2 ou 3 correspondante, et on continue la procédure. S'il est toujours dans la région hyperbolique, mais contient à présent le point $\frac{2}{3}$, on procède de la façon suivante. Soit $L = T_{k+1}^{k+n}(K)$ le nouvel itéré et soient L_g et L_d les parties de cet intervalle respectivement à gauche et à droite de $\frac{2}{3}$. On a soit $|L_l| > \frac{1}{3}|L|$ ou $|L_r| > \frac{1}{3}|L|$.
Dans le premier cas, après une itération, l'image de L_g sera contenue dans $[\frac{2}{3}, 1]$, avec l'extrémité droite égale à 1. Après m étapes, il couvrira l'intervalle unité, et la distorsion est bien contrôlée durant ces $m+1$ itérations, car les applications sont dérivables sur les itérés de L_g. On a donc

$$\begin{aligned} P_{k+n+1}^{k+n+m+1} \mathbb{1}_L \geq P_{k+n+1}^{k+n+m+1} \mathbb{1}_{L_l} &\geq \mathbb{1}_{T_{k+n+1}^{k+n+m+1}(L_l)} \frac{|L_l|}{|T_{k+n+1}^{k+n+m+1}(L_l)|} \exp\left\{-c_1 |T_{k+n+1}^{k+n+m+1}(L_l)|\right\} \\ &\geq \frac{1}{3}|L| \exp\{-c_1\}. \end{aligned}$$

On pose $n_1 = n + m + 1$ et on a

$$P_{k+1}^{k+n_1} \mathbb{1}_K = P_{k+n+1}^{k+n_1} P_{k+1}^{k+n} \mathbb{1}_K \geq P_{k+n+1}^{k+n_1} \mathbb{1}_L \frac{|K|}{|L|} \exp\{-c_1|L|\} \geq \frac{1}{3}|K| \exp\{-c_1 - c_1|L|\}.$$

Dans le second cas, si $|L_d| > \frac{1}{3}|L|$, alors après une itération (par une application affine), l'image de L_d sera de la forme $[0, x]$ et on poursuit alors avec l'étape 3 et on applique le contrôle sur la distorsion donnée à cette étape.
Dans ces deux cas, on obtient un facteur multiplicatif $\frac{1}{3}$, mais le premier cas amène à la fin de l'algorithme, et comme on le verra plus loin, l'étape 3 amène elle aussi à la fin de l'algorithme,

ce qui fait qu'on ne rencontrera au plus qu'une seule fois le point $\frac{2}{3}$. Ainsi, ce facteur $\frac{1}{3}$ n'apparait au plus qu'une seule fois, et ne pourra donc pas compromettre l'estimation finale.

Cas 2. On suppose que K est inclus dans un intervalle $I^k_{l,+}$. Comme $T^{k+l}_{k+1}(I^k_{l,+}) = [a_2^{k+l}, 1] \subset [a_2^0, 1]$, après exactement l itérations, l'image de K sera incluse dans la région hyperbolique, et on continue alors avec le cas 1. Durant ce temps, la distorsion sera contrôlée grâce au lemme de Koebe, dont on rappelle l'énoncé, ou plutôt une de ses conséquences :

Lemme 2.2.12 (Lemme de Koebe [49, Théorème IV.1.2])**.** *Pour tout $\tau > 0$, il existe une constante $C = C(\tau) > 0$ telle que si g un difféomorphisme croissant de classe C^3 dont la dérivée Schwarzienne est négative, c'est à dire*

$$\frac{g'''(x)}{g'(x)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)}\right)^2 \leq 0,$$

si $J_1 \subset J_2$ sont deux intervalles tels que $g(J_2)$ contient un voisinage τ-échelonné de $g(J_1)$ (c'est-à-dire que les intervalles à la droite et à la gauche de $g(J_1)$ dans $g(J')$ sont de longueur au moins $\tau|g(J_2)|$), alors pour tout $x, y \in J_1$,

$$\frac{g'(x)}{g'(y)} \leq \exp\left\{C\frac{|g(x) - g(y)|}{|g(J_1)|}\right\}.$$

On va appliquer le lemme de Koebe à g définie comme l'extension analytique à $]0, +\infty[$ de T^{k+l}_{k+1}, avec $J_1 = I^k_{l,+}$ et $J_2 = [\delta, 2]$, où $\delta = \delta(k, l)$ est choisi suffisamment petit de telle sorte que $\delta < a^k_{l+2}$ et $T^l_0\delta < \frac{1}{2}a^0_2$. Clairement, chaque (extension analytique à $]0, +\infty[$ de) T_β a une dérivée Schwarzienne négative, car $T'''_\beta < 0$ et $T'_\beta > 0$, et par conséquent g a une dérive Schwarzienne négative, comme composée d'applications à dérivées Schwarzienne négatives.
On a $g(J_1) = [a_2^{k+l}, 1] \subset [a^0_2, 1]$ et $g(J_2) = [T^{k+l}_{k+1}\delta, g(2)] \supset [T^l_0\delta, 2] \supset [\frac{1}{2}a^0_2, 2]$.
On pose $\tau = \min\left\{\frac{a^0_2}{2(1-a^0_2)}, \frac{1}{1-a^0_2}\right\}$ qui ne dépend pas de la concaténation d'applications ou du nombres d'étapes l.
L'intervalle à la gauche de $g(J_1)$ dans $g(J_2)$ contient $[\frac{1}{2}a^0_2, a^0_2]$ et est donc de longueur plus grande que

$$\frac{1}{2}a^0_2 \geq \tau(1 - a^0_2) \geq \tau|g(J_1)|.$$

De façon similaire, l'intervalle à la droite de $g(J_1)$ dans $g(J_2)$ contient $[1, 2]$ et a donc une longueur plus grande que

$$1 \geq \tau(1 - a^0_2) \geq \tau|g(J_1)|.$$

On a démontré que $g(J_2)$ contient un voisinage τ-échelonné de $g(J_1)$, donc d'après le lemme de Koebe, il existe $C = C(\tau)$ telle que pour tout $a, b \in J_1$, on a

$$\frac{(T^{k+l}_{k+1})'(a)}{(T^{k+l}_{k+1})'(b)} \leq \exp\left\{C\frac{|T^{k+l}_{k+1}(a) - T^{k+l}_{k+1}(b)|}{|T^{k+l}_{k+1}(J_1)|}\right\} \leq \exp\left\{c_2|T^{k+l}_{k+1}(a) - T^{k+l}_{k+1}(b)|\right\},$$

avec $c_2 = 3C(\tau)$, puisque $T_{k+1}^{k+l}(J_1) \supset [\frac{2}{3}, 1]$.
Comme $K \subset J_1$, on a pour tout a, b dans K

$$\frac{(T_{k+1}^{k+l})'(a)}{(T_{k+1}^{k+l})'(b)} \leq \exp\left\{c_2 |T_{k+1}^{k+l}(K)|\right\},$$

et donc

$$P_{k+1}^{k+l} \mathbb{1}_K \geq \mathbb{1}_{T_{k+1}^{k+l}(K)} \frac{|K|}{|T_{k+1}^{k+l}(K)|} \exp\left\{-c_2 |T_{k+1}^{k+l}(K)|\right\}.$$

Cas 3. On suppose que K intersecte la région intermittente, et contient au moins deux a_l^k. Si plus d'un tiers de K est contenu dans $[\frac{2}{3}, 1]$ et est de la forme $[a, 1]$, alors on peut considérer $K \cap [\frac{2}{3}, 1]$, et le cas 1 aura lieu jusqu'à ce que l'image de cet intervalle recouvre l'intervalle unité, et on perd un facteur $\frac{1}{3}$ (seulement une fois). Autrement, on définit l_- comme le plus petit entier tel que $I_{l_-}^k$ soit inclus dans K. On a $l_- = \mathcal{O}(|K|^{-\alpha})$ d'après le lemme 2.2.11, puisque l_- est le temps nécessaire à K pour recouvrir $[0, 1]$.
On distingue deux cas, selon que la partie de K qui se trouve à droite de $a_{l_-}^k$ soit de longueur au moins $\frac{|K|}{3}$ ou non.
Dans le premier cas, on pose $K' = K \cap [a_{l_-+1}^k, 1]$, qui vérifie $|K'| \geq \frac{|K|}{3}$. Comme $K' \subset I_{l_--1,+}^k$ et $T_{k+1}^{k+l_--1}(K') \supset [a_2^{k+l_--1}, a_1^{k+l_--1}]$, on a d'après le cas 2 :

$$P_{k+1}^{k+l_--1} \mathbb{1}_K \geq \frac{1}{3} \mathbb{1}_{[a_2^{k+l_--1}, a_1^{k+l_--1}]} |K| e^{-c_2}.$$

Comme $T_{k+l_-}^{k+l_-+1}([a_2^{k+l_--1}, a_1^{k+l_--1}]) = [0, 1]$, et que $|T'_\beta(x)|$ est borné supérieurement uniformément en β et x par M, on trouve

$$P_{k+1}^{k+l_-+1} \mathbb{1}_K \geq \frac{1}{3M^2} |K| e^{-c_2}.$$

Dans le second cas, on choisit K' de telle sorte que $|K'| \geq \frac{|K|}{3}$ et que l'extrémité droite de K' soit $a_{l_-}^k$ et son extrémité gauche se trouve à droite de 0. On subdivise K' en sous-intervalles $I_{l_-}^k, \ldots, I_{l_+}^k$ de telle sorte à ce que leur union soit de longueur plus grande que $\frac{|K'|}{3} = \frac{|K|}{9}$ avec l_+ minimimal. Ce choix de découper K' plutôt que K va permettre d'estimer l_+. En effet, si on pose $K = [a, a_{l_-}^k]$, comme l_+ a été choisi minimal, on a que $[a, a_{l_+-1}^k]$ est de longueur au moins $\frac{2|K'|}{3} \geq \frac{2|K|}{9}$. Ainsi, $\frac{C}{(l_+-1)^{\frac{1}{\alpha}}} \geq a_{l_+-1}^k \geq a_{l_+-1}^k - a \geq \frac{2|K|}{9}$ et par conséquent $l_+ = \mathcal{O}(|K|^{-\alpha})$.
Grâce à l'étude menée pour le cas 2, on a

$$\begin{aligned} P_{k+1}^{k+l_++1} \mathbb{1}_K &\geq \sum_{l=l_-}^{l_+} P_{k+1}^{k+l_++1} \mathbb{1}_{I_l^k} = \sum_{l=l_-}^{l_+} P_{k+l+1}^{k+l_++1} P_{k+1}^{k+l} \mathbb{1}_{I_l^k} \\ &\geq \sum_{l=l_-}^{l_+} P_{k+l+1}^{k+l_++1} \mathbb{1}_{T_{k+1}^{k+l}(I_l^k)} \frac{|I_l^k|}{|T_{k+1}^{k+l}(I_l^k)|} \exp\left\{-c_2 |T_{k+1}^{k+l}(I_l^k)|\right\}. \end{aligned}$$

Comme $T_{k+1}^{k+l}(I_l^k) = [\frac{2}{3}, 1]$, qui est envoyé après une itération sur l'espace entier, on a grâce au corollaire 2.2.5

$$P_{k+1}^{k+l_+ +1}\mathbb{1}_K \geq \sum_{l=l_-}^{l_+} P_{k+l+1}^{k+l_+ +1}\mathbb{1}_{[\frac{2}{3},1]}\frac{|I_l^k|}{1/3}\exp\left\{-\frac{c_2}{3}\right\} \geq \frac{c_3}{3}\exp\left\{-\frac{c_2}{3}\right\}\sum_{l=l_-}^{l_+}|I_l^k| \geq \frac{\delta}{27}\exp\left\{-\frac{c_2}{3}\right\}|K|,$$

où δ est la constante donnée par le corollaire 2.2.5, puisque $P_{k+l+1}^{k+l_+ +1}\mathbb{1}_{[\frac{2}{3},1]} \geq \frac{1}{3}P_{k+l+2}^{k+l_+ +1}\mathbb{1} \geq \frac{\delta}{3}$.

Conclusion. Soit J un intervalle de longueur au moins 2ϵ. On associe à J une suite d'entiers $n_1, m_1, n_2, m_2 \ldots, n_p$ telle que pour n_1 étapes, les itérées de J soient dans la région hyperbolique (avec éventuellement $n_1 = 0$) et vérifient donc le cas 1 , puis qu'ensuite pour m_1 étapes, on se trouve dans la situation du cas 2, puis qu'ensuite on se retrouve à nouveau dans la situation 1 pour n_2 étapes (on rappelle qu'après la situation 2 on peut seulement sortir de l'algorithme ou aller dans la situation 1), et ainsi de suite, jusqu'à ce qu'un itéré de J rencontre $\frac{2}{3}$, ou satisfasse le cas 3. Ces deux cas de figure mènent à la fin de l'algorithme. On ne traitera que le cas où 3 est vérifié, dans la situation où la partie de K à gauche de $a_{l_-}^k$ n'est pas de longueur plus grande que $\frac{|K|}{3}$, les autres cas se traitant de façon similaire.
Pour $n \geq n_1 + m_1 + \ldots + n_p + l_+ + 1$, on a

$$\begin{aligned}
P_1^n\mathbb{1}_J &\geq P_{n_1+\ldots+l_++2}^n P_{n_1+m_1+\ldots+n_p+1}^{n_1+\ldots+n_p+l_++1}\ldots P_{n_1+1}^{n_1+m_1}P_1^{n_1}\mathbb{1}_J \\
&\geq P_{n_1+\ldots+l_++2}^n P_{n_1+m_1+\ldots+n_p+1}^{n_1+\ldots+n_p+l_++1}\ldots P_{n_1+1}^{n_1+m_1}\mathbb{1}_{T_1^{n_1}(J)}\frac{|J|}{|T_1^{n_1}(J)|}e^{-c_1|T_1^{n_1}(J)|} \\
&\geq P_{n_1+\ldots+l_++2}^n P_{n_1+m_1+\ldots+n_p+1}^{n_1+\ldots+n_p+l_++1}\ldots \mathbb{1}_{T_1^{n_1+m_1}(J)}\frac{|T_1^{n_1}(J)|}{|T_1^{n_1+m_1}(J)|}\frac{|J|}{|T_1^{n_1}(J)|}e^{-c_1|T_1^{n_1}(J)|-c_2|T_1^{n_1+m_1}(J)|}
\end{aligned}$$

Par récurrence on a donc

$$\begin{aligned}
P_1^n\mathbb{1}_J &\geq (P_{n_1+\ldots+l_++2}^n\mathbb{1})\frac{\delta}{27}|T_1^{n_1+\ldots+n_p}(J)|\frac{|T_1^{n_1+\ldots+m_{p-1}}(J)|}{|T_1^{n_1+\ldots+n_p}(J)|}\cdots\frac{|T_1^{n_1}(J)|}{|T_1^{n_1+m_1}(J)|}\frac{|J|}{|T_1^{n_1}(J)|} \\
&\quad\times e^{-c_1|T_1^{n_1}(J)|-c_2|T_1^{n_1+m_1}(J)|-\ldots-c_1|T_1^{n_1+\ldots+n_p}(J)|-\frac{c_2}{3}} \\
&\geq (P_{n_1+\ldots+l_++2}^n\mathbb{1})\frac{\delta}{27}|J|e^{-2(\frac{c_2}{3}+c_1)(r^{n_2+\ldots+n_p}+\ldots+r^{n_2})} \\
&\geq \frac{\delta^2}{27}e^{-\frac{2(\frac{c_3}{3}+c_1)r}{1-r}}|J| =: \gamma|J|.
\end{aligned}$$

D'autre part, on a $n_1 + m_1 + \ldots + n_p + l_+ + 1 = \mathcal{O}(\epsilon^{-\alpha})$. En effet, en $n_1 + m_1 + \ldots + n_p$ itérations, l'intervalle J n'a pas encore recouvert $[0, 1]$, et donc $n_1+m_1\ldots+n_p \leq n'_\epsilon = \mathcal{O}(\epsilon^{-\alpha})$ par le lemme 2.2.11. D'autre part, on a vu au cours du cas 3 que $l_+ = \mathcal{O}(|K|^{-\alpha}) = \mathcal{O}(\epsilon^{-\alpha})$, ce qui termine la démonstration. □

Démonstration du lemme 2.2.11. On doit estimer le temps minimal n'_ϵ nécessaire pour qu'un intervalle J de longueur au moins 2ϵ recouvre l'intervalle $[0,1]$. On va examiner toutes les situations possibles. On notera tout d'abord que toutes les estimations qui vont suivre sont indépendantes de l'intervalle J et du choix des applications composées, à condition que chaque β soit dans $[0,\alpha]$, et que donc, si à partir d'un cas donné, on est renvoyé après itération à un cas traité auparavant, on pourra appliquer l'estimation déjà établie au nouvel itéré de J (qui pourra avoir une longueur différente). On rappelle qu'après une itération, l'image de J est toujours de longueur strictement supérieure à celle de J.
On supposera que J est de longueur au moins ϵ.

(a) $J \supset [0,\epsilon]$: il existe $l \geq 1$ tel que $a^1_{l+1} \leq \epsilon \leq a^1_l$. Ainsi, $T_1^{l+1}J = [0,1]$, et comme $a^1_l \leq \frac{C}{l^{\frac{1}{\alpha}}}$, on obtient $l = \mathcal{O}(\epsilon^{-\alpha})$.

(b) $J \supset [1-\epsilon,\epsilon]$: clairement, après $\mathcal{O}(\log\frac{1}{\epsilon})$ étapes, l'image de J recouvrira l'intervalle $[0,1]$.

(c) $J \supset [1-\delta,1] \cup [0,\delta']$: si $\delta \geq \frac{\epsilon}{2}$, après $\mathcal{O}(\log\frac{1}{\epsilon})$ itérations, l'image de J recouvrira $[0,1]$. Si $\delta' \geq \frac{\epsilon}{2}$, alors on est dans la situation du cas a, avec un intervalle de longueur au moins $\frac{\epsilon}{2}$, et donc en moins de $\mathcal{O}(\epsilon^{-\alpha})$ étapes on recouvrira $[0,1]$.

(d) $J \supset [\delta,\delta+\epsilon]$, avec $\delta \leq \frac{2}{3} \leq \delta+\epsilon$: après une itération, on se retrouve dans la situation du cas a ou b, avec un intervalle de longueur au moins $\frac{\epsilon}{2}$, et donc on recouvrira $[0,1]$ en moins de $\mathcal{O}(\epsilon^{-\alpha})$ étapes.

(e) $J \subset]0,\frac{2}{3}[$ et $J = [a,b]$ contient au moins deux a^1_l, avec $l \geq 2$: on choisira l comme le plus petit entier tel que $[a^1_{l+1}, a^1_l] \subset J$. En particulier, on a $a \leq a^1_{l+1} < a^1_l \leq b < a^1_{l-1}$. On distingue deux sous-cas : soit plus d'un tiers de J se trouve à droite de a^1_l, soit plus d'un tiers de J se trouve à gauche de a^1_l. Dans le premier sous-cas, on pose $\tilde{J} = J \cap [a^1_l, a^1_{l-1}]$. Après $l-1$ itérations de $\tilde{J}$, on est dans la situation du cas d avec un intervalle de longueur au moins $\frac{\epsilon}{3}$, qui se termine donc en $\mathcal{O}(\epsilon^{-\alpha})$ étapes. Pour estimer $l-1$, on remarque que $\frac{\epsilon}{3} \leq \frac{|J|}{3} \leq |\tilde{J}| \leq a^1_{l-1} - a^1_l \leq a^1_{l-1} \leq \frac{C}{(l-1)^{\frac{1}{\alpha}}}$, et donc $l-1 = \mathcal{O}(\epsilon^{-\alpha})$. Dans le second sous-cas, qui nécessite plus de travail, on pose $\tilde{J} = J \cap [0,a^1_l] = [a,a^1_l]$. Après l itérations, $T_1^l\tilde{J}$ contiendra $T_1^l I^1_l = [\frac{2}{3},1]$, et donc il faut $l+1$ itérations pour recouvrir $[0,1]$. Afin d'estimer l, on pose $\hat{J} = [0,|\tilde{J}|]$, qui est de même longueur que $\tilde{J}$, et on choisit k tel que $a^1_{k+1} \leq |\hat{J}| \leq a^1_k$. Comme pour le cas a, on a $k = \mathcal{O}(|\hat{J}|^{-\alpha}) = \mathcal{O}(\epsilon^{-\alpha})$. On prétend que $l < k$, ce qui conclut le sous-cas. On remarque que $[0,a^1_i]$ est le domaine de définition d'une branche régulière de T_1^i, et donc pour chaque $i = 0,\dots l$, $\tilde{J}$ et $\hat{J}$ sont inclus dans le domaine de définition d'une même branche de T_1^i. Comme chaque branche des applications T_β est croissante et convexe, et que la somme de ces deux propriétés est stable par composition, on obtient que chaque branche de T_1^i est convexe, et que donc sa dérivée y est croissante. Pour tout $i = 0,\dots,l$ et tout $x \in \tilde{J}$, on a alors $(T_1^i)'(x) \geq (T_1^i)'(x-a)$, et après intégration sur $\tilde{J}$, on trouve $|T_1^i\tilde{J}| \geq |T_1^i\hat{J}|$. Or, $|T_1^i\tilde{J}| < 1$ pour tout $i = 0,\dots,l-1$, et $|T_1^{k+1}\hat{J}| = 1$, et on obtient donc $k > l$.

(f) $J \subset]0,\frac{2}{3}[$, et J contient exactement un seul a^1_l, $l \geq 2$: après l itérations, on se trouve

dans la situation du cas c qui se termine en $\mathcal{O}(\epsilon^{-\alpha})$ étapes. Pour estimer l, on remarque que $J \subset [a^1_{l+1}, a^1_{l-1}]$, et donc $\epsilon \leq |J| \leq a^1_{l-1} - a^1_{l+1} \leq a^1_{l-1} \leq \frac{C}{(l-1)^{\frac{1}{\alpha}}}$, et donc $l = \mathcal{O}(\epsilon^{-\alpha})$.

(g) $J \subset]0, \frac{2}{3}[$ et J ne contient aucun terme a^1_l : on a donc $J \subset]a^1_{l+1}, a^1_l[$ pour un certain $l \geq 1$. Après, l itérations, on aura $T_1^l J \subset]\frac{2}{3}, 1[$, qui est traité au cas h. Pour estimer l, on remarque comme au cas précédent que $\epsilon \leq |J| \leq a^1_l - a^1_{l+1} \leq a^1_l \leq \frac{C}{l^{\frac{1}{\alpha}}}$ et donc $l = \mathcal{O}(\epsilon^{-\alpha})$.

(h) $J \subset]\frac{2}{3}, 1[$: après $\mathcal{O}(\log \frac{1}{\epsilon})$ étapes, l'image de J va rencontrer $]0, \frac{2}{3}[$, et se retrouver dans un des cas détaillés auparavant. L'image de J sera de longueur au moins 3ϵ.

Les seules situations où l'étude n'a pas été menée à son terme sont g et h. Le seul cas de figure non couvert par cette analyse correspond au cas où l'on commence dans le cas g ou h et que l'on rentre dans une boucle entre ces deux cas. On finira par sortir de cette boucle, car la longueur de l'itérée de J croit strictement avec le temps, et on se retrouvera alors dans un cas étudié auparavant, et on couvre ensuite $[0, 1]$ en $\mathcal{O}(\epsilon^{-\alpha})$ étapes. Il faut donc estimer le nombre d'étapes que l'on effectue au sein de la boucle. On remarque qu'un intervalle de longueur au moins ϵ effectue $\mathcal{O}(\epsilon^{-\alpha})$ étapes au cours d'un seul passage dans la boucle, grâce à l'étude déjà effectuée. Or, en passant une fois dans la boucle, la longueur de l'intervalle est multipliée par un facteur au moins 3. Le nombre d'étapes effectuées au total dans cette boucle est donc de l'ordre de $\sum_{i \geq 0} \mathcal{O}((3^i \epsilon)^{-\alpha}) = \mathcal{O}(\epsilon^{-\alpha})$ (comme on l'a rappelé plus haut, les constantes intervenants implicitement dans les $\mathcal{O}$ sont uniformes vis-à-vis de l'intervalle et de la composition, et donc cette sommation est légitime). Ceci termine la démonstration. □

2.3 Remarques et questions ouvertes

1. Pour les systèmes non-autonomes unidimensionnels, avec perte de mémoire exponentielle, peut-on démontrer des théorèmes limites plus fins, comme une inégalité de Berry-Esseen, un principe des grandes déviations ou un théorème de la limite locale. Les deux premiers types de résultats sont présentés dans [74] pour des applications surjectives par morceaux.
2. Bien que nouveau, le résultat présenté au théorème 2.2.1 est relativement faible, et est susceptible d'être amélioré sur plusieurs points :
   (a) la vitesse de convergence vers 0 n'est certainement pas optimale, le facteur logarithmique $(\log n)^{\frac{1}{\alpha}}$ étant absent du résultat dans le cas autonome ;
   (b) la dépendance vis-à-vis de la norme C^1 des observables ϕ et ψ est faible, et on pourrait s'attendre à ce que le terme $\mathcal{F}(\|\phi\|_{C^1} + \|\psi\|_{C^1})$ soit remplacé par $\|\phi - \psi\|_{C^1}$.

   Une conjecture vraisemblable est la suivante :

   Conjecture 2.3.1. *Soit $0 < \alpha < 1$. Alors il existe $C > 0$ telle que pour toute suite (β_n) avec $0 < \beta_n \leq \alpha$, et tout ϕ de classe C^1 avec $\int \phi \, dm = 0$, on a*

$$\|P_1^n \phi\|_{L^1_m} \leq C \frac{1}{n^{\frac{1}{\alpha}-1}} \|\phi\|_{C^1}.$$

Il serait souhaitable de disposer d'une méthode plus performante que celle utilisée pour démontrer le théorème 2.2.1 afin d'établir cette conjecture. Il est difficile d'utiliser l'induction dans le contexte non-autonome, ce qui exclut la puissante méthode du renouvellement pour les opérateurs [128, 61, 60]. Il existe au moins deux possibilités à explorer :

(a) Etendre au cas non-autonome le cadre fonctionnel développé par Conze et Raugi [45] pour les systèmes autonomes, où les auteurs ont construit un formalisme basé sur une chaîne d'espaces fonctionnels et l'étude de l'action de l'opérateur de transfert sur ces derniers afin d'étudier des systèmes avec une vitesse de décroissance des corrélations non-exponentielle. Cette méthode a en particulier été utilisée pour des applications de type intermittente par Raugi [125].

(b) Etendre au cas non-autonome la méthode de la métrique de Hilbert utilisée par Maume-Deschamps [108, 109] dans le cas autonome.

3. Le théorème 2.2.1 entraine par un argument de densité que le système séquentiel correspondant est exact. Est-il possible de démontrer cela par un argument élémentaire, sans étudier auparavant la vitesse de perte de mémoire pour des observables réguliers ?

4. Que peut-on dire si β_n converge vers 0 ? Le théorème 2.2.1 entraine que la vitesse de perte de mémoire est super-polynomiale. Est-il néanmoins possible d'obtenir une vitesse de convergence meilleure pour des cas particuliers de suites (β_n), comme par exemple $\beta_n = \lambda^n$ avec $0 < \lambda < 1$. De même, que se passe-t-il si β_n converge vers 1 ? Le système séquentiel est-il toujours exact, et si oui, peut-on estimer la vitesse de perte de mémoire pour des suites (β_n) particulières ?

5. Peut-on démontrer des propriétés statistiques pour ces systèmes non-autonomes avec perte de mémoire polynomiale, comme par exemple la loi des grands nombres, le théorème de la limite centrale, des inégalités de grandes déviations, etc ?

Chapitre 3

Inégalités de concentration pour les systèmes dynamiques aléatoires et non-autonomes

Dans ce chapitre, on présentera la notion d'inégalités de concentration pour les systèmes dynamiques, avant de démontrer de telles inégalités pour des systèmes non-autonomes et aléatoires en dimension 1. La deuxième section sur les systèmes non-autonomes présente les résultats de l'article [6] écrit en collaboration avec Jérôme Rousseau, en cours de soumission, tandis que la troisième section expose le résultat sur les systèmes aléatoires issu du travail [3] coécrit avec Matthew Nicol et Sandro Vaienti, à paraître dans *Probability Theory and Related Fields*.

3.1 Inégalités de concentration : définition

Les inégalités de concentration permettent d'estimer la probabilité que des variables aléatoires de la forme $Y_n = K(Z_0, \ldots, Z_{n-1})$ dévient de leur espérance pour des fonctions K suffisamment régulières. Avant d'aller plus loin, on a besoin de préciser la régularité de K. Une possibilité est d'introduire les fonctions séparément Lipschitz :

Définition 3.1.1. *Soit (X, d) un espace métrique. Une fonction $K : X^n \to \mathbb{R}$ est dite séparément Lipschitz si il existe des constantes $\mathrm{Lip}_i(K)$, $i = 0, \ldots, n-1$, telles que*

$$|K(x_0, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_{n-1}) - K(x_0, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_{n-1})| \leq \mathrm{Lip}_i(K) d(x_i, x_i')$$

pour tous choix de $x_0, \ldots, x_{n-1}, x_i'$ dans X.

Afin d'estimer la probabilité qu'une variable aléatoire dévie de son espérance, on a besoin d'estimations sur ses moments centrés. La situation la plus favorable correspond à un contrôle sur les moments exponentiels :

Définition 3.1.2. *On dira qu'un processus stochastique $(Z_n)_{n\geq 0}$ à valeurs dans X satisfait une inégalité de concentration exponentielle si il existe $C > 0$ telle que pour tout $n \geq 1$ et toute fonction $K : X^n \to \mathbb{R}$ séparément Lipschitz, on a*

$$\mathbb{E}(e^{K(Z_0,\ldots,Z_{n-1})-\mathbb{E}(K(Z_0,\ldots,Z_{n-1}))}) \leq e^{C\sum_{j=0}^{n-1}\mathrm{Lip}_j^2(K)}.$$

Proposition 3.1.3. *Si (Z_n) satisfait une inégalité de concentration exponentielle avec la constante C, alors pour tout $n \geq 1$ et toute fonction $K : X^n \to \mathbb{R}$ séparément Lipschitz, on a*

$$\mathbb{P}\left(K(Z_0,\ldots,Z_{n-1}) - \mathbb{E}(K(Z_0,\ldots,Z_{n-1})) > t\right) \leq e^{-\frac{t^2}{4C\sum_{j=0}^{n-1}\mathrm{Lip}_j^2(K)}}$$

pour tout $t > 0$.

Démonstration. On pose $Y = K(Z_0,\ldots,Z_{n-1}) - \mathbb{E}(K(Z_0,\ldots,Z_{n-1}))$. Par l'inégalité de Markov, pour tout $\lambda > 0$, on a

$$\mathbb{P}(Y > t) = \mathbb{P}(e^{\lambda Y} > e^{\lambda t}) \leq e^{-\lambda t}\mathbb{E}(e^{\lambda Y}).$$

Comme λY est séparément Lipschitz, avec des constantes de Lipschitz données par $\lambda\mathrm{Lip}_i(K)$, $i = 0,\ldots,n-1$, on trouve

$$\mathbb{P}(Y > t) \leq e^{-\lambda t + C\lambda^2\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)}.$$

Une optimisation sur $\lambda > 0$ suggère de prendre $\lambda = \frac{t}{2C\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)}$ d'où

$$\mathbb{P}(Y > t) \leq e^{-\frac{t^2}{4C\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)}}.$$

□

On peut aussi seulement exiger un contrôle sur les moments d'ordre fini, pour des estimations plus faibles :

Définition 3.1.4. *On dira qu'un processus stochastique $(Z_n)_{n\geq 0}$ satisfait une inégalité de concentration polynomiale d'ordre $p \geq 2$ si il existe $C > 0$ telle que pour tout $n \geq 1$ et toute fonction $K : X^n \to \mathbb{R}$ séparément Lipschitz, on a*

$$\mathbb{E}\left|K(Z_0,\ldots,Z_{n-1}) - \mathbb{E}(K(Z_0,\ldots,Z_{n-1}))\right|^p \leq C\left(\sum_{i=0}^{n-1}\mathrm{Lip}_i^2(K)\right)^{\frac{p}{2}}.$$

Une application directe de l'inégalité de Markov donne cette fois-ci :

Proposition 3.1.5. *Si (Z_n) satisfait une inégalité de concentration polynomiale d'ordre $p \geq 2$ avec la constante C, alors pour tout $n \geq 1$ et toute fonction $K : X^n \to \mathbb{R}$ séparément Lipschitz, on a*

$$\mathbb{P}\left(K(Z_0, \ldots, Z_{n-1}) - \mathbb{E}(K(Z_0, \ldots, Z_{n-1})) > t\right) \leq C \frac{1}{t^p} \left(\sum_{i=0}^{n-1} \mathrm{Lip}_i^2(K) \right)^{\frac{p}{2}}.$$

Exemple 3.1.6. *L'exemple le plus élémentaire consiste à prendre des fonctionnelles additives. Si $\varphi : X \to \mathbb{R}$ est Lipschitz, on pose $K(x_0, \ldots, x_{n-1}) = \frac{1}{n} \sum_{k=0}^{n-1} \varphi(x_k)$. Alors K est séparément Lipschitz, avec $\mathrm{Lip}_k(K) \leq \frac{1}{n}\mathrm{Lip}(\varphi)$. Ainsi, si (Z_n) est un processus à valeurs dans X, et si $S_n = \sum_{k=0}^{n-1} Z_k$, on a pour tout $t > 0$ et tout $n \geq 1$:*

$$\mathbb{P}(S_n > nt) \leq e^{-\frac{t^2 n}{4C\mathrm{Lip}^2(\varphi)}},$$

si (Z_n) satisfait une inégalité de concentration exponentielle et

$$\mathbb{P}(S_n > nt) \leq C \frac{\mathrm{Lip}^p(\varphi)}{t^p n^{\frac{p}{2}}},$$

si (Z_n) satisfait une inégalité de concentration polynomiale d'ordre $p \geq 2$. On verra plus loin des exemples plus sophistiqués, qui justifient l'intérêt des inégalités de concentration, pour les systèmes dynamiques.

Si (Z_n) est une suite de variables aléatoires indépendantes à valeurs réelles, bornée dans L^p, avec $p \in [2, +\infty]$, alors elle satisfait une inégalité de concentration :

Théorème 3.1.7. *Soit (Z_n) une suite de variables aléatoires indépendantes, à valeurs réelles.*

1. *Si (Z_n) est bornée dans L^∞, alors elle satisfait une inégalité de concentration exponentielle.*
2. *Si (Z_n) est bornée dans L^p, $p \geq 2$, alors elle satisfait une inégalité de concentration polynomiale d'ordre p.*

Bien que le résultat soit classique, on en redonne la démonstration afin d'illustrer la stratégie qui sera réutilisée dans le cadre des systèmes dynamiques.

Démonstration. Soit $n \geq 1$ et $K : \mathbb{R}^n \to \mathbb{R}$ séparément Lipschitz. Pour $k = 0, \ldots, n-1$, on pose $\mathcal{F}_k = \sigma(Z_k, \ldots, Z_{n-1})$ et $\mathcal{F}_n = \{\emptyset, \Omega\}$.
Soit $Y = K(Z_0, \ldots, Z_{n-1})$, et $Y_k = \mathbb{E}(Y|\mathcal{F}_k)$. On a $Y_0 = Y$ et $Y_n = \mathbb{E}(Y)$. Par conséquent,

$$Y - \mathbb{E}(Y) = Y_0 - Y_n = \sum_{k=0}^{n-1} (Y_k - Y_{k+1}) = \sum_{k=0}^{n-1} H_k,$$

avec $H_k = Y_k - Y_{k+1}$. (H_k) forme une suite de différence de martingales inversée par rapport à la filtration décroissante $(\mathcal{F}_k)_k$.

Soit $(Z'_n)_n$ une suite de variables aléatoires, indépendante de $(Z_n)_n$, et de même loi (en tant que processus). Par indépendance des (Z_n), on a $Y_k = f_k(Z_0, \ldots, Z_{n-1})$, avec

$$f_k(x_0, \ldots, x_{n-1}) = \mathbb{E}(K(Z'_0, \ldots, Z'_{k-1}, x_k, \ldots, x_{n-1})).$$

Ainsi, $H_k = h_k(Z_0, \ldots, Z_{n-1})$, avec $h_k = f_k - f_{k+1}$ qui vérifie

$$|h_k(x_0, \ldots, x_{n-1})| \leq \mathrm{Lip}_k(K)\mathbb{E}|x_k - Z'_k|. \tag{3.1}$$

Si (Z_n) est bornée dans L^∞, on utilisera l'inégalité de Azuma-Hoeffding :

Proposition 3.1.8 (Inégalité de Azuma-Hoeffding [78, 13])**.** *Soit (X_k) une suite de différences de martingale inversée vérifiant $a_k \leq X_k \leq b_k$ pour tout k. Soit $S_n = \sum_{k=0}^{n-1} X_k$. Alors pour tout n, on a*

$$\mathbb{E}(e^{S_n}) \leq e^{\frac{1}{8}\sum_{k=0}^{n-1}(b_k - a_k)^2}.$$

En appliquant l'inégalité de Azuma-Hoeffding à (H_k), on obtient $\mathbb{E}(e^{Y-\mathbb{E}(Y)}) \leq e^{\frac{1}{2}\sum_{k=0}^{n-1}\sup|H_k|^2}$. Soit $M = \sup_n \|Z_n\|_\infty$. D'après 3.1, on a $|H_k| \leq 2M\mathrm{Lip}_k(K)$. Par conséquent

$$\mathbb{E}(e^{Y-\mathbb{E}(Y)}) \leq e^{2M^2\sum_{k=0}^{n-1}\mathrm{Lip}_k^2(K)},$$

d'où le résultat avec $C = 2M^2$ qui ne dépend ni de n ni de K.
Dans le cas où (Z_n) est bornée dans L^p, $p \geq 2$, on utilisera l'inégalité de Burkholder :

Proposition 3.1.9 (Inégalité de Burkholder [33])**.** *Pour tout $p \in]1, \infty[$, il existe une constante $C_p > 0$ telle que pour toute suite (X_k) de différences de martingale inversée, et tout n, on a*

$$C_p^{-1}\|D_n\|_p \leq \|S_n\|_p \leq C_p\|D_n\|_p,$$

où on a noté $S_n = \sum_{k=0}^{n-1} X_k$, et $D_n = \sqrt{\sum_{k=0}^{n-1} X_k^2}$.

On a donc

$$\begin{aligned}
\|Y - \mathbb{E}(Y)\|_p \leq C_p \left\| \sqrt{\sum_{k=0}^{n-1} H_k^2} \right\|_p &= C_p \left(\mathbb{E} \left| \sum_{k=0}^{n-1} H_k^2 \right|^{\frac{p}{2}} \right)^{\frac{1}{p}} \\
&= C_p \left(\left\| \sum_{k=0}^{n-1} H_k^2 \right\|_{\frac{p}{2}} \right)^{\frac{1}{2}} \\
&\leq C_p \left(\sum_{k=0}^{n-1} \|H_k^2\|_{\frac{p}{2}} \right)^{\frac{1}{2}}.
\end{aligned}$$

Bien que l'inégalité de Burkholder soit valable pour tout $p > 1$, on remarquera que l'on a utilisé l'inégalité triangulaire pour la norme $\|.\|_{\frac{p}{2}}$, d'où la nécessité de supposer que $p \geq 2$.

Soit $M = \sup_n \|Z_n\|_p$. D'après (3.1) et l'inégalité de Jensen, on a

$$\|H_k^2\|_{\frac{p}{2}} = (\mathbb{E}|H_k|^p)^{\frac{2}{p}} \leq 4M^2 \mathrm{Lip}_k^2(K),$$

d'où $\|Y - \mathbb{E}(Y)\|_p \leq 2C_p M \left(\sum_{k=0}^{n-1} \mathrm{Lip}_k^2(K)\right)^{\frac{1}{2}}$ et le résultat. □

Lorsque la suite (Z_n) n'est plus indépendante, alors la relation $Y_k = f_k(Z_0, \ldots, Z_{n-1})$, avec $f_k(x_0, \ldots, x_{n-1}) = \mathbb{E}(K(Z'_0, \ldots, Z'_{k-1}, x_k, \ldots, x_{n-1}))$ cesse d'être vérifiée, et on doit donc estimer l'erreur commise lorsque l'on remplace Y_k par $f_k(X_0, \ldots, X_{n-1})$.

Le cas des systèmes dynamiques. Soit à présent un système dynamique (X, T, μ), où (X, d) est un espace métrique, et μ une probabilité invariante par T.

Définition 3.1.10. *On dira que le système dynamique (X, T, μ) satisfait une inégalité de concentration exponentielle (resp. polynomiale d'ordre $p \geq 2$) si le processus défini par $Z_n = T^n$ sur l'espace probabilisé (X, μ) satisfait une inégalité de concentration exponentielle (resp. polynomiale d'ordre $p \geq 2$).*

Cette notion a été pour la première fois introduite par Collet, Martinez et Schmitt [44], qui ont étudié le cas des applications uniformément dilatantes par morceaux de l'intervalle :

Théorème 3.1.11 (Collet, Martinez, Schmitt, [44])**.** *Soit $T : [0, 1] \to [0, 1]$ une application de classe C^2 par morceaux, uniformément dilatante et topologiquement mélangeante. Soit μ l'unique mesure de probabilité invariante par T et absolument continue par rapport à Lebesgue. On suppose que la densité de μ est bornée uniformément loin de 0. Alors le système dynamique (X, T, μ) satisfait une inégalité de concentration exponentielle.*

De nombreux autres travaux ont suivi, visant essentiellement à étendre ce résultat à des systèmes non-uniformément dilatants ou hyperboliques. On donne le résultat de Chazottes et Gouëzel [39], qui en est une bonne synthèse, et qui concerne les systèmes pouvant être représentés par une tour de Young. On renvoit à [143, 144] pour une définition précise des tours de Young.

Théorème 3.1.12 (Chazottes, Gouëzel, [39])**.** *Soit (X, T, μ) un système dynamique représenté par une tour de Young (inversible ou non-inversible), de base Δ_0 et de temps de retour donné par R.*

1. *S'il existe $c > 0$ tel que $\int_{\Delta_0} e^{cR}\, d\mu < \infty$, alors (X, T, μ) satisfait une inégalité de concentration exponentielle.*
2. *Si $\int_{\Delta_0} R^p\, d\mu < \infty$ pour $p \geq 2$, alors (X, T, μ) satisfait une inégalité de concentration polynomiale d'ordre $2p - 2$.*

Dans le cas des systèmes dynamiques perturbés aléatoirement, on peut citer le résultat de Maldonado [106] sur le bruit observationnel.

3.2 Inégalités de concentration pour les systèmes dynamiques non-autonomes

3.2.1 Le résultat et sa démonstration

On reprend le cadre décrit à la section 2.1.

Définition 3.2.1. *On dira qu'un système dynamique séquentiel $(T_n)_{n\geq 1}$ sur un espace de probabilité (X, m), où (X, d) est un espace métrique, satisfait une inégalité de concentration exponentielle (resp. polynomiale) si le processus $Z_n = T_1^n$ défini sur (X, m) satisfait une inégalité de concentration exponentielle (resp. polynomiale).*

On considèrera le cas où $X = [0, 1]$ et m est la mesure de Lebesgue, et on supposera que les applications (T_n) appartiennent à une classe $\mathcal{C}$ de transformations qui satisfait les conditions suivantes :

Hypothèses sur la classe $\mathcal{C}$.
$\mathcal{C}$ est une classe d'applications de classe C^2 par morceaux sur $[0, 1]$ telle que :

1. Il existe $\lambda > 1$ et $M > 0$ tels que $\inf |T'| \geq \lambda$ et $\sup |T''| \leq M$ pour tout $T \in \mathcal{C}$.
2. Il existe $N \geq 1$ tel que $\sharp\mathcal{A}_T \leq N$ pour tout $T \in \mathcal{C}$, où $\mathcal{A}_T$ désigne la partition de monotonicité et régularité de T.
3. $\mathcal{C}$ satisfait $\boldsymbol{(D_r)}$ pour un certain $r \geq 1$ et **(Dec)**.

Le résultat principal est le suivant :

Théorème 3.2.2. *Soit $(T_n)_{n\geq 1}$ un système dynamique séquentiel appartenant à une classe $\mathcal{C}$ comme décrite ci-dessus. Si (T_n) satisfait la condition* **(Min)***, alors il vérifie une inégalité de concentration exponentielle.*

Ce résultat s'applique en particulier lorsque T_n est la β_n-transformation, avec tous les β_n dans un voisinage convenable d'un $\beta_\star > 1$ donné.
Afin de démontrer ce théorème, on va utiliser la méthode des différences bornées de McDiarmid [110, 111], comme cela a été fait dans [44] et [39], mais adaptée au contexte non-stationnaire.
On prolonge d'abord naturellement K en une fonction $K : X^{\mathbb{N}} \to \mathbb{R}$ qui ne dépend que des n premières coordonnées. $X^{\mathbb{N}}$ est muni de la tribu produit $\tilde{\mathcal{F}}$ et de la mesure de probabilité $\tilde{m}$ qui est l'image de m par l'application $\Phi : X \to X^{\mathbb{N}}$ définie par $\Phi(x) = (T_1^n x)_{n\geq 0}$, avec la convention $T_1^0 = \mathrm{Id}$. Avec cette notation,

$$\mathbb{E}_m(e^{K(Z_0,\ldots,Z_{n-1})-\mathbb{E}_m(K(Z_0,\ldots,Z_{n-1}))} = \mathbb{E}_{\tilde{m}}(e^{K-\mathbb{E}_{\tilde{m}}(K)}).$$

Soit $\tilde{\mathcal{F}}_p$ la tribu sur $X^{\mathbb{N}}$ engendrée par les coordonnées $(x_k)_{k\geq p}$ et $K_p = \mathbb{E}_{\tilde{m}}(K|\tilde{\mathcal{F}}_p)$. L'hypothèse **(Dec)** entraine que le système séquentiel (T_n) est exact, et par conséquent la tribu $\mathcal{F}_\infty$

est triviale modulo m. Ainsi, $\tilde{\mathcal{F}}_\infty := \cap_{p\geq 0}\tilde{\mathcal{F}}_p$ est triviale modulo $\tilde{m}$, puisque $\Phi^{-1}(\tilde{\mathcal{F}}_\infty) \subset \mathcal{F}_\infty$. D'après le théorème de convergence pour les martingales de Doob, on obtient que K_p converge $\tilde{m}$-p.p. vers $\mathbb{E}_{\tilde{m}}(K)$, et dans chaque L^p, avec $1 \leq p < \infty$. On a donc

$$K - \mathbb{E}_{\tilde{m}}(K) = \sum_{p\geq 0} D_p$$

avec $D_p = K_p - K_{p+1}$. En utilisant l'inégalité de Azuma-Hoeffding 3.1.8, on en déduit l'existence d'une constante $C > 0$ universelle telle que pour tout $P \geq 0$,

$$\mathbb{E}_{\tilde{m}}(e^{\sum_{p=0}^{P} D_p}) \leq e^{C\sum_{p=0}^{P} \sup|D_p|^2}.$$

Il reste à estimer D_p :

Proposition 3.2.3. *Il existe $0 < \rho < 1$ et $C > 0$ dépendant seulement de $(T_n)_{n\geq 1}$ telles que pour tout p et tout K, on a*

$$|D_p| \leq C \sum_{j=0}^{p} \rho^{p-j} \mathrm{Lip}_j(K).$$

Cette proposition, combinée avec l'inégalité de Cauchy-Schwarz, entraine l'inégalité de concentration, de la même manière que dans [39]. En effet, on a

$$\left(\sum_{j=0}^{p} \rho^{p-j}\mathrm{Lip}_j(K)\right)^2 \leq \left(\sum_{j=0}^{p} \rho^{p-j}\mathrm{Lip}_j^2(K)\right)\left(\sum_{j=0}^{p} \rho^{p-j}\right) \leq C\sum_{j=0}^{p} \rho^{p-j}\mathrm{Lip}_j^2(K).$$

Après sommation sur p, on obtient $\sum_{p=0}^{P} \sup|D_p|^2 \leq C\sum_j \mathrm{Lip}_j^2(K)$.
La proposition 3.2.3 découle immédiatement de la condition de Lipschitz sur K et du lemme suivant :

Lemme 3.2.4. *Il existe $0 < \rho < 1$ et $C > 0$ dépendant seulement de $(T_n)_{n\geq 1}$ telles que pour tout p, tout K, et $\tilde{m}$-p.t. $\underline{x} = (x_0, x_1, \ldots) \in X^{\mathbb{N}}$, on a*

$$\left|K_p(\underline{x}) - \int_X K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots)\, dm(y)\right| \leq C \sum_{j=0}^{p-1} \mathrm{Lip}_j(K)\rho^{p-j}.$$

Le reste de cette section est consacrée à la démonstration de ce lemme.
Premièrement, on remarque que

$$K_p(\underline{x}) = \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{T_1^p y = x_p} \frac{K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots)}{|(T_1^p)'(y)|}. \tag{3.2}$$

La démonstration de cette identité est renvoyée à la fin de cette section.

On fixe un $x_\star \in X$ et on décompose K_p comme

$$K_p(\underline{x}) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{i=0}^{p-1} \sum_{T_1^p y = x_p} \frac{H_i(y)}{|(T_1^p)'(y)|},$$

où $H_i(y) = K(y, \ldots, T_1^i y, x_\star, \ldots, x_\star, x_p, \ldots) - K(y, \ldots, T_1^{i-1} y, x_\star, \ldots, x_\star, x_p, \ldots)$.
En utilisant la règle de dérivation pour les fonctions composées, on trouve

$$K_p(\underline{x}) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{i=0}^{p-1} P_{i+1}^p f_i(x_p)$$

avec

$$f_i(y) = \sum_{T_1^i z = y} \frac{H_i(z)}{|(T_1^i)'(z)|} = P_1^i H_i(y).$$

Comme $\int f_i \, dm = \int H_i \, dm$, on a

$$\sum_{i=0}^{p-1} \int f_i \, dm = \int K(y, \ldots, T_1^{p-1} y, x_p, \ldots) \, dm(y) - K(x_\star, \ldots, x_\star, x_p, \ldots).$$

Par conséquent,

$$\begin{aligned} & K_p(\underline{x}) - \int K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots) \, dm(y) \\ = \; & \sum_{i=0}^{p-1} \left(\frac{P_{i+1}^p \mathbb{1}(x_p)}{P_1^p \mathbb{1}(x_p)} - 1 \right) \int f_i \, dm + \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{i=0}^{p-1} \left(P_{i+1}^p f_i(x_p) - \left(\int f_i \, dm \right) P_{i+1}^p \mathbb{1}(x_p) \right). \end{aligned}$$

Pour le premier terme, on a

$$\left| \frac{P_{i+1}^p \mathbb{1}(x_p)}{P_1^p \mathbb{1}(x_p)} - 1 \right| = \frac{|P_{i+1}^p \mathbb{1}(x_p) - P_1^p \mathbb{1}(x_p)|}{P_1^p \mathbb{1}(x_p)} \leq C \| P_{i+1}^p (P_1^i \mathbb{1} - \mathbb{1}) \|_{\mathrm{BV}},$$

où on a utilisé la condition **(Min)** et $\|.\|_{\sup} \leq \|.\|_{\mathrm{BV}}$. De plus, d'après **(Dec)**, on a

$$\| P_{i+1}^p (P_1^i \mathbb{1} - \mathbb{1}) \|_{\mathrm{BV}} \leq K \theta^{p-i-1} \| P_1^i \mathbb{1} - \mathbb{1} \|_{\mathrm{BV}} \leq C \theta^{p-i-1}$$

puisque $(P_1^i \mathbb{1})_i$ est bornée dans BV d'après l'hypothèse (H3).
Finalement, comme $|\int f_i \, dm| \leq \mathrm{Lip}_i(K)$, on obtient

$$\left| \sum_{i=0}^{p-1} \left(\frac{P_{i+1}^p \mathbb{1}(x_p)}{P_1^p \mathbb{1}(x_p)} - 1 \right) \int f_i \, dm \right| \leq C \sum_{i=0}^{p-1} \theta^{p-i-1} \mathrm{Lip}_i(K). \tag{3.3}$$

Pour le second terme, l'hypothèse **(Dec)** donne

$$\left\| P_{i+1}^p f_i - \left(\int f_i \, dm \right) P_{i+1}^p \mathbb{1} \right\|_{\mathrm{BV}} \leq K \theta^{p-i-1} \left\| f_i - \int f_i \, dm \right\|_{\mathrm{BV}} \leq C \theta^{p-i-1} \| f_i \|_{\mathrm{BV}}. \tag{3.4}$$

En utilisant la condition **(Min)**, (3.4) et $\|.\|_{\sup} \leq \|.\|_{\mathrm{BV}}$, on trouve

$$\left| \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{i=0}^{p-1} \left(P_{i+1}^p f_i(x_p) - \left(\int f_i \, dm \right) P_{i+1}^p \mathbb{1}(x_p) \right) \right| \leq C \sum_{i=0}^{p-1} \theta^{p-i-1} \|f_i\|_{\mathrm{BV}}. \tag{3.5}$$

Ainsi, il reste à estimer $\|f_i\|_{\mathrm{BV}}$. On peut observer que $\|f_i\|_{\mathrm{BV}} \leq \|f_i\|_{\sup} + \mathrm{Var}(f_i)$ et que pour tout y, on a

$$|f_i(y)| \leq P_1^i \mathbb{1}(y) \|H_i\|_{\sup} \leq P_1^i \mathbb{1}(y) \mathrm{Lip}_i(K) \leq C \mathrm{Lip}_i(K), \tag{3.6}$$

puisque $(P_1^i \mathbb{1})_i$ est bornée dans BV. Le point crucial réside donc dans l'estimation de la variation de f_i.
Afin de procéder, on va d'abord établir un contrôle sur la distorsion. Bien que cette propriété soit classique, on en reproduit la démonstration ici par complétude. Si T est une application de classe C^2 par morceaux sur l'intervalle, de partition de monotonicité $\mathcal{A}_T$, on définit sa distorsion $\mathrm{Dist}(T)$ comme étant la plus petite constante $C > 0$ telle que

$$|T'(x) - T'(y)| \leq C |T'(x)| |Tx - Ty|$$

pour tous $x, y \in I$ et $I \in \mathcal{A}_T$. On remarque que $(1+C)^{-1} \leq \frac{|T'(x)|}{|T'(y)|} \leq (1+C)$ pour tous $x, y \in I$.

Lemme 3.2.5. *Il existe $C > 0$ telle que $\mathrm{Dist}(T_1^n) \leq C$ pour tout n.*

Démonstration. Comme auparavant, on désigne par $\mathcal{A}_n$ la partition de monotonicité de T_n, et par $\mathcal{A}_n^m$ la partition de T_n^m. On a pour tous $x, y \in I \in \mathcal{A}_1^n$:

$$\left| \log \frac{|(T_1^n)'(x)|}{|(T_1^n)'(y)|} \right| \leq \sum_{j=1}^{n} \left| \log |T_j'(T_1^{j-1}x)| - \log |T_j'(T_1^{j-1}y)| \right| \leq C \sum_{j=1}^{n} |T_1^{j-1}x - T_1^{j-1}y|,$$

où on a utilisé le fait que la constante de Lipschitz de $\log |T_j'|$ est bornée indépendamment de j, puisque $|T_j''| \leq M$. Comme toutes les applications sont uniformément dilatantes, avec un facteur de dilatation au moins λ, et comme x et y appartiennent au même élément de la partition $\mathcal{A}_1^n$, on a

$$|T_1^{j-1}x - T_1^{j-1}y| \leq \lambda^{-(n-j-1)} |T_1^n x - T_1^n y|,$$

d'où le lemme, puisque $\sum_j \lambda^{-(n-j-1)}$ est borné indépendamment de n. □

On aura besoin du lemme technique suivant, qui est adapté du lemme II.4 de [44].

Lemme 3.2.6. *Il existe une constante $C > 0$ telle que pour tout $n \geq 1$,*

$$\sum_{I \in \mathcal{A}_1^n} \sup_I \frac{1}{|(T_1^n)'|} \leq C.$$

Démonstration. Pour tout $I \in \mathcal{A}_1^n$, il existe un plus petit entier $p = p_I \leq n-1$ tel que $T_1^p(I) \cap \partial\mathcal{A}_{p+1} \neq \emptyset$. On définit $\mathcal{A}_1^{n,p}$ comme l'ensemble de tous les intervalles $I \in \mathcal{A}_1^n$ pour lesquels $p_I = p$. Soit $I \in \mathcal{A}_1^{n,p}$. Il existe $a \in \partial I$ tel que $b = T_1^p a \in \partial\mathcal{A}_{p+1}$. D'après le lemme 3.2.5, on en déduit l'existence d'un $C > 0$ tel que pour tout $x \in I$,

$$|(T_1^n)'(x)| \geq C|(T_1^n)'(a)| = C|(T_{p+1}^n)'(T_1^p a)||(T_1^p)'(a)| \geq C\lambda^{n-p}|(T_1^p)'(a)|.$$

Ainsi, on a

$$\sup_I \frac{1}{|(T_1^n)'|} \leq C^{-1}\lambda^{-(n-p)}\frac{1}{|(T_1^p)'(a)|}.$$

Comme une pré-image par T_1^p d'un élément $b \in \mathcal{A}_{p+1}$ ne peut appartenir qu'à au plus deux $I \in \mathcal{A}_1^n$ différents, il s'ensuit que

$$\begin{aligned}
\sum_{I\in\mathcal{A}_1^n} \sup_I \frac{1}{|(T_1^n)'|} &\leq 2C^{-1}\sum_{p=0}^{n-1}\lambda^{-(n-p)}\sum_{b\in\partial\mathcal{A}_{p+1}}\sum_{T_1^p a=b}\frac{1}{|(T_1^p)'(a)|} \\
&= 2C^{-1}\sum_{p=0}^{n-1}\lambda^{-(n-p)}\sum_{b\in\partial\mathcal{A}_{p+1}} P_1^p\mathbb{1}(b) \leq 2C^{-1}\sum_{p=0}^{n-1}\lambda^{-(n-p)}\|P_1^p\mathbb{1}\|_{\mathrm{BV}}\sharp\partial\mathcal{A}_{p+1}.
\end{aligned}$$

Cette quantité est bornée indépendamment de n, puisque $(P_1^p)_p$ est bornée dans BV, et le nombre d'éléments de $\mathcal{A}_{p+1}$ est borné indépendamment de p par hypothèse sur la classe $\mathcal{C}$. □

Comme conséquence des deux précédents lemmes, on obtient :

Corollaire 3.2.7. *Il existe $C > 0$ tel que pour tout $n \geq 1$,*

$$\sum_{I\in\mathcal{A}_1^n} \mathrm{Var}_I\left(\frac{1}{|(T_1^n)'|}\right) \leq C,$$

où $\mathrm{Var}_I(f)$ désigne la variation totale de f sur le sous-intervalle $I \subset [0,1]$.

Démonstration. Soit $x_0 < \ldots < x_j$ une suite d'éléments de I. Alors

$$\begin{aligned}
\sum_{i=0}^{j-1}\left|\frac{1}{|(T_1^n)'(x_{i+1})|} - \frac{1}{|(T_1^n)'(x_i)|}\right| &= \sum_{i=0}^{j-1}\frac{\big|\,|(T_1^n)'(x_{i+1})| - |(T_1^n)'(x_i)|\,\big|}{|(T_1^n)'(x_i)||(T_1^n)'(x_{i+1})|} \\
&\leq \sum_{i=0}^{j-1}\frac{\mathrm{Dist}(T_1^n)|T_1^n x_{i+1} - T_1^n x_i|}{|(T_1^n)'(x_i)|} \leq C\sup_I\frac{1}{|(T_1^n)'|},
\end{aligned}$$

où pour la dernière inégalité, on a utilisé le lemme 3.2.5 et le fait que $(T_1^n x_i)_i$ est une suite monotone dans $[0,1]$. Une application directe du lemme 3.2.6 prouve le corollaire □

On est à présent en mesure d'estimer $\mathrm{Var}(f_i)$. Pour $I \in \mathcal{A}_1^i$, on désigne par $S_{i,I}$ la branche inverse de la restriction de T_1^i à I. On peut alors écrire

$$f_i = P_1^i H_i = \sum_{I \in \mathcal{A}_1^i} \left(\frac{H_i}{|(T_1^i)'|} \right) \circ S_{i,I} \, \mathbb{1}_{T_1^i(I)}.$$

En utilisant des propriétés standards de la variation totale, il vient

$$\mathrm{Var}(f_i) \leq \sum_{I \in \mathcal{A}_1^i} \mathrm{Var}_I \left(\frac{H_i}{|(T_1^i)'|} \right) + 2 \sum_{a \in \partial \mathcal{A}_1^i} \frac{|H_i(a)|}{|(T_1^i)'(a)|} \leq (\mathrm{I})_i + (\mathrm{II})_i + (\mathrm{III})_i,$$

avec

$$\begin{aligned}
(\mathrm{I})_i &= \sum_{I \in \mathcal{A}_1^i} \mathrm{Var}_I \left(\frac{1}{|(T_1^i)'|} \right) \sup_I |H_i|, \\
(\mathrm{II})_i &= \sum_{I \in \mathcal{A}_1^i} \sup_I \frac{1}{|(T_1^i)'|} \mathrm{Var}_I(H_i), \\
(\mathrm{III})_i &= 2 \sum_{a \in \partial \mathcal{A}_1^i} \frac{|H_i(a)|}{|(T_1^i)'(a)|}.
\end{aligned}$$

En utilisant la condition de Lipschitz sur K, on obtient

$$(\mathrm{I})_i \leq C \mathrm{Lip}_i(K) \sum_{I \in \mathcal{A}_1^i} \mathrm{Var}_I \left(\frac{1}{|(T_1^i)'|} \right),$$

ce qui donne $(\mathrm{I})_i \leq C \mathrm{Lip}_i(K)$ par le corollaire 3.2.7.
On estime à présent $(\mathrm{II})_\mathrm{i}$. Soit $y_0 < \ldots < y_l$ une suite de points dans I. Afin d'estimer $\sum_{j=0}^{l-1} |H_i(y_{j+1}) - H_i(y_j)|$, on sépare H_i en deux termes de façon évidente, et on traite le premier, le second étant complètement similaire. On a

$$\begin{aligned}
&\sum_{j=0}^{l-1} \sum_{k=0}^{i} \left| K(y_{j+1}, \ldots, T_1^k y_{j+1}, T_1^{k+1} y_j, \ldots, T_1^i y_j, \ldots) - K(y_{j+1}, \ldots, T_1^{k-1} y_{j+1}, T_1^k y_j, \ldots, T_1^i y_j, \ldots) \right| \\
&\qquad \leq \sum_{j=0}^{l-1} \sum_{k=0}^{i} \mathrm{Lip}_k(K) |T_1^k y_{j+1} - T_1^k y_j| \leq \sum_{k=0}^{i} \mathrm{Lip}_k(K) m(T_1^k(I)).
\end{aligned}$$

Comme I appartient à $\mathcal{A}_1^i$, l'intervalle $T_1^k(I)$ est inclus dans un intervalle de monotonicité pour T_{k+1}^i, et sa longueur est donc inférieure à $\lambda^{-(i-k)}$. Par conséquent, on a

$$\mathrm{Var}_I(H_i) \leq 2 \sum_{k=0}^{i} \lambda^{-(i-k)} \mathrm{Lip}_k(K)$$

et une application du lemme 3.2.6 donne $(\mathrm{II})_i \leq C \sum_{k=0}^{i} \lambda^{-(i-k)} \mathrm{Lip}_k(K)$.
Une nouvelle application du lemme 3.2.6 entraine que le troisième terme est majoré par $(\mathrm{III})_i \leq C \mathrm{Lip}_i(K)$.
En rassemblant toutes les estimations, on trouve donc que

$$\mathrm{V}(f_i) \leq C \sum_{k=0}^{i} \lambda^{-(i-k)} \mathrm{Lip}_k(K) \tag{3.7}$$

Finalement, (3.6) et (3.7) donnent la même estimation pour $\|f_i\|_{\mathrm{BV}}$.
En revenant à $K_p(\underline{x})$, on a par (3.3) et (3.5) :

$$\left| K_p(\underline{x}) - \int K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots)\, dm(y) \right| \leq C \sum_{i=0}^{p-1} \theta^{p-i-1} \sum_{k=0}^{i} \lambda^{-(i-k)} \mathrm{Lip}_k(K),$$

qui est inférieur à $C \sum_{k=0}^{p-1} \rho^{p-k} \mathrm{Lip}_k(K)$, comme le montre un simple calcul. Ceci prouve le lemme 3.2.4 et termine donc la démonstration du théorème 3.2.2. □

Démonstration de la relation (3.2). Il s'agit de démontrer que

$$K_p(\underline{x}) = \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{T_1^p y = x_p} \frac{K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots)}{|(T_1^p)'(y)|}, \tag{3.8}$$

pour $\tilde{m}$-p.t. $\underline{x} \in X^{\mathbb{N}}$.
Premièrement, on rappelle quelques propriétés classiques des opérateurs de transfert. Soient $(X, \mathcal{A}, \mu)$ et $(Y, \mathcal{B}, \nu)$ deux espaces de probabilité, et $T : X \to Y$ une application mesurable non-singulière. Son opérateur de transfert associé $P_T : L^1(\mu) \to L^1(\nu)$ est défini de la façon suivante : pour $f \in L^1(\mu)$, $P_T(f)$ est la dérivée de Radon-Nykodim $\frac{d(\lambda \circ T^{-1})}{d\nu}$, où $d\lambda = f d\mu$.
Il satisfait les propriétés élémentaires suivantes :

Lemme 3.2.8. *Pour tout $f \in L^1(\mu)$ et tout $g \in L^\infty(\nu)$, on a*

1. $\int_X f\, g \circ T d\mu = \int_Y P_T(f)\, g\, d\nu$,
2. $P_T(g \circ T) = g P_T \mathbb{1}$,
3. $\mathbb{E}_\mu(f | T^{-1}\mathcal{B}) = \left(\frac{P_T(f)}{P_T \mathbb{1}}\right) \circ T$, *où* $\frac{P_T(f)}{P_T \mathbb{1}}$ *est défini comme valant* 0 *sur* $\{P_T \mathbb{1} = 0\}$. *En particulier, on a* $\left(\frac{P_T(f_1)}{P_T(f_2)}\right) \circ T = \frac{\mathbb{E}_\mu(f_1 | T^{-1}\mathcal{B})}{\mathbb{E}_\mu(f_2 | T^{-1}\mathcal{B}}$) *pour tout* $f_1, f_2 \in L^1(\mu)$.

On rappelle que $\Phi : X \to \mathbb{N}$ est défini par $\Phi(x) = (T_1^n x)_{n \geq 0}$, et que $\tilde{\mathcal{F}}_p$ est la tribu sur $X^{\mathbb{N}}$ engendrée par les coordonnées $(x_k)_{k \geq p}$. Clairement, $\tilde{\mathcal{F}}_p = \sigma^{-p} \tilde{\mathcal{F}}$, où $\sigma : X^{\mathbb{N}} \to X^{\mathbb{N}}$ est le décalage à gauche. On a $K_p = \mathbb{E}_{\tilde{m}}(K | \tilde{F}_p)$, où $K : X^{\mathbb{N}} \to \mathbb{R}$ est séparément Lipschitz, et ne dépend que des n premières variables.

Lemme 3.2.9. *On a* $K_p \circ \Phi = \left(\frac{P_1^p (K \circ \Phi)}{P_1^p \mathbb{1}}\right) \circ T_1^p$, *m-p.p.*

Démonstration. Par le point (3) du lemme 3.2.8, on a

$$K_p = \mathbb{E}_{\tilde{m}}(K|\sigma^{-p}\tilde{\mathcal{F}}) = \left(\frac{P_\sigma^p K}{P_\sigma^p \mathbb{1}}\right) \circ \sigma^p.$$

On introduit l'application $\Phi_p : X \to X^{\mathbb{N}}$ définie par $\Phi_p(x) = (x, T_{p+1}x, T_{p+1}^{p+2}x, \ldots) = (T_{p+1}^{p+n}x)_{n\geq 0}$. Cette application satisfait $\sigma^p \circ \Phi = \Phi_p \circ T_1^p$. Elle n'est pas non-singulière, et par conséquent son opérateur de transfert n'est pas défini, mais la relation précédente entraine que $P_{\Phi_p}(P_1^p f)$ est bien défini pour tout $f \in L^1(m)$ comme étant égal à $P_\sigma^p P_\Phi f$, et satisfait les propriétés décrites au lemme 3.2.8.
On a alors

$$K_p \circ \Phi = \left(\frac{P_\sigma^p K}{P_\sigma^p \mathbb{1}}\right) \circ \Phi_p \circ T_1^p = \left(\frac{P_\sigma^p P_\Phi(K \circ \Phi)}{P_\sigma^p P_\Phi \mathbb{1}}\right) \circ \Phi_p \circ T_1^p,$$

par le point (2) du lemme 3.2.8, puisque $P_\Phi \mathbb{1} = \mathbb{1}$.
Mais comme $P_\sigma^p P_\Phi = P_{\Phi_p} P_1^p$, on peut utiliser le point (3) du lemme 3.2.8 pour obtenir

$$K_p \circ \Phi = \left(\frac{P_{\Phi_p} P_1^p(K \circ \Phi)}{P_{\Phi_p} P_1^p \mathbb{1}}\right) \circ \Phi_p \circ T_1^p = \left(\frac{((P_{\Phi_p}\mathbb{1}) \circ \Phi_p)\mathbb{E}_m(P_1^p(K \circ \Phi)|(\Phi_p)^{-1}\tilde{\mathcal{F}})}{((P_{\Phi_p}\mathbb{1}) \circ \Phi_p)\mathbb{E}_m(P_1^p \mathbb{1}|(\Phi_p)^{-1}\tilde{\mathcal{F}})}\right) \circ T_1^p.$$

Le lemme est prouvé dès que l'on remarque que $(\Phi_p)^{-1}\tilde{\mathcal{F}} = \mathcal{F}$. □

Le lemme entraine la relation (3.8). En effet, si on définit $\pi_p : X^{\mathbb{N}} \to X$ comme l'application qui envoie $\underline{x}$ sur sa p-ième coordonnée x_p, alors $T_1^p = \pi_p \circ \Phi$, et le lemme 3.2.9 nous dit que $K_p \circ \Phi = \left(\frac{P_1^p(K\circ\Phi)}{P_1^p \mathbb{1}}\right) \circ \pi_p \circ \Phi$, m-p.p.
Cela implique que $K_p = \left(\frac{P_1^p(K\circ\Phi)}{P_1^p \mathbb{1}}\right) \circ \pi_p$, $\tilde{m}$-p.p., et prouve la relation désirée, puisque, $\tilde{m}$-p.p.

$$\left(\frac{P_1^p(K \circ \Phi)}{P_1^p \mathbb{1}}\right) \circ \pi_p(\underline{x}) = \frac{1}{P_1^p \mathbb{1}(x_p)} \sum_{T_1^p y = x_p} \frac{K(y, T_1 y, \ldots, T_1^{p-1} y, x_p, \ldots)}{|(T_1^p)'(y)|}.$$

Pour le voir, on pose $A = \{K_p \circ \Phi = \left(\frac{P_1^p(K\circ\Phi)}{P_1^p \mathbb{1}}\right) \circ \pi_p \circ \Phi\}$. On a $A \in \mathcal{F}$, et $m(A) = 1$ d'après le lemme 3.2.9. Comme $(X, \mathcal{F})$ et $(X^{\mathbb{N}}, \tilde{\mathcal{F}})$ sont tous deux des espaces polonais munis de leurs tribus boréliennes respectives, et que l'application $\Phi : X \to X^{\mathbb{N}}$ est mesurable et injective, on obtient que $\Phi(A) \in \tilde{\mathcal{F}}$ par le théorème 8.3.7 de [42].
Ainsi, $\tilde{m}(\Phi(A)) = m(\Phi^{-1}(\Phi(A))) \geq m(A) = 1$, ce que montre que $K_p = \left(\frac{P_1^p(K\circ\Phi)}{P_1^p \mathbb{1}}\right) \circ \pi_p$, $\tilde{m}$-p.p., puisque l'ensemble sur lequel cette relation est vraie contient $\Phi(A)$. □

3.2.2 Applications

Soit $(T_n)_{n\geq 1}$ un système dynamique séquentiel sur $[0, 1]$ qui satisfait les hypothèses du théorème 3.2.2. On donne des applications de l'inégalité de concentration, en suivant les idées de [44] et [37].

Grandes déviations pour les sommes ergodiques.

Si on considère des observables K de la forme

$$K(x_0, \dots, x_{n-1}) = \sum_{k=0}^{n-1} f_k(x_k)$$

avec des f_k Lipschitz, une application directe du théorème 3.2.2 donne

Proposition 3.2.10. *Il existe $C > 0$ telle que pour toute suite $f_n : [0,1] \to \mathbb{R}$ d'observables Lipschitz, tout $n \geq 1$ et tout $t > 0$:*

$$m\left(\frac{1}{n}\sum_{k=0}^{n-1}\left[f_k \circ T_1^k - \int f_k \circ T_1^k \, dm\right] > t\right) \leq e^{-\frac{Cn^2t^2}{\sum_{k=0}^{n-1}\mathrm{Lip}(f_k)^2}}.$$

Cet énoncé raffine la proposition 2.1.10 car il précise la dépendance du facteur τ vis-à-vis de la suite (f_n), sous l'hypothèse que celle-ci est formée de fonctions Lipschitz.

Mesure empirique.

Pour $x \in [0,1]$, on définit pour $n \geq 1$ la mesure empirique

$$\mathcal{E}_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{T_1^k x}.$$

Par la loi des grands nombres 2.1.7, on sait que pour p.t. x, cette mesure se rapproche, dans la topologie faible, lorsque $n \to \infty$, de la mesure m_n définie par

$$m_n = \frac{1}{n}\sum_{k=0}^{n-1} (T_1^k)_\star m.$$

Il est naturel de chercher à quantifier la vitesse de convergence. Pour cela, on doit d'abord introduire une notion de distance sur l'espace des mesures de probabilité. On va considérer la distance de Kantorovich, qui, pour deux mesures de probabilité μ_1 et μ_2 sur $[0,1]$, peut être définie par

$$\kappa(\mu_1, \mu_2) = \int_0^1 |F_{\mu_1}(t) - F_{\mu_2}(t)| \, dt, \tag{3.9}$$

où $F_{\mu_i}(t) = \mu_i([0,t])$ est la fonction de répartition de μ_i. On a alors

Proposition 3.2.11. *Il existe $t_0 > 0$ et $C > 0$ tels que pour tout $t > t_0$ et tout $n \geq 1$:*

$$m\left(\kappa(\mathcal{E}_n, m_n) > \frac{t}{\sqrt{n}}\right) \leq e^{-Ct^2}.$$

Démonstration. Pour $n \geq 1$, on définit

$$K_n(x_0, \ldots, x_{n-1}) = \int_0^1 \left| \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{[0,t]}(x_k) - F_{m_n}(t) \right| dt.$$

On a clairement $\kappa(\mathcal{E}_n(x), m_n) = K_n(x, \ldots, T_1^{n-1}x)$ et $\mathrm{Lip}_j(K_n) \leq \frac{1}{n}$ pour tout $0 \leq j \leq n-1$. Par l'inégalité de concentration 3.2.2 et la proposition 3.1.3, on en déduit

$$m\left(\kappa(\mathcal{E}_n, m_n) - \mathbb{E}_m(\kappa(\mathcal{E}_n, m_n)) > \frac{t}{\sqrt{n}}\right) \leq e^{-Ct^2}.$$

Pour conclure, il est suffisant de prouver que $\mathbb{E}_m(\kappa(\mathcal{E}_n, m_n))$ est d'ordre $\frac{1}{\sqrt{n}}$.
Soit χ_t la fonction caractéristique $\mathbb{1}_{[0,t]}$. En utilisant l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned}
\mathbb{E}_m(\kappa(\mathcal{E}_n, m_n)) &= \int \left(\int_0^1 \left| \frac{1}{n} \sum_{k=0}^{n-1} \chi_t(T_1^k x) - F_{m_n}(t) \right| dt \right) dm(x) \\
&\leq \left[\int_0^1 \left(\int \left| \frac{1}{n} \sum_{k=0}^{n-1} \left(\chi_t \circ T_1^k - F_{(T_1^k)_\star m}(t) \right) \right|^2 dm \right) dt \right]^{\frac{1}{2}}.
\end{aligned}$$

En développant le carré, on trouve

$$\begin{aligned}
&\int \left| \frac{1}{n} \sum_{k=0}^{n-1} \left(\chi_t \circ T_1^k - F_{(T_1^k)_\star m}(t) \right) \right|^2 dm \\
&= \frac{1}{n^2} \sum_{k,l=0}^{n-1} \int \left(\chi_t \circ T_1^k - F_{(T_1^k)_\star m}(t) \right) \left(\chi_t \circ T_1^l - F_{(T_1^l)_\star m}(t) \right) dm.
\end{aligned}$$

Si $k \leq l$, par les propriétés des opérateurs de transfert, et car $F_{(T_1^k)_\star m}(t) = \int \chi_t \circ T_1^k \, dm$,

$$\begin{aligned}
&\int \left(\chi_t \circ T_1^k - F_{(T_1^k)_\star m}(t) \right) \left(\chi_t \circ T_1^l - F_{(T_1^l)_\star m}(t) \right) dm \\
&= \int \left(\chi_t - \int \chi_t \circ T_1^l \right) P_{k+1}^l \left(\left(\chi_t - \int \chi_t \circ T_1^k \right) P_1^k \mathbb{1} \right) dm.
\end{aligned}$$

Comme $\left(\chi_t - \int \chi_t \circ T_1^k \right) P_1^k \mathbb{1}$ est d'intégrale nulle, $\{\chi_t\}_t$ est une famille bornée dans BV, et $(P_1^k \mathbb{1})_k$ est aussi bornée dans BV, on peut utiliser la propriété **(Dec)** pour obtenir

$$\left| \int \left(\chi_t - \int \chi_t \circ T_1^l \right) P_{k+1}^l \left(\left(\chi_t - \int \chi_t \circ T_1^k \right) P_1^k \mathbb{1} \right) dm \right| \leq C\theta^{l-k}.$$

Il s'ensuit que

$$\mathbb{E}_m(\kappa(\mathcal{E}_n, m_n)) \leq \frac{C}{n} \left(\sum_{k,l=0}^{n-1} \theta^{-|k-l|} \right)^{\frac{1}{2}} = \mathcal{O}\left(\frac{1}{\sqrt{n}} \right).$$

□

Approximation des orbites.

Soit A un sous-ensemble mesurable de $[0,1]$, avec $m(A) > 0$. Comment peut-on approximer la trajectoire d'un point $x \in [0,1]$ par une trajectoire démarrant dans A ? Le résultat suivant fournit une estimation sur l'efficacité en moyenne d'une telle approximation :

Proposition 3.2.12. *Il existe deux constantes $C_1, C_2 > 0$ telles que pour tout A avec $m(A) > 0$ et tout $n \geq 1$, la suite de fonctions $(\mathcal{Z}_n)$ définies par*

$$\mathcal{Z}_n(x) = \frac{1}{n} \inf_{y \in A} \sum_{k=0}^{n-1} |T_1^k x - T_1^k y|,$$

satisfait pour tout $t > 0$:

$$m\left(\mathcal{Z}_n > C_1 \frac{\sqrt{|\log m(A)|}}{\sqrt{n}} + \frac{t}{\sqrt{n}}\right) \leq e^{-C_2 t^2}.$$

Démonstration. On applique l'inégalité de concentration aux fonctions

$$K_n(x_0, \ldots, x_{n-1}) = \frac{1}{n} \inf_{y \in A} \sum_{k=0}^{n-1} |T_1^k y - x_k|.$$

Elles vérifient $\mathrm{Lip}_k(K_n) \leq \frac{1}{n}$, et donc

$$m\left(\mathcal{Z}_n - \mathbb{E}_m(\mathcal{Z}_n) > \frac{t}{\sqrt{n}}\right) \leq e^{-Ct^2}.$$

On estime à présent $\mathbb{E}_m(\mathcal{Z}_n)$. On a

$$m(A) \leq m\left(\mathbb{E}_m(\mathcal{Z}_n) - \mathcal{Z}_n > \frac{1}{\sqrt{n}} \frac{\sqrt{n}|\mathbb{E}_m(\mathcal{Z}_n)|}{2}\right),$$

puisque $\mathcal{Z}_n = 0$ sur A. Par l'inégalité de concentration, on en déduit que $m(A) \leq e^{-Cn(\mathbb{E}_m(\mathcal{Z}_n))^2}$, d'où

$$\mathbb{E}_m(\mathcal{Z}_n) \leq C \frac{\sqrt{|\log m(A)|}}{\sqrt{n}}.$$

□

On peut aussi démontrer un résultat similaire pour le nombre d'écarts entre une trajectoire quelconque et une trajectoire issue de A, pour une erreur $\epsilon > 0$ donnée.

Proposition 3.2.13. *Il existe deux constantes $C_1, C_2 > 0$ telles que pour tout A avec $m(A) > 0$, tout $n \geq 1$ et tout $0 < \epsilon < \frac{1}{2}$, la suite de fonctions $(\mathcal{Z}'_{n,\epsilon})$ de fonctions définies par*

$$\mathcal{Z}'_{n,\epsilon}(x) = \frac{1}{n} \inf_{y \in A} \sharp\{0 \leq k < n \,:\, |T_1^k x - T_1^k y| > \epsilon\},$$

satisfait pour tout $t > 0$:

$$m\left(\mathcal{Z}'_{n,\epsilon} > C_1 \frac{\sqrt{|\log m(A)|}}{\epsilon\sqrt{n}} + \frac{t}{\epsilon\sqrt{n}}\right) \leq e^{-C_2 t^2}.$$

Démonstration. On ne peut pas appliquer directement l'inégalité de concentrations aux fonctions

$$K'_{n,\epsilon}(x_0, \ldots, x_{n-1}) = \frac{1}{n} \inf_{y \in A} \sum_{k=0}^{n-1} \mathbb{1}_{[0,1] \setminus B_\epsilon(T_1^k y)}(x_j),$$

car celles-ci ne sont pas séparément Lipschitz. On approxime donc l'indicatrice du complémentaire de l'intervalle $[-\epsilon, +\epsilon]$ par une fonction linéaire par morceaux : on pose

$$g_\epsilon(s) = \begin{cases} \frac{|s|}{\epsilon} & \text{si } |s| < \epsilon \\ 1 & \text{sinon} \end{cases}.$$

On définit alors

$$\tilde{K}_{n,\epsilon}(x_0, \ldots, x_{n-1}) = \frac{1}{n} \inf_{y \in A} \sum_{k=0}^{n-1} g_\epsilon(x_k - T_1^k y).$$

Ces fonctions sont séparément Lipschitz, de constantes $\mathrm{Lip}_k(\tilde{K}_{n,\epsilon}) \leq \frac{1}{\epsilon n}$, et vérifient $K'_{n,\epsilon} \leq \tilde{K}_{n,\epsilon}$. On pose aussi $\tilde{\mathcal{Z}}_{n,\epsilon}(x) = \tilde{K}_{n,\epsilon}(x, \ldots, T_1^{n-1} x)$. Comme $\mathcal{Z}'_{n,\epsilon} \leq \tilde{\mathcal{Z}}_{n,\epsilon}$, on obtient grâce à l'inégalité de concentration :

$$m\left(\mathcal{Z}'_{n,\epsilon} > m(\tilde{\mathcal{Z}}_{n,\epsilon}) + \frac{t}{\epsilon\sqrt{n}}\right) \leq m\left(\tilde{\mathcal{Z}}_{n,\epsilon} > m(\tilde{\mathcal{Z}}_{n,\epsilon}) + \frac{t}{\epsilon\sqrt{n}}\right) \leq e^{-Ct^2}.$$

On peut alors estimer $m(\tilde{\mathcal{Z}}_{n,\epsilon})$ par la même méthode qu'à la proposition 3.2.12. □

Version presque sure du théorème de la limite centrale.

Soit (Y_n) un processus aléatoire à valeurs réelles qui satisfait le théorème de la limite centrale : il existe $A_n \in \mathbb{R}$ et $B_n > 0$ telles que $S_n = \sum_{k=0}^{n-1} Y_k$ vérifie

$$\frac{S_n - A_n}{B_n} \to W$$

en loi, où W est distribuée selon la loi normale centrée réduite $\mathcal{N}(0, 1)$.

Par définition, cela signifie que pour tout $t \in \mathbb{R}$,

$$\mathbb{E}\left(\mathbb{1}_{\left\{\frac{S_n - A_n}{B_n} \leq t\right\}}\right) \to \mathbb{P}(W \leq t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{s^2}{2}}\, ds.$$

On peut alors se demander si l'on peut dire quelque chose à propos de la convergence presque sure de $\mathbb{1}_{\left\{\frac{S_n - A_n}{B_n} \leq t\right\}}$. Si cette quantité ne peut pas en général converger presque surement, il s'avère qu'après un procédé de moyenne approprié, le plus souvent logarithmique, on retrouve la convergence presque sure.
Plus précisément, on dira que (Y_n) satisfait la version presque sure du théorème de la limite centrale si, presque surement,

$$\frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k}\, \delta_{\frac{S_k - A_k}{B_k}} \to \mathcal{N}(0,1)$$

dans la topologie faible, avec $H_n = \sum_{k=1}^{n} \frac{1}{k} \sim \log n$.
On renvoie à [22, 85] pour des survols particulièrement intéressants de ce sujet, et à [38] pour une description des applications aux systèmes dynamiques.
Soit à présent un système dynamique séquentiel $(T_n)_{n \geq 1}$ qui satisfait les hypothèses du théorème 3.2.2. Soit $f_n : [0,1] \to \mathbb{R}$ une suite d'observables Lipschitz, bornée dans Lip. On définit

$$S_n = \sum_{k=0}^{n-1} \left[f_k \circ T_1^k - \int f_k \circ T_1^k\, dm \right].$$

Si $\|S_n\|_{L^2_m}$ est non-bornée, on sait d'après le théorème 2.1.8 que $\frac{S_n}{\|S_n\|_{L^2_m}}$ converge en loi vers $W \sim \mathcal{N}(0,1)$.
Afin de prouver une version presque sure de cette convergence, on aura besoin d'une hypothèse supplémentaire sur la croissance de la variance de la somme : plus précisément, on supposera que $\|S_n\|_{L^2_m} \geq C n^{\frac{1}{2}}$ pour tout n, pour un certain $C > 0$.

Proposition 3.2.14. *Soit (f_n) une suite d'observables bornée dans* Lip *et soit S_n définie comme ci-dessus. On suppose qu'il existe $C > 0$ telle que $\|S_n\|_{L^2_m} \geq C n^{\frac{1}{2}}$ pour tout $n \geq 1$. Alors S_n satisfait la version presque sure du théorème de la limite centrale : pour presque tout $x \in [0,1]$*

$$\frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k}\, \delta_{\frac{S_k(x)}{\|S_k\|_2}} \to \mathcal{N}(0,1)$$

dans la topologie faible.

Démonstration. On doit montrer que presque partout,

$$\frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k}\, g\left(\frac{S_k}{\|S_k\|_2}\right) \to \mathbb{E}(g(W))$$

pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ continue et bornée. Par des arguments standards de densité et de séparabilité (voir par exemple le théorème 2.4 de [85]), il est suffisant de montrer que pour toute fonction g bornée et Lipschitz, cette convergence a lieu presque partout. On remarque aussi que l'on peut supposer sans perdre de généralité que $g(0) = 0$. Par conséquent, il existe $C > 0$ telle que $|g(x)| \leq C|x|$ pour tout x.
On définit

$$K_n(x_0, \ldots, x_{n-1}) = \frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k} g\left(\frac{\sum_{j=0}^{k-1} \left[f_k(x_j) - \int f_k \circ T_1^j \right]}{\|S_k\|_2} \right).$$

On a alors $K_n(x, \ldots, T_1^{n-1}x) = \frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2}\right)$, et l'hypothèse sur $\|S_n\|_{L^2_m}$ donne

$$\mathrm{Lip}_j(K_n) \leq \frac{1}{H_n} \sum_{k>j} \frac{1}{k} \mathrm{Lip}(g) \frac{\mathrm{Lip}(f_k)}{\|S_k\|_2} = \mathcal{O}\left(\frac{1}{H_n} \sum_{k>j} k^{-\frac{3}{2}} \right) = \mathcal{O}\left(\frac{1}{j^{\frac{1}{2}} H_n} \right).$$

Ainsi,

$$\sum_{j=0}^{n-1} \mathrm{Lip}_j^2(K_n) = \mathcal{O}\left(\frac{1}{H_n^2} \sum_{j=0}^{n-1} j^{-1} \right) = \mathcal{O}\left(\frac{1}{\log n} \right).$$

En utilisant l'inégalité de concentration, on en déduit que

$$m\left(\left| \frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) - \mathbb{E}_m\left(\frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) \right) \right| \geq t \right) \leq e^{-Ct^2 \log n}.$$

Par conséquent, si l'on définit la sous-suite $n_j = e^{j^\alpha}$ avec $\alpha > 0$, on a que

$$\frac{1}{H_{n_j}} \sum_{k=1}^{n_j} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right)$$

converge presque surement vers $\mathbb{E}(g(W))$ par le lemme de Borel-Cantelli, puisque

$$\mathbb{E}_m\left(\frac{1}{H_{n_j}} \sum_{k=1}^{n_j} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) \right)$$

converge vers $\mathbb{E}(g(W))$ lorsque $j \to \infty$, par le TCL.
Il reste à estimer ce qu'il se passe entre deux termes consécutifs de la sous-suite.
Pour $n_j < n \leq n_{j+1}$, on a

$$\begin{aligned} &\frac{1}{H_n} \sum_{k=1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) - \frac{1}{H_{n_j}} \sum_{k=1}^{n_j} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) \\ &= \frac{1}{H_n} \sum_{k=n_j+1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right) + \left(\frac{H_{n_j}}{H_n} - 1 \right) \frac{1}{H_{n_j}} \sum_{k=1}^{n_j} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2} \right). \end{aligned}$$

D'après les précédentes observations, la borne supérieure sur tous les n vérifiant $n_j < n \leq n_{j+1}$ du second terme tend vers 0 lorsque $j \to \infty$. Pour le premier terme, l'hypothèse faite sur g donne

$$\sup_{n_j<n\leq n_{j+1}} \left| \frac{1}{H_n} \sum_{k=n_j+1}^{n} \frac{1}{k} g\left(\frac{S_k}{\|S_k\|_2}\right) \right| \leq \frac{C}{H_{n_j}} \sum_{k=n_j+1}^{n_{j+1}} \frac{|S_k|}{k\|S_k\|_2}.$$

On prouve à présente la convergence presque-sure vers 0 de $G_j = \frac{1}{H_{n_j}} \sum_{k=n_j+1}^{n_{j+1}} \frac{|S_k|}{k\|S_k\|_2}$. En utilisant l'inégalité de Cauchy-Schwarz, on a

$$\begin{aligned} \mathbb{E}_m(G_j^2) \leq \frac{1}{H_{n_j}^2} \sum_{k,l=n_j+1}^{n_{j+1}} \frac{\mathbb{E}_m(S_k^2)^{\frac{1}{2}} \mathbb{E}_m(S_l^2)^{\frac{1}{2}}}{k\|S_k\|_2 l\|S_l\|_2} &= \frac{1}{H_{n_j}^2} \left(\sum_{k=n_j+1}^{n_{j+1}} k^{-1} \right)^2 \\ &= \frac{(\log n_{j+1} - \log n_j + \mathcal{O}(1))^2}{H_{n_j}^2} \leq \frac{\mathcal{O}(1)}{j^2}. \end{aligned}$$

Le résultat découle donc du lemme de Borel-Cantelli, puisque $\mathbb{E}_m(G_j^2)$ est sommable. □

3.3 Inégalités de concentration pour les systèmes dynamiques aléatoires

3.3.1 Le résultat et sa démonstration

On reprend les notations du chapitre 1, et plus précisément de la section 1.1.2. On considère un ensemble fini Ω avec un vecteur de probabilité $\mathbb{P} = \{p_\omega\}_{\omega\in\Omega}$, et une famille d'applications $T_\omega : [0,1] \to [0,1]$ de type Lasota-Yorke, toutes uniformément dilatantes : $\lambda(T_\omega) = \inf |T'_\omega| > 1$. On fait aussi l'hypothèse que le système aléatoire possède la propriété de recouvrement. Sous ces hypothèses, il existe une unique mesure de probabilité μ stationnaire et absolument continue, et sa densité h appartient à BV et est bornée uniformément loin de 0.

Théorème 3.3.1. *Le processus $Z_n(\underline{\omega}, x) = T_{\underline{\omega}}^n x$, défini sur $(\tilde{\Omega} \times X, \tilde{\mathbb{P}} \otimes \mu)$, satisfait une inégalité de concentration exponentielle.*

On remarquera qu'il s'agit d'une inégalité de concentration pour le système produit gauche $(\tilde{\Omega} \times X, S, \tilde{\mathbb{P}} \otimes \mu)$, mais uniquement pour des observables K qui ne dépendent que de la composante x. Il s'agit aussi d'un résultat de type annealed qui ne dit rien à propos du processus $x \mapsto T_{\underline{\omega}}^n x$, pour $\underline{\omega}$ fixé, générique.
Afin de démontrer ce résultat, on va, comme dans le cas non-autonome, adapter la méthode des différences bornées de McDiarmid. On notera par P l'opérateur de transfert en moyenne du système aléatoire, défini par rapport à la mesure de Lebesgue m, et par L l'opérateur de

transfert en moyenne défini par rapport à la mesure stationnaire μ. On rappelle que L agit sur $L^1(\mu)$ par $L(f) = \frac{P(fh)}{h}$, d'où

$$Lf(x) = \sum_{\omega\in\Omega} p_\omega \sum_{T_\omega y = x} \frac{h(y)f(y)}{h(x)|T'_\omega(y)|}.$$

Comme h et $\frac{1}{h}$ appartiennent à BV qui est une algèbre de Banach, L agit sur BV et y admet un trou spectral.
On renvoie à la section 1.3 pour la construction du système symbolique $(\Omega^\star, \tau, \mu_c)$ et de la filtration décroissante $\{\mathcal{F}_p\}_{p\geq 0}$. On prolonge K en une fonction sur $X^{\mathbb{N}}$ ne dépendant que des n premières coordonnées. On a alors $\mathbb{E}_{\mu_c}(K) = \mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}\left(K(x, T^1_{\underline{\omega}}x, \ldots, T^{n-1}_{\underline{\omega}}x)\right)$ et

$$\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}\left(e^{K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x)-\mathbb{E}_{\tilde{\mathbb{P}}\otimes\mu}\left(K(x,T^1_{\underline{\omega}}x,\ldots,T^{n-1}_{\underline{\omega}}x)\right)}\right) = \mathbb{E}_{\mu_c}(e^{K-\mathbb{E}_{\mu_c}(K)}),$$

d'après le lemme 1.3.1.
On définit $K_p = \mathbb{E}_{\mu_c}(K|\mathcal{F}_p)$ et $D_p = K_p - K_{p+1}$. Comme $(\Omega^\star, \tau, \mu_c)$ est un facteur du système $(\tilde{\Omega} \times X, F, \tilde{\mathbb{P}} \otimes \mu)$ qui exact d'après la proposition 1.1.4, on en déduit que la tribu $\mathcal{F}_\infty := \bigcap_{p\geq 0} \mathcal{F}_p = \bigcap_{p\geq 0} \tau^{-p}\mathcal{F}$ est triviale modulo μ_c. Par conséquent, d'après le théorème de convergence de Doob, K_p converge μ_c-p.s. et dans L^p, $1 \leq p < \infty$ vers $\mathbb{E}_{\mu_c}(K)$, d'où

$$K - \mathbb{E}_{\mu_c}(K) = \sum_{p\geq 0} D_p.$$

D'après l'inégalité de Azuma-Hoeffding 3.1.8, on en déduit l'existence de $C > 0$ telle que pour tout $P \geq 0$,

$$\mathbb{E}_{\mu_c}(e^{\sum_{p=0}^{P} D_p}) \leq e^{C\sum_{p=0}^{P} \sup|D_p|^2}.$$

Il reste donc à borner D_p :

Proposition 3.3.2. *Il existe $0 < \rho < 1$ et $C > 0$ telles que pour tout p,*

$$|D_p| \leq C \sum_{j=0}^{p} \rho^{p-j} \mathrm{Lip}_i(K).$$

Comme pour le cas non-autonome, cette proposition entraine le théorème principal, et découle du lemme suivant :

Lemme 3.3.3. *Il existe $0 < \rho < 1$ et $C > 0$ telles que pour tout p et μ_c-p.t. $\underline{x} \in X^{\mathbb{N}}$:*

$$\left|K_p(\underline{x}) - \int_{\tilde{\Omega}} \int_X K(y, T^1_{\underline{\omega}}y, \ldots, T^{p-1}_{\underline{\omega}}y, x_p, \ldots) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega})\right| \leq C \sum_{j=0}^{p-1} \mathrm{Lip}_j(K)\rho^{p-j}.$$

Le reste de cette section est consacrée à la démonstration du lemme. Pour une suite $\underline{\omega} \in \tilde{\Omega}$, on pose $g_{\underline{\omega}}^{(p)}(y) = \frac{h(y)}{h(T_{\underline{\omega}}^p y)} \frac{1}{|(T_{\underline{\omega}}^p)'(y)|}$. On a alors

$$K_p(\underline{x}) = \sum_{\underline{\omega} \in \Omega^p} p_{\underline{\omega}}^p \sum_{T_{\underline{\omega}}^p y = x} g_{\underline{\omega}}^{(p)}(y) K(y, T_{\underline{\omega}}^1 y, \ldots, T_{\underline{\omega}}^{p-1} y, x_p \ldots). \tag{3.10}$$

En effet, on a $K_p(\underline{x}) = \tilde{P}^p(K)(\tau^p(\underline{x})) = \tilde{P}^p(K)(x_p, \ldots)$, où $\tilde{P}$ est l'opérateur de transfert associé au système $(\Omega^\star, \tau, \mu_c)$, qui est donné par la formule valable pour tout $K \in L^1(\mu_c)$

$$\tilde{P}(K)(\underline{x}) = \sum_{\underline{\omega} \in \Omega^p} p_{\underline{\omega}}^p \sum_{T_{\underline{\omega}}^p y = x_0} g_{\underline{\omega}}^{(p)}(y) K(y, T_{\underline{\omega}}^1 y, \ldots, T_{\underline{\omega}}^{p-1} y, x_0 \ldots),$$

comme on le vérifie aisément par un changement de variables.
On fixe $x_\star \in X$ et on décompose K_p comme

$$K_p(\underline{x}) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \sum_{i=0}^{p-1} \sum_{\underline{\omega} \in \Omega^p} p_{\underline{\omega}}^p \sum_{T_{\underline{\omega}}^p y = x} g_{\underline{\omega}}^{(p)}(y) H_i(y, \ldots, T_{\underline{\omega}}^i y),$$

avec $H_i(y_0, \ldots, y_i) = K(y_0, \ldots, y_i, x_\star, \ldots, x_\star, x_p, \ldots) - K(y_0, \ldots, y_{i-1}, x_\star, \ldots, x_\star, x_p, \ldots)$.
Un simple calcul permet alors d'établir que

$$K_p(\underline{x}) = K(x_\star, \ldots, x_\star, x_p, \ldots) + \sum_{i=0}^{p-1} L^{p-i} f_i(x_p)$$

avec

$$f_i(y) = \sum_{\underline{\omega} \in \Omega^i} p_{\underline{\omega}}^i \sum_{T_{\underline{\omega}}^i z = y} g_{\underline{\omega}}^{(i)}(z) H_i(z, \ldots, T_{\underline{\omega}}^i z).$$

Comme L a un trou spectral sur BV, il existe $C > 0$ et $\rho < 1$ ne dépendant que du système aléatoire tels que $\|L^{p-i} f_i - \int_X f_i d\mu\|_{\mathrm{BV}} \leq C\rho^{p-i} \|f_i\|_{\mathrm{BV}}$. D'une part, $\|.\|_{\sup} \leq \|.\|_{\mathrm{BV}}$, on a $|L^{p-i} f_i(x_p) - \int_X f_i d\mu| \leq C\rho^{p-i} \|f_i\|_{\mathrm{BV}}$. D'autre part, on vérifie facilement $\int_X f_i d\mu = \int_{\tilde{\Omega}} \int_X H_i(y, \ldots, T_{\underline{\omega}}^i y) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega})$, d'où l'on déduit, en sommant toutes les relations :

$$\left| K_p(\underline{x}) - \int_{\tilde{\Omega}} \int_X K(y, T_{\underline{\omega}}^1 y, \ldots, T_{\underline{\omega}}^{n-1} y, x_p, \ldots) d\mu(y) d\tilde{\mathbb{P}}(\underline{\omega}) \right| \leq C \sum_{i=0}^{p-1} \rho^{p-i} \|f_i\|_{\mathrm{BV}}. \tag{3.11}$$

Il faut donc estimer $\|f_i\|_{\mathrm{BV}}$. Pour cela, on aura besoin d'un lemme technique. Pour $\omega \in \Omega$, on note $\mathcal{A}_\omega$ la partition de monotonicité de T_ω, et pour $\underline{\omega} \in \Omega^{\mathbb{N}}$, on définit $\mathcal{A}_{\underline{\omega}}^{n-1} = \bigvee_{k=0}^{n-1} \left(T_{\underline{\omega}}^k\right)^{-1} \left(\mathcal{A}_{\omega_{k+1}}\right)$ qui est la partition de monotonicité de $T_{\underline{\omega}}^n$. On rappelle aussi que si T est une application de type Lasota-Yorke, $\mathrm{Dist}(T)$ désigne sa distorsion.

Lemme 3.3.4. *Il existe $\lambda > 1$ et $C > 0$ tels que pour tout $\underline{\omega} \in \Omega^{\mathbb{N}}$ et tout $n \geq 0$:*

1. $\lambda(T^n_{\underline{\omega}}) \geq \lambda^n$,
2. $\mathrm{Dist}(T^n_{\underline{\omega}}) \leq C$,
3. $\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \sup_I \frac{1}{|(T^n_{\underline{\omega}})'|} \leq C$,
4. $\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \mathrm{Var}_I \left(\frac{1}{|(T^n_{\underline{\omega}})'|}\right) \leq C$.

Démonstration. Le point 1 est évident car Ω est fini, et chaque T_ω est uniformément dilatante. Le point 2 se démontre exactement de la même manière que le lemme 3.2.5. Pour démontrer le point 3, on adapte la preuve du lemme II.4 de [44]. Pour chaque $\underline{\omega} \in \Omega^n$ et chaque $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$, il existe un plus petit entier $p = p_{\underline{\omega},I} \leq n-1$ tel que $T^p_{\underline{\omega}}(I) \cap \partial\mathcal{A}_{\omega_{p+1}} \neq \emptyset$. Soit $\mathcal{A}^{n-1,p}_{\underline{\omega}}$ l'ensemble de tous les $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$ pour lesquels $p = p_{\underline{\omega},I}$. On définit $\partial = \cup_{\omega\in\Omega}\partial\mathcal{A}_\omega$. Soit alors $I \in \mathcal{A}^{n-1,p}_{\underline{\omega}}$. Il existe $a \in \partial I$ tel que $b = T^p_{\underline{\omega}} a \in \partial$. D'après le point 2, il existe une constante $C > 0$, ne dépendant que du système aléatoire, telle que pour tout $x \in I$,

$$|(T^n_{\underline{\omega}})'(x)| \geq C|(T^n_{\underline{\omega}})'(a)| = C|(T_{\omega_n} \circ \ldots \circ T_{\omega_{p+1}})'(b)||(T^p_{\underline{\omega}})'(a)| \geq C\lambda^{n-p}|(T^p_{\underline{\omega}})'(a)|.$$

On a alors

$$\sup_I \frac{1}{|(T^n_{\underline{\omega}})'|} \leq C^{-1}\lambda^{-(n-p)} \frac{1}{|(T^p_{\underline{\omega}})'(a)|}.$$

Comme une pré-image par $T^p_{\underline{\omega}}$ d'un élément $b \in \partial$ ne peut appartenir qu'à au plus deux $I \in \mathcal{A}^{n-1}_{\underline{\omega}}$ distincts, il s'ensuit que

$$\begin{aligned}
\sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{I\in\mathcal{A}^{n-1}_{\underline{\omega}}} \sup_I \frac{1}{|(T^n_{\underline{\omega}})'|} &\leq 2C^{-1} \sum_{p=0}^{n-1} \lambda^{-(n-p)} \sum_{b\in\partial} \sum_{\underline{\omega}\in\Omega^n} p^n_{\underline{\omega}} \sum_{T^p_{\underline{\omega}} a = b} \frac{1}{|(T^p_{\underline{\omega}})'(a)|} \\
&= 2C^{-1} \sum_{p=0}^{n-1} \lambda^{-(n-p)} \sum_{b\in\partial} P^p \mathbb{1}(b).
\end{aligned}$$

Cette quantité est bornée, puisque P a ses itérés bornés dans la norme d'opérateur sur BV, et ∂ est un ensemble fini. Le point 4 découle des trois précédents, de la même manière que pour le corollaire 3.2.7. □

Comme $L^i\mathbb{1} = \mathbb{1}$, on a $\|f_i\|_{\sup} \leq \|H_i\|_{\sup} \leq \mathrm{Lip}_i(K)$. On estime à présente la variation totale de f_i. On remarque tout d'abord que

$$\mathrm{Var}(f_i) \leq \mathrm{Var}(\frac{1}{h})\|hf_i\|_{\sup} + \|\frac{1}{h}\|_{\sup}\mathrm{Var}(hf_i).$$

Comme $\|hf_i\|_{\sup} \leq \mathrm{Lip}_i(K)\|P^i h\|_{\sup} \leq C\mathrm{Lip}_i(K)$, on doit seulement estimer $\mathrm{Var}(hf_i)$.
Pour $\underline{\omega} \in \Omega^i$ et $I \in \mathcal{A}^{i-1}_{\underline{\omega}}$, on note $S_{i,I,\underline{\omega}}$ la branche inverse de $T^i_{\underline{\omega}}$ restreinte à I. On définit aussi $H_{i,\underline{\omega}}(z) = H_i(z, \ldots, T^i_{\underline{\omega}} z)$.
On peut alors écrire

$$hf_i = \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}} \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \left(\frac{hH_{i,\underline{\omega}}}{|(T^i_{\underline{\omega}})'|}\right)\circ S_{i,I,\underline{\omega}}\,\mathbb{1}_{T^i_{\underline{\omega}}(I)}.$$

Il s'ensuit que

$$\begin{aligned}\mathrm{Var}(hf_i) \leq & \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}} \left(\sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \mathrm{Var}_I\left(\frac{hH_{i,\underline{\omega}}}{|(T^i_{\underline{\omega}})'|}\right) + 2\sum_{a\in\partial\mathcal{A}^{i-1}_{\underline{\omega}}} \frac{|h(a)||H_{i,\underline{\omega}}(a)|}{|(T^i_{\underline{\omega}})'(a)|}\right)\\ \leq & \sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}} \left((\mathrm{I})_{\underline{\omega},i} + (\mathrm{II})_{\underline{\omega},i} + (\mathrm{III})_{\underline{\omega},i} + (\mathrm{IV})_{\underline{\omega},i}\right) \qquad ,\end{aligned}$$

avec

$$\begin{aligned}(\mathrm{I})_{\underline{\omega},i} &= \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \mathrm{Var}_I(h)\sup_I \frac{1}{|(T^i_{\underline{\omega}})'|}\sup_I |H_{i,\underline{\omega}}|,\\ (\mathrm{II})_{\underline{\omega},i} &= \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \sup_I h\,\mathrm{Var}_I\left(\frac{1}{|(T^i_{\underline{\omega}})'|}\right)\sup_I |H_{i,\underline{\omega}}|,\\ (\mathrm{III})_{\underline{\omega},i} &= \sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \sup_I h\sup_I \frac{1}{|(T^i_{\underline{\omega}})'|}\mathrm{Var}_I(H_{i,\underline{\omega}}),\\ (\mathrm{IV})_{\underline{\omega},i} &= 2\sum_{a\in\partial\mathcal{A}^{i-1}_{\underline{\omega}}} \frac{|h(a)||H_{i,\underline{\omega}}(a)|}{|(T^i_{\underline{\omega}})'(a)|}.\end{aligned}$$

En utilisant la condition de Lipschitz sur K, on a

$$(\mathrm{I})_{\underline{\omega},i} \leq C\mathrm{Lip}_i(K)\sum_{I\in\mathcal{A}^{i-1}_{\underline{\omega}}} \sup_I \frac{1}{|(T^i_{\underline{\omega}})'|},$$

ce qui donne $\sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}\,(\mathrm{I})_{\underline{\omega},i} \leq C\mathrm{Lip}_i(K)$ par le lemme 3.3.4. Le même argument s'applique pour donner $\sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}\,(\mathrm{II})_{\underline{\omega},i} \leq C\mathrm{Lip}_i(K)$.
On estime à présent le terme $(\mathrm{III})_{\underline{\omega},i}$. Soit $y_0 < \ldots < y_l$ une suite de points de I. Afin d'estimer $\sum_{j=0}^{l-1} |H_{i,\underline{\omega}}(y_{j+1}) - H_{i,\underline{\omega}}(y_j)|$, on découpe $H_{i,\underline{\omega}}$ en deux termes de façon naturelle, et on considère le premier, le deuxième étant totalement similaire. On a

$$\begin{aligned}&\sum_{j=0}^{l-1}\sum_{k=0}^{i} |K(y_{j+1},\ldots,T^k_{\underline{\omega}}y_{j-1},T^{k+1}_{\underline{\omega}}y_j,\ldots,T^i_{\underline{\omega}}y_j,\ldots) - K(y_{j+1},\ldots,T^{k-1}_{\underline{\omega}}y_{j+1},T^k_{\underline{\omega}}y_j,\ldots,T^i_{\underline{\omega}}y_j,\ldots)|\\ &\leq \sum_{j=0}^{l-1}\sum_{k=0}^{i} \mathrm{Lip}_k(K)\left|T^k_{\underline{\omega}}y_{j+1} - T^k_{\underline{\omega}}y_j\right| = \sum_{k=0}^{i}\mathrm{Lip}_k(K)m(T^k_{\underline{\omega}}(I)).\end{aligned}$$

Comme $I\in\mathcal{A}^{i-1}_{\underline{\omega}}$, $T^k_{\underline{\omega}}(I)$ est inclus dans un intervalle de monotonicité de $T_{\omega_i}\circ\ldots\circ T_{\omega_{k+1}}$, et donc sa longueur est plus petite ou égale à $\left(\lambda(T_{\omega_i}\circ\ldots\circ T_{\omega_{k+1}})\right)^{-1} \leq \lambda^{-(i-k)}$. De plus, on a

$\mathrm{Var}_I(H_{i,\underline{\omega}}) \leq \sum_{k=0}^{i} \lambda^{-(i-k)}\mathrm{Lip}_k(K)$. Une application du lemme 3.3.4 montre que

$$\sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}\,(\mathrm{III})_{\underline{\omega},i} \leq C\sum_{k=0}^{i} \lambda^{-(i-k)}\mathrm{Lip}_k(K).$$

Une autre application du lemme 3.3.4 donne grâce à la condition de Lipschitz sur K que $\sum_{\underline{\omega}\in\Omega^i} p^i_{\underline{\omega}}\,(\mathrm{IV})_{\underline{\omega},i} \leq C\mathrm{Lip}_i(K)$.
En regroupant toutes les inégalités, on obtient $\mathrm{Var}(hf_i) \leq C\sum_{k=0}^{i} \lambda^{-(i-k)}\mathrm{Lip}_k(K)$, ce qui donne $\mathrm{Var}(f_i) \leq C\sum_{k=0}^{i} \lambda^{-(i-k)}\mathrm{Lip}_k(K)$, et la même inégalité pour $\|f_i\|_{\mathrm{BV}}$.
Par (3.11), on trouve

$$\left|K_p(\underline{x}) - \int_{\tilde{\Omega}}\int_X K(y, T^1_{\underline{\omega}}y, \ldots, T^{p-1}_{\underline{\omega}}y, x_p, \ldots)d\mu(y)d\tilde{\mathbb{P}}(\underline{\omega})\right| \leq C\sum_{i=0}^{p-1}\rho^{p-i}\sum_{k=0}^{i}\lambda^{-(i-k)}\mathrm{Lip}_k(K).$$

Un simple calcul montre que ce terme est inférieur à $C\sum_{k=0}^{p-1}(\rho')^{p-k}\mathrm{Lip}_k(K)$, pour $\max(\rho, \lambda^{-1}) < \rho' < 1$ ce qui prouve le lemme 3.3.3 et achève la démonstration du théorème 3.3.1. □

3.3.2 Applications

On décrit ici quelques applications du théorème 3.3.1.

Grandes déviations pour les sommes ergodiques

En considérant des observables K de la forme

$$K(x_0, \ldots, x_{n-1}) = \sum_{k=0}^{n-1} f_k(x_k),$$

on obtient :

Proposition 3.3.5. *Il existe $C > 0$ telle que pour toute suite $f_n : [0,1] \to \mathbb{R}$ d'observables Lipschitz, tout $n \geq 1$ et tout $t > 0$:*

$$\tilde{\mathbb{P}}\otimes\mu\left(\left\{(\underline{\omega}, x)\ :\ \frac{1}{n}\sum_{k=0}^{n-1}\left[f_k(T^k_{\underline{\omega}}x) - \int f_k\circ T^k_{\underline{\omega}}\,d\mu\right] > t\right\}\right) \leq e^{-\frac{Cn^2t^2}{\sum_{k=0}^{n-1}\mathrm{Lip}^2(f_k)}}.$$

Mesure empirique

Pour $(\underline{\omega}, x) \in \tilde{\Omega}\times X$, on définit pour $n \geq 1$ la mesure empirique

$$\mathcal{E}_n(\underline{\omega}, x) = \frac{1}{n}\sum_{j=0}^{n-1}\delta_{T^j_{\underline{\omega}}x}.$$

D'après la loi des grands nombres 1.2.2, $\mathcal{E}_n(\underline{\omega}, x)$ converge faiblement vers μ pour $\tilde{\mathbb{P}} \otimes \mu$-p.t. $(\underline{\omega}, x)$. On peut alors estimer la vitesse de convergence dans la métrique de Kantorovich κ définie par (3.9). La même preuve que pour la proposition 3.2.11 donne le résultat suivant, la seule différence dans la démonstration réside dans le fait qu'on utilisera la décroissance exponentielle des corrélations 1.2.1 au lieu de la condition **(Dec)**.

Proposition 3.3.6. *Il existe $t_0 > 0$ et $C > 0$ telles que pour tout $t > t_0$ et tout $n \geq 1$,*

$$\tilde{\mathbb{P}} \otimes \mu \left(\kappa(\mathcal{E}_n, \mu) > \frac{t}{\sqrt{n}} \right) \leq e^{-Ct^2}.$$

Approximation des orbites

Les mêmes démonstrations que pour les propositions 3.2.12 et 3.2.13 donnent les résultats suivants :

Proposition 3.3.7. *Il existe deux constantes $C_1, C_2 > 0$ telles que pour tout $A \subset [0,1]$ avec $\mu(A) > 0$ et tout $n \geq 1$, la suite de fonctions $(\mathcal{Z}_n)$ définies par*

$$\mathcal{Z}_n(\underline{\omega}, x) = \inf_{y \in A} \sum_{k=0}^{n-1} \left| T_{\underline{\omega}}^k x - T_{\underline{\omega}}^k y \right|,$$

satisfait pour tout $t > 0$:

$$\tilde{\mathbb{P}} \otimes \mu \left(\mathcal{Z}_n > C_1 \frac{\sqrt{|\log \mu(A)|}}{\sqrt{n}} + \frac{t}{\sqrt{n}} \right) \leq e^{-C_2 t^2}.$$

Proposition 3.3.8. *Il existe deux constantes $C_1, C_2 > 0$ telles que pour tout $A \subset [0,1]$ avec $\mu(A) > 0$, tout $n \geq 1$ et tout $0 < \epsilon < \frac{1}{2}$, la suite de fonctions $(\mathcal{Z}'_{n,\epsilon})$ de fonctions définies par*

$$\mathcal{Z}'_{n,\epsilon}(\underline{\omega}, x) = \frac{1}{n} \inf_{y \in A} \sharp\{0 \leq k < n \,:\, |T_{\underline{\omega}}^k x - T_{\underline{\omega}}^k y| > \epsilon\},$$

satisfait pour tout $t > 0$:

$$\tilde{\mathbb{P}} \otimes \mu \left(\mathcal{Z}'_{n,\epsilon} > C_1 \frac{\sqrt{|\log \mu(A)|}}{\epsilon\sqrt{n}} + \frac{t}{\epsilon\sqrt{n}} \right) \leq e^{-C_2 t^2}.$$

3.4 Remarques et questions ouvertes

1. On trouvera de nombreuses applications des inégalités de concentration pour les processus stationnaires dans [37]. Ces conséquences s'appliquent en particulier aux processus concernés par le théorème 3.3.1.

2. Est-il possible de démontrer des inégalités de concentration polynomiale pour les systèmes séquentiels avec un point fixe neutre étudiés au paragraphe 2.2 ? La méthode de [39] repose sur l'induction et la théorie du renouvellement pour les opérateurs, et semble donc difficilement applicable dans ce contexte. La méthode de [35], basée sur le couplage, pourrait peut-être être adaptée, mais elle ne donne dans le cas autonome qu'un résultat très faible, avec une inégalité de concentration d'ordre 2 pour une classe restreinte de paramètres α.
3. On a fait l'hypothèse au théorème 3.3.1 que l'ensemble des applications T_ω sont uniformément dilatantes. Le résultat reste-t-il vrai si on suppose uniquement que le système est dilatant en moyenne ? On dispose toujours du trou spectral sur BV, qui était un élément important de la démonstration, mais le lemme 3.3.4 n'est plus vérifié, et il est donc nécessaire de trouver un argument permettant de le remplacer.
4. Le théorème 3.3.1 est un résultat de type annealed. Peut-on en démontrer une version quenched, c'est-à-dire : est-ce que le processus $Z_{n,\underline{\omega}} = T^n_{\underline{\omega}}$, défini sur $([0,1],\mu)$, satisfait une inégalité de concentration pour $\tilde{\mathbb{P}}$-p.t. $\underline{\omega}$?
5. Est-il possible, dans le cas des systèmes déterministes et autonomes, d'utiliser des espaces fonctionnels plus compliqués, comme Quasi-Hölder ou BV en dimension supérieure, afin de démontrer des inégalités de concentration exponentielle pour des systèmes uniformément dilatants multidimensionnels ? La même question se pose pour les systèmes uniformément hyperboliques, avec comme possibilité l'utilisation des espaces fonctionnels développés dans [24, 62, 20] et bien d'autres références.

Chapitre 4

Lemmes de Borel-Cantelli et statistiques de récurrence pour l'induction de Rauzy-Veech-Zorich

Dans ce chapitre, on s'intéressera aux lemmes de Borel-Cantelli et aux statistiques de récurrence pour l'application de Rauzy-Veech-Zorich. La présentation de celle-ci s'inspire pour beaucoup des textes de Pollicott [124] et Viana [141]. Les résultats de ce chapitre sont issus de l'article [4] coécrit avec Matthew Nicol et Michael Todd, en cours de soumission.

4.1 L'induction de Rauzy-Veech-Zorich

Dans cette section, on suit de près la présentation de Viana [141], auquel on renvoie pour plus de précisions.

4.1.1 Échanges d'intervalle

Soit $I \subset \mathbb{R}$ un intervalle, dont l'extrémité gauche est 0. Par la suite, tous les intervalles seront bornés, fermés à gauche et ouverts à droite. Soit $\mathcal{A}$ un alphabet fini avec $d \geq 2$ symboles, et $\{I_a : a \in \mathcal{A}\}$ une partition de I en intervalles indexés par $\mathcal{A}$. Un échange d'intervalles est une application bijective $f = f_{(\pi,\lambda)} : I \to I$ qui est une translation sur chaque sous-intervalle I_a, préserve la mesure de Lebesgue, et est déterminée par les données combinatoires suivantes :

(a) Un couple $\pi = (\pi_0, \pi_1)$ de bijections $\pi_\epsilon : \mathcal{A} \to \{1, \ldots, d\}$ qui décrivent l'ordre des sous-intervalles I_a avant et après l'application de f :

$$\begin{pmatrix} a_1^0 & a_2^0 & \ldots & a_d^0 \\ a_1^1 & a_2^1 & \ldots & a_d^1 \end{pmatrix}$$

où $a_j^\epsilon = \pi_\epsilon^{-1}(j)$ pour $\epsilon \in \{0, 1\}$ et $j \in \{1, \ldots, d\}$.

(b) Un vecteur $\lambda = (\lambda_a)_{a\in\mathcal{A}}$ où chaque élément $\lambda_a > 0$ représente la longueur du sous-intervalle I_a. On notera $|\lambda| = |I| = \sum_a \lambda_a$.

La permutation $p = \pi_1 \circ \pi_0^{-1}$ est appelée l'invariant de monodronie du couple $\pi = (\pi_0, \pi_1)$. Comme observé par Viana [141], on peut toujours se ramener au cas où $\mathcal{A} = \{1, \dots, d\}$ et π_0, en remplaçant π_1 par p. Néanmoins, on conservera cette notation qui est invariante sous les algorithmes d'induction et de renormalisation décrits plus loin.
On peut donner une expression explicite pour l'échange d'intervalles $f_{(\pi,\lambda)}$. Pour $a \in \mathcal{A}$, on définit

$$w_a = \sum_{\{b:\pi_1(b)<\pi_1(a)\}} \lambda_b - \sum_{\{b:\pi_0(b)<\pi_0(a)\}} \lambda_b.$$

Alors

$$f_{(\pi,\lambda)}(x) = x + \sum_a w_a \mathbb{1}_{I_a}(x).$$

4.1.2 Induction de Rauzy-Veech et classes de Rauzy

Comme il est courant avec des familles de systèmes dynamiques présentant un comportement de type parabolique, une manière de procéder à leur étude est de définir un bon schéma de renormalisation sur l'espace des paramètres. Dans le contexte des échanges d'intervalle, une possibilité est d'utiliser l'induction de Rauzy, décrite ci-après.
Pour un échange d'intervalle f de données combinatoires (π, λ) et pour $\epsilon \in \{0, 1\}$, on notera $a(\epsilon)$ le dernier symbole dans l'expression de π_ϵ, c'est-à-dire $a(\epsilon) = \pi_\epsilon^{-1}(d) = a_d^\epsilon$. En supposant que l'on est dans la situation générique où $I_{a(0)}$ et $I_{a(1)}$ sont de longueurs différentes, on dira que

$$(\pi, \lambda) \text{ est de } \begin{cases} \text{type } 0 & \text{si} \quad \lambda_{a(0)} > \lambda_{a(1)}, \\ \text{type } 1 & \text{si} \quad \lambda_{a(0)} < \lambda_{a(1)}. \end{cases}$$

Maintenant, on pose

$$J = \begin{cases} I \setminus f_{(\pi,\lambda)}(I_{a(1)}) & \text{si } (\pi, \lambda) \text{ est de type } 0, \\ I \setminus I_{a(0)} & \text{si } (\pi, \lambda) \text{ est de type } 1. \end{cases}$$

On considère à présent l'application de premier retour de f dans J. Comme expliqué à la section 1.2 de [141], il s'agit à nouveau d'un échange d'intervalle, défini sur l'intervalle J, avec le même alphabet à d symboles $\mathcal{A}$, et de données combinatoires (π', λ'). Celles-ci peuvent être décrites de la façon suivante : si (π, λ) est de type 0, alors pour $k = \pi_1^{-1}(a(0))$ on a

$$\pi' = \begin{pmatrix} \pi'_0 \\ \pi'_1 \end{pmatrix} = \begin{pmatrix} a_1^0 & \dots & a_{k-1}^0 & a_k^0 & a_{k+1}^0 & \dots & \dots & a(0) \\ a_1^1 & \dots & a_{k-1}^1 & a(0) & a(1) & a_{k+1}^0 & \dots & a_{d-1}^1 \end{pmatrix}$$

et $\lambda' = (\lambda'_a)_{a\in\mathcal{A}}$ vérifie

$$\lambda'_a = \lambda_a \text{ pour } a \neq a(0), \text{ et } \lambda'_{a(0)} = \lambda_{a(0)} - \lambda_{a(1)}.$$

De même, si (π, λ) est de type 1, alors pour $k = \pi_0^{-1}(a(0))$ on a

$$\pi' = \begin{pmatrix} \pi'_0 \\ \pi'_1 \end{pmatrix} = \begin{pmatrix} a_1^0 & \dots & a_{k-1}^0 & a(1) & a(0) & a_{k+1}^0 & \dots & a_{d-1}^0 \\ a_1^1 & \dots & a_{k-1}^1 & a_k^1 & a_{k+1}^1 & \dots & \dots & a(1) \end{pmatrix}$$

et $\lambda' = (\lambda'_a)_{a \in \mathcal{A}}$ vérifie

$$\lambda'_a = \lambda_a \text{ pour } a \neq a(1), \text{ et } \lambda'_{a(1)} = \lambda_{a(1)} - \lambda_{a(0)}.$$

On notera $\Theta = \Theta_{(\pi,\lambda)}$ la matrice constituée uniquement de 0 et de 1 telle que $\lambda' = \Theta^{-1*}(\lambda)$. Celle-ci est donnée par les formules (1.9) et (1.10) de [141].
On pose

$$\hat{\mathcal{T}}_0(\pi, \lambda) = (\pi', \lambda').$$

Comme on l'a vu précédemment, $\hat{\mathcal{T}}_0(\pi, \lambda)$ est bien défini seulement si $I_{a(0)}$ et $I_{a(1)}$ sont de longueur différente. On s'intéresse à l'ensemble des (π, λ) pour lesquels $\hat{\mathcal{T}}_0^n(\pi, \lambda)$ est bien défini pour tout $n \geq 0$.
On supposera dans toute la suite que π est irréductible, c'est-à-dire qu'il n'existe pas de $k \in \{1, \dots, d-1\}$ tel que $p(\{1, \dots, k\}) = \{1, \dots, k\}$.
On dira que (π, λ) satisfait la condition de Keane si

$$f^n_{(\pi,\lambda)}(\partial I_a) \neq \partial I_b \text{ pour tout } n \geq 1 \text{ et } a, b \in \mathcal{A} \text{ avec } \pi_0(b) \neq 1,$$

où ∂I_a est l'extrémité gauche de l'intervalle I_a.
L'induction de Rauzy $\hat{\mathcal{T}}_0^n(\pi, \lambda)$ est définie pour tout $n \geq 0$ si et seulement si (π, λ) satisfait la condition de Keane, voir la section 1.3 et le corollaire 1.22 de [141]. C'est en particulier le cas si π est irréductible, et si le vecteur λ est irrationnel, c'est-à-dire que

$$\sum_{a \in \mathcal{A}} \eta_a \lambda_a \neq 0$$

pour tout vecteur $\eta = (\eta_a)_{a \in \mathcal{A}} \in \mathbb{Z}^{\mathcal{A}}$ non nul, puisque de telles données combinatoires vérifient la condition de Keane, d'après la proposition de 1.10 de [141].
Pour un entier $d \geq 2$ fixé, on définit la classe de Rauzy $\mathcal{R} = \mathcal{R}(\pi)$ d'un couple π comme l'ensemble des couples π' pour lesquels il existe $n \geq 0$ et deux vecteurs λ, λ' tels que $\hat{\mathcal{T}}_0^n(\pi, \lambda) = (\pi', \lambda')$. Les classes de Rauzy forment une partition de l'ensemble des couples π, et on peut considérer l'induction de Rauzy $\hat{\mathcal{T}}_0$ comme une application agissant sur les ensembles de la forme $\mathcal{R} \times \mathbb{R}_+^{\mathcal{A}}$. Pour $d = 2$ et $d = 3$, il n'existe qu'une seule classe de Rauzy, mais ce n'est plus nécessairement vrai lorsque $d \geq 4$. On renvoie là encore à Viana [141] pour une description de celles-ci.
L'application de renormalisation $\mathcal{T}_0$ de Rauzy-Veech est simplement la transformation $\hat{\mathcal{T}}_0$ renormalisée de telle sorte à ce qu'elle agisse sur l'ensemble des échanges d'intervalle définis sur $I = [0, 1[$. Plus précisément, si on note

$$\Delta = \Delta_{\mathcal{A}} = \{\lambda = (\lambda_a)_{a \in \mathcal{A}} \,:\, \lambda_a > 0, \sum_a \lambda_a = 1\}$$

l'ensemble des vecteurs positifs de longueur 1, alors pour une classe de Rauzy $\mathcal{R}$ donnée, $\mathcal{T}_0 : \mathcal{R} \times \Delta_{\mathcal{A}} \to \mathcal{R} \times \Delta_{\mathcal{A}}$ est définie par

$$\mathcal{T}_0(\pi, \lambda) = (\pi', \frac{\lambda'}{|\lambda'|}),$$

si $\hat{\mathcal{T}}_0(\pi, \lambda) = (\pi', \lambda')$. On remarque que $|\lambda'| = 1 - \lambda_{a(1-\epsilon)}$ si (π, λ) est de type ϵ.
Pour $\pi \in \mathcal{R}$ et $\epsilon \in \{0, 1\}$, on pose

$$\Delta_{\pi,\epsilon} := \left\{\lambda \in \Delta_{\mathcal{A}} \,:\, \lambda_{a(\epsilon)} > \lambda_{a(1-\epsilon)}\right\}.$$

Alors $\mathcal{T}_0 : \{\pi\} \times \Delta_{\pi,\epsilon} \to \{\pi'\} \times \Delta_{\mathcal{A}}$ est une bijection.
Soit $d\pi$ la mesure de comptage sur la classe de Rauzy $\mathcal{R}$ et m la mesure de Lebesgue (de dimension $d-1$) sur le simplexe $\Delta_{\mathcal{A}}$. Masur [107] et Veech [140] ont indépendamment montré qu'il existe, à un facteur multiplicatif près, une unique mesure μ_0 σ-finie sur $\mathcal{R} \times \Delta_{\mathcal{A}}$ absolument continue par rapport à $d\pi \times m$ et invariante par $\mathcal{T}_0$. De plus cette mesure est conservative et ergodique, et est équivalente à $d\pi \times m$, avec une densité bornée uniformément loin de 0.

4.1.3 Induction de Zorich

En général, la mesure μ_0 de Masur et Veech n'est pas de masse totale finie. Afin de se ramener au cas d'une mesure de probabilité invariante, Zorich [145] a introduit une accélération de l'induction de Rauzy, que l'on va maintenant décrire.
Soit $\mathcal{R}$ une classe de Rauzy. Pour un couple $(\pi, \lambda) \in \mathcal{R} \times \Delta_{\mathcal{A}}$ satisfaisant la condition de Keane, on note $(\pi^n, \lambda^n) = \hat{\mathcal{T}}_0^n(\pi, \lambda)$, et ϵ^n le type de (π^n, λ^n). On pose aussi $\epsilon = \epsilon^0$ le type de (π, λ). On définit alors n_1 par

$$n_1(\pi, \lambda) = \inf\{k \geq 0 \,:\, \epsilon^k \neq \epsilon\}.$$

Comme μ_0 est conservative pour $\mathcal{T}_0$, n_1 est bien définie μ_0-p.p. (et donc $d\pi \times m$-p.p.)
L'induction de Zorich $\hat{\mathcal{T}}_1$ est alors définie par

$$\hat{\mathcal{T}}_1(\pi, \lambda) = \hat{\mathcal{T}}_0^{n_1(\pi,\lambda)}(\pi, \lambda),$$

et de même, l'application de renormalisation de Zorich $\mathcal{T}_1$ est définie par $\mathcal{T}_1 = \mathcal{T}_0^{n_1}$.
Cette application admet une partition de Markov dénombrable : en effet, si on pose

$$\Delta_{\pi,\epsilon,n} = \{\lambda \in \Delta_{\pi,\epsilon} \,:\, \epsilon = \epsilon^1 = \ldots = \epsilon^{n-1} \neq \epsilon^n\}$$

alors pour tout $\pi \in \mathcal{R}$, $\mathcal{T}_1 : \{\pi\} \times \Delta_{\pi,\epsilon,n} \to \{\pi^n\} \times \Delta_{\pi^n,1-\epsilon}$ est une bijection. De plus,

$$\lambda^n = c_n \Theta^{-n*}(\lambda),$$

où c_n et Θ^{-n*} ne dépendent que de π, ϵ et n.
Zorich [145] a démontré qu'il existe une unique mesure de probabilité μ_1 sur $\mathcal{R} \times \Delta_{\mathcal{A}}$ absolument continue par rapport à $d\pi \times m$, et invariante par $\mathcal{T}_1$. De plus cette mesure est ergodique, et admet deux composantes mélangeantes : si pour $\epsilon \in \{0, 1\}$, on pose $\Delta_\epsilon = \cup_{\pi \in \mathcal{R}} \{\pi\} \times \Delta_{\pi,\epsilon}$, alors $\mathcal{T}_1(\Delta_\epsilon) = \Delta_{1-\epsilon}$ et $\mathcal{T}_1^2 : \Delta_\epsilon \to \Delta_\epsilon$ est mélangeante par rapport à $2\mu_1$ restreinte à Δ_ϵ.

4.1.4 Une nouvelle induction

L'application $\mathcal{T}_1$ n'est pas uniformément dilatante, et admet des directions neutres sur le bord du simplexe. Afin de se ramener au cas d'une application uniformément dilatante, on induit sur un sous-ensemble B de la partition dynamique loin du bord. Cette approche a par exemple été décrite par Morita [118] et Pollicott [124].
On définit la partition

$$\mathcal{P} = \{\{\pi\} \times \Delta_{\pi,\epsilon} \,:\, \pi \in \mathcal{R}, \epsilon \in \{0,1\}\},$$

et pour $n \geq 1$ son rafinement

$$\mathcal{P}_n = \bigvee_{k=0}^{n-1} \mathcal{T}_1^{-k}\mathcal{P}.$$

On peut alors choisir un entier $n_B > 1$ et un élément $B \in \mathcal{P}_{n_B}$ tels que la branche inverse de la restriction de $\mathcal{T}_1^{n_B}$ à B est une contraction stricte, voir par exemple le corollaire 1.21 de Viana [141]. On définit alors n_2 comme le premier temps de retour de $\mathcal{T}_1$ dans B :

$$n_2(\pi,\lambda) = \inf\{k \geq 1 \,:\, \mathcal{T}_1^k(\pi,\lambda) \in B\},$$

et $\mathcal{T}_2 : B \to B$ est définie comme l'application de premier retour correspondante :

$$\mathcal{T}_2(\pi,\lambda) = \mathcal{T}_1^{n_2(\pi,\lambda)}(\pi,\lambda).$$

Remarque 4.1.1. *On notera que pour tout élément $(\pi,\lambda) \in \mathcal{R}\times\Delta_{\mathcal{A}}$ qui satisfait la condition de Keane, on peut trouver un tel élément B contenant (π,λ).*

L'ensemble B a une partition naturelle dénombrable $\mathcal{Q} = \{B_i\}_{i\in\mathcal{I}}$ en sous-ensembles sur lesquels n_2 est constante. L'application $\mathcal{T}_2 : B_i \to B$ est un difféomorphisme pour chaque $i \in \mathcal{I}$, d'après [118, Lemme 3.1]. Il existe sur B une unique mesure de probabilité μ_2 absolument continue et invariante par $\mathcal{T}_1$, qui est en fait la restriction, convenablement renormalisée, de μ_1 à B. En particulier, sa densité est bornée inférieurement loin de 0.
Soit $\mathrm{Jac}(\mathcal{T}_2^n)$ le jacobien de $\mathcal{T}_2^n$ définie par rapport à la mesure de Lebesgue m, c'est à dire que pour tout sous-ensemble $A \subset B$ sur lequel $\mathcal{T}_2^n$ est injective, on a $m(\mathcal{T}_2^n(A)) = \int_A \mathrm{Jac}(\mathcal{T}_2^n)\,dm$.
Soit $\mathcal{Q}_n = \bigvee_{k=0}^{n-1} \mathcal{T}_2^{-k}\mathcal{Q}$.
On a le contrôle suivant sur la distorsion et la dilatation de $\mathcal{T}_2$:

Lemme 4.1.2 ([118, Lemme 3.4], [124, Lemme 2.2])**.** *Il existe $C > 0$, $\theta > 1$ et $D_1, D_2 > 0$ telles que pour tout $n \geq 1$ et tout x, y dans le même élément de la partition $\mathcal{Q}_n$:*

1. $d(\mathcal{T}_2^n x, \mathcal{T}_2^n y) \geq C\theta^n d(x,y)$;
2. $\left|\log\left(\frac{\mathrm{Jac}(\mathcal{T}_2^n)(x)}{\mathrm{Jac}(\mathcal{T}_2^n)(y)}\right)\right| \leq D_1 d(\mathcal{T}_2^n x, \mathcal{T}_2^n y)$;
3. $\frac{1}{D_2} \leq \mu_2(A)|\mathrm{Jac}(\mathcal{T}_2^n)(x)| \leq D_2$ *pour tout* $x \in A \in \mathcal{Q}_n$.

Remarque 4.1.3. *La distance d peut désigner indifféremment la distance euclidienne ou la distance de Hilbert, qui sont équivalentes sur B. D'autre part, on peut énoncer le point (3) du lemme précédent pour la mesure de Lebesgue m, puisque μ_2 est équivalente à m, avec une densité h_B vérifiant $c^{-1} \leq h_B \leq c$, pour une constante $c > 0$.*

4.1.5 L'opérateur de transfert de $\mathcal{T}_2$

On montre dans cette section que l'opérateur de transfert associé à $\mathcal{T}_2$ admet un trou spectral sur l'espace Quasi-Hölder. On utilisera une définition de cet espace plus proche de celle de Keller [90] que de celle de Saussol [129], qui a été rappelée à la section 1.1.3.
On en donne la définition. Soient $0 < \alpha \leq 1$ et $\epsilon_0 > 0$. Pour $f \in L^1_m(B)$, on définit l'oscillation de f sur le borélien $S \subset B$ par

$$\mathrm{osc}(f, S) = \mathrm{ess\,sup}_S f - \mathrm{ess\,inf}_S f.$$

On définit

$$|f|_\alpha := \sup_{0<\epsilon\leq\epsilon_0} \epsilon^{-\alpha} \int_B \mathrm{osc}(f, B_\epsilon(x) \cap B) dx$$

et on pose $V_\alpha := \{f \in L^1_m(B) \ : \ |f|_\alpha < \infty\}$. La principale différence avec la définition de Saussol est qu'on ne considère plus le compact B comme un sous-ensemble d'un espace euclidien, en prolongeant les fonctions initialement définies sur B en posant 0 en dehors de B. En particulier, on ne tient plus compte des discontinuités "artificielles" au bord de B, ce qui nous sera utile ici, puisque $\mathcal{T}_2$ est surjective par morceaux.
Muni de la norme $\|.\|_\alpha = |.|_\alpha + \|.\|_{L^1_m}$, V_α est un espace de Banach qui s'injecte de façon compacte dans $L^1_m(B)$, de façon continue dans $L^\infty_m(B)$, et est une algèbre de Banach satisfaisant $|fg|_\alpha \leq |f|_\alpha \|g\|_\infty + \|f\|_\infty |g|_\alpha$ pour tous $f, g \in V_\alpha$.
On remarque que la norme $\|.\|_\alpha$ dépend de façon implicite du choix de ϵ_0, bien que l'espace V_α n'en dépende pas, et que deux choix différents pour ϵ_0 donnent deux normes équivalentes.

Remarque 4.1.4. *Pour tout sous-ensemble borélien $A \subset B$, la norme Quasi-Hölder de sa fonction caractéristique vérifie $\|\mathbb{1}_A\|_\alpha \leq m(A) + \sup_{0<\epsilon\leq\epsilon_0} \frac{m(B_\epsilon(\partial A))}{\epsilon^\alpha}$. En particulier, toute suite de fonctions caractéristiques de boules est bornée dans V_α puisque B est de diamètre fini.*

Soit P l'opérateur de transfert de $\mathcal{T}_2$, défini par rapport à la mesure de Lebesgue m. Pour $\phi \in L^1_m(B)$, on a

$$P\phi(x) = \sum_i \frac{\phi(x_i)}{\mathrm{Jac}(\mathcal{T}_2)(x_i)},$$

où $x_i \in B_i$ vérifie $\mathcal{T}_2(x_i) = x$.
On va prouver qu'il existe des constantes $C_1, C_2 > 0$, $0 < \theta < 1$ telles que pour tout $\phi \in V_\alpha$ et tout $\psi \in L^1(\mu_2)$:
(a) $\|P^n\phi - (\int \phi dm)\, h_B\|_\alpha \leq C_1 \theta^n \|\phi\|_\alpha$;
(b) $|\int \phi\psi \circ \mathcal{T}_2^n d\mu_2 - \mu_2(\phi)\mu_2(\psi)| \leq C_2 \theta^n \|\phi\|_\alpha \|\psi\|_{L^1_{\mu_2}}$.
On va déduire ces deux propriétés d'une inégalité de Lasota-Yorke pour l'opérateur P agissant sur V_α :

Lemme 4.1.5. *Il existe $C, D > 0$ et $0 < \eta < 1$ telles que pour tout $\phi \in V_\alpha$ et tout $n \geq 0$,*

$$\|P^n\phi\|_\alpha \leq C\eta^n \|\phi\|_\alpha + D\|\phi\|_{L^1_m}.$$

Démonstration. Comme P^n est une contraction sur $L^1_m(B)$, il suffit d'estimer $|P^n\phi|_\alpha$. Pour $Q \in \mathcal{Q}_n$, on note $I_{n,Q} : B \to Q$ la branche inverse de la restriction de $\mathcal{T}_2^n$ à Q. L'opérateur P^n peut alors s'écrire :

$$P^n\phi(x) = \sum_{Q\in\mathcal{Q}_n} g_n(I_{n,Q}x)\phi(I_{n,Q}x),$$

où $g_n = \frac{1}{\mathrm{Jac}(\mathcal{T}_2^n)}$. Soit $M_{n,Q}$ l'opérateur défini par $M_{n,Q}\phi(x) = g_n(I_{n,Q}x)\phi(I_{n,Q}x)$. On énonce d'abord le lemme suivant, qui nous sera aussi utile par la suite :

Lemme 4.1.6. *Il existe $C > 0$ telle que pour tout $n \geq 1$, $Q \in \mathcal{Q}_n$ et $\phi \in V_\alpha$, on a $\|M_{n,Q}\phi\|_{L^1_m} = \int_Q |\phi|dm$ et*

$$\int osc(M_{n,Q}\phi, B_\epsilon(x))dx \leq C\int_Q osc(\phi, B_{c_{n,Q}\epsilon}(x))dx + C\epsilon\int_Q |\phi|dm,$$

où $c_{n,Q}$ est la constante de Lipschitz de $I_{n,Q} : B \to Q$.

En admettant temporairement ce lemme, on observe que $c_{n,Q} \leq C\theta^{-n} \leq C$ avec $\theta > 1$, d'après le point (1) du lemme 4.1.2. En écrivant $P^n = \sum_{Q\in\mathcal{Q}_n} M_{n,Q}$ et en sommant toutes les inégalités du lemme 4.1.6, on obtient par la proposition 3.2 (i) de [129] :

$$\begin{aligned}\int \mathrm{osc}(P^n\phi, B_\epsilon(x))dx &\leq C\int \mathrm{osc}(\phi, B_{C\theta^{-n}\epsilon}(x))dx + C\epsilon\|\phi\|_{L^1_m}\\ &\leq C\epsilon^\alpha\left(\theta^{-\alpha n}|\phi|_\alpha + \epsilon_0^{1-\alpha}\|\phi\|_{L^1_m}\right),\end{aligned}$$

Pour tout $0 < \epsilon \leq \frac{\epsilon_0}{C} =: \epsilon_1$, de telle sorte que $C\theta^{-n}\epsilon \leq \epsilon_0$ et que l'estimation

$$\int \mathrm{osc}(\phi, B_{C\theta^{-n}\epsilon}(x))dx \leq C\epsilon^\alpha\theta^{-\alpha n}|\phi|_\alpha$$

soit vérifiée.
On en déduit que $\|P^n\phi\|_{\alpha,\epsilon_1} \leq C\theta^{-\alpha n}\|\phi\|_{\alpha,\epsilon_0} + C\|\phi\|_{L^1_m}$. On indique en indice ϵ_0 ou ϵ_1 pour préciser si la norme Quasi-Hölder utilisée est définie avec ϵ_0 ou ϵ_1. Comme ces deux normes sont équivalentes, cela conclut la preuve. □

Démonstration du lemme 4.1.6. La relation $\int_B |M_{n,Q}\phi|dm = \int_Q |\phi|dm$ provient d'un simple changement de variables.
On observe que $\mathrm{osc}(M_{n,Q}\phi, B_\epsilon(x)) = \mathrm{osc}(g_n\phi, I_{n,Q}B_\epsilon(x))$. En utilisant la proposition 3.2 (iii) de [129], on trouve que pour tout $x \in B$,

$$\mathrm{osc}(M_{n,Q}\phi, B_\epsilon(x)) \leq \mathrm{osc}(\phi, I_{n,Q}B_\epsilon(x)) \sup_{I_{n,Q}B_\epsilon(x)} g_n + \mathrm{osc}(g_n, I_{n,Q}B_\epsilon(x)) \operatorname*{ess\,inf}_{I_{n,Q}B_\epsilon(x)} |\phi|.$$

Par le contrôle de la distorsion fourni par le lemme 4.1.2, on a $\operatorname*{ess\,sup}_{I_{n,Q}B_\epsilon(x)} g_n \leq Cg_n(I_{n,Q}x)$ et $\mathrm{osc}(g_n, I_{n,Q}B_\epsilon(x)) \leq Cg_n(I_{n,Q}x)\epsilon$ pour une constante $C > 0$. On a aussi $\mathrm{osc}(\phi, I_{n,Q}B_\epsilon(x)) \leq$

$\mathrm{osc}(\phi, B_{c_{n,Q}\epsilon}(I_{n,Q}x))$ et $\operatorname*{ess\,inf}_{I_{n,Q}B_\epsilon(x)} |\phi| \leq |\phi(I_{n,Q}x)|$ pour presque tout $x \in B$. En réunissant toutes ces inégalités, on obtient que pour presque tout x,

$$\mathrm{osc}(M_{n,Q}\phi, B_\epsilon(x)) \leq C\mathrm{osc}(\phi, B_{c_{n,Q}\epsilon}(I_{n,Q}x))g_n(I_{n,Q}x) + C\epsilon|\phi(I_{n,Q}x)|g_n(I_{n,Q}x).$$

Une intégration sur B suivie d'un changement de variables conclut la démonstration. □

D'après le théorème de Hennion [75], le lemme 4.1.5 entraine que P, vu comme opérateur sur V_α est quasi-compact et a un rayon spectral essentiel strictement plus petit que 1. Afin de démontrer le point (a), il suffit donc de montrer que 1 est une valeur propre simple de P, et qu'il n'existe pas d'autre valeur propre sur le cercle unité. Soit donc $\phi \in V_\alpha$ un vecteur propre de P associé à la valeur propre $\lambda \in \mathbb{C}$ avec $|\lambda| = 1$. D'après les résultats de Morita [118] et Pollicott [124], l'opérateur P est aussi quasi-compact sur l'espace des fonctions Hölder, avec un rayon spectral essentiel strictement inférieur à 1. Comme l'espace des fonctions Hölder est dense dans $L^1_m(B)$, on en déduit que ϕ admet une version Hölder. D'après le lemme 2.3 de [124], ϕ est un multiple de h_B, et $\lambda = 1$, ce qui démontre le point (a).
Le point (b) en découle immédiatement, comme pour le point (3) de la proposition 1.2.1.

4.2 Lemmes dynamiques de Borel-Cantelli

Dans cette section, on démontre des lemmes dynamiques de Borel-Cantelli pour les transformations $\mathcal{T}_2$ et $\mathcal{T}_1$. On rappelle tout d'abord les définitions. Soit (X, T, μ) un système dynamique où μ est une mesure de probabilité T-invariante et ergodique.

Définition 4.2.1. *Soit (A_n) une suite d'ensembles mesurables, vérifiant $\sum_n \mu(A_n) = \infty$. On dira que*

1. *(A_n) vérifie la propriété de Borel-Cantelli si pour μ-presque tout point $x \in X$, $T^n x \in A_n$ infiniment souvent ;*
2. *(A_n) vérifie la propriété forte de Borel-Cantelli si pour μ-presque tout $x \in X$,*

$$\frac{1}{E_n}\sum_{k=0}^{n-1} \mathbb{1}_{A_k}(T^k x) \to 1,$$

avec $E_n = \sum_{k=0}^{n-1} \mu(A_k)$.

La propriété forte de Borel-Cantelli s'étend aux fonctions positives : si (f_n) est une suite de fonctions mesurables positives avec $\sum_n \int_X f_n \, d\mu = \infty$, on dira qu'elle satisfait la propriété forte de Borel-Cantelli si

$$\frac{1}{E_n}\sum_{k=0}^{n-1} f_k(T^k x) \to 1$$

pour μ-p.t. $x \in X$, avec $E_n = \sum_{k=0}^{n-1} \int_X f_k \, d\mu$.

4.2.1 Lemmes dynamiques de Borel-Cantelli pour $\mathcal{T}_2$

Les lemmes de Borel-Cantelli pour $\mathcal{T}_2$ sont une conséquence directe de la décroissance des corrélations pour les observables dans V_α contre $L^1(\mu_2)$, et du théorème de Sprindzuk (Proposition 1.4.1) :

Proposition 4.2.2. *Soit $(\phi_n)_n$ une suite de fonctions positives, définies sur B et telle que $\sup_n \|\phi_n\|_\alpha < \infty$. Si $E_n = \sum_{k=0}^{n-1} \int_B \phi_k \, d\mu_2$ est non bornée, alors $(\phi_n)_n$ vérifie la propriété forte de Borel-Cantelli pour $\mathcal{T}_2$.*

En particulier, toute suite (A_n) de boules contenues dans B vérifie la propriété forte de Borel-Cantelli pour $\mathcal{T}_2$, d'après la remarque 4.1.4.

Démonstration. On applique le théorème de Sprindzuk avec $f_k = \phi_k \circ \mathcal{T}_2^k$, $\bar{f}_k = \varphi_k = \int_B \phi_k \, d\mu_2$. On a

$$\begin{aligned}
\left| \sum_{i=m}^{n} \sum_{j=i+1}^{n} \int_B \phi_j \circ \mathcal{T}_2^j \phi_i \circ \mathcal{T}_2^i \, d\mu_2 - \mu_2(\phi_j)\mu_2(\phi_i) \right| &= \left| \sum_{i=m}^{n} \sum_{j=i+1}^{n} \int_B \phi_j \circ \mathcal{T}_2^{j-i} \phi_i \, d\mu_2 - \mu_2(\phi_j)\mu_2(\phi_i) \right| \\
&\leq \sum_{i=m}^{n} \sum_{j=i+1}^{n} C\theta^{j-i} \|\phi_j\|_\alpha \|\phi_i\|_{L^1_{\mu_2}} \\
&\leq C \sum_{i=m}^{n} \|\phi_i\|_{L^1_{\mu_2}},
\end{aligned}$$

d'où le résultat. □

4.2.2 Lemmes dynamiques de Borel-Cantelli pour $\mathcal{T}_1$

Dans cette section et la suivante, on considère le problème de Borel-Cantelli pour $\mathcal{T}_1$ (ou plus précisément pour $\mathcal{T}_1^2$). Notre premier résultat concerne le cas d'une suite décroissante d'ensembles rétrécissants autour d'un point, et découle essentiellement d'un résultat de Kim [95]. Dans la section suivante, on considèrera des suites de fonctions d'un type un peu plus général.

Remarque 4.2.3. *D'après le théorème 6.1 de [70], si (U_n) est une suite de boules dans Δ_ϵ, $\epsilon \in \{0, 1\}$, vérifiant $\mu_1(U_n) \geq \frac{C}{n}$, alors (U_n) vérifie la propriété de Borel-Cantelli pour $\mathcal{T}_1^2$, c'est-à-dire que $\mathcal{T}_1^{2n} x \in U_n$ pour μ_2-p.t. $x \in \Delta_\epsilon$, puisque $(\mathcal{T}_1^2, \Delta_\epsilon, \mu_1)$ a une décroissance exponentielle des corrélations pour les observables Lipschitz, d'après Bufetov [32]. On s'intéresse ici à des résultats de nature quantitative, c'est-à-dire la propriété forte de Borel-Cantelli.*

On peut déduire de la proposition 4.2.2 une propriété forte de Borel-Cantelli pour $\mathcal{T}_1$ pour les suites décroissantes de fonctions à support dans B et bornées dans V_α. On utilisera pour cela le résultat suivant de Kim :

Théorème 4.2.4 ([95, Théorème 3.1])**.** *Soit (X, T, μ) un système dynamique ergodique, où μ est une mesure de probabilité T-invariante. Soit $E \subset X$ avec $\mu(E) > 0$, et $T_E : E \to E$ l'application de premier retour. Soit (f_n) une suite décroissante de fonctions mesurables positives à support dans E et telle que $\sum_n \int_X f_n \, d\mu = \infty$. Si toute sous-suite (f_{n_k}) avec $\sum_k \int_E f_{n_k} \, d\mu = \infty$ vérifie la propriété forte de Borel-Cantelli pour T_E, alors (f_n) vérifie la propriété forte de Borel-Cantelli pour T.*

Comme corollaire immédiat, on obtient :

Théorème 4.2.5. *Soit (f_n) une suite décroissante de fonctions mesurables positives, supportées dans B, et bornée dans V_α, avec $\sum_n \int_B f_n \, d\mu_1 = \infty$. Alors (f_n) vérifie la propriété forte de Borel-Cantelli pour $\mathcal{T}_1$.*

Démonstration. On rappelle que la mesure μ_2 invariante par $\mathcal{T}_2$ est la restriction de μ_1 à B, renormalisée. Comme toute sous-suite (f_{n_k}) avec $\sum_k \int_B f_{n_k} \, d\mu_2 = \infty$ est bornée dans V_α, d'après la proposition 4.2.2, une telle sous-suite vérifie la propriété forte de Borel-Cantelli pour $\mathcal{T}_2$. Une application du théorème de Kim conclut la démonstration. □

Corollaire 4.2.6. *Soit (A_n) une suite décroissante de boules rétrécissante autour d'un point p^* vérifiant la condition de Keane, avec $\sum_n \mu_1(A_n) = \infty$. Alors (A_n) vérifie la propriété forte de Borel-Cantelli pour $\mathcal{T}_1$.*

Démonstration. Comme p^* vérifie la condition de Keane, on peut trouver un bon domaine d'induction B qui contient p^*. Par conséquent, pour n assez grand, on a $A_n \subset B$. D'après la remarque 4.1.4, la suite des fonctions caractéristiques $\mathbb{1}_{A_n}$ est bornée dans V_α d'où le résultat. □

4.2.3 Lemmes de Borel-Cantelli pour des suites de fonctions quelconques pour $\mathcal{T}_1$

On considère à présent des suites de fonctions positives supportées dans B, et on démontre des lemmes dynamiques de Borel-Cantelli pour $(\mathcal{T}_1^2, \Delta_\epsilon, 2\mu_1)$. On pose $G = \mathcal{T}_1^2$ et $\tilde{\mu}_1 = 2\mu_1$ restreinte à Δ_ϵ pour simplifier les notations.
Soit $B \subset \Delta_\epsilon$ un bon domaine d'induction. On remarque que $\mathcal{T}_2$ est aussi l'application de premier retour de G dans B, mais avec un temps de retour $\tilde{n}_2 = \frac{n_2}{2}$. L'ingrédient clé est une estimation sur la décroissance des corrélations pour G pour des observables dans V_α à support dans B :

Théorème 4.2.7. *Il existe $0 < \kappa < 1$ et $C > 0$ telles que pour tout $\phi \in V_\alpha$ et tout $\psi \in L^1(\tilde{\mu}_1)$ à support dans B :*

$$\left| \int \phi\psi \circ G^n d\tilde{\mu}_1 - \left(\int \phi d\tilde{\mu}_1 \right) \left(\int \psi d\tilde{\mu}_1 \right) \right| \leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1(\tilde{\mu}_1)}.$$

On en déduit :

Corollaire 4.2.8. *Soit (ϕ_n) une suite de fonctions positives, à support dans B et telle que $\sup_n \|\phi_n\|_\alpha < \infty$. Si $\sum_n \int_B \phi_n \, d\tilde{\mu}_1 = \infty$, alors (ϕ_n) possède la propriété forte de Borel-Cantelli pour G.*

En particulier, toute suite de boules non nécessairement décroissante, mais "localisée", au sens où elle doit être contenue dans un bon domaine d'induction B, satisfait la propriété forte de Borel-Cantelli pour G.

Démonstration du corollaire 4.2.8. On utilise la proposition 1.4.1 avec $f_k = \phi_k \circ G^k$ et $\bar{f}_k = \varphi_k = \int_B \phi_k \, d\tilde{\mu}_1$. Un réarrangement des termes montre qu'il est suffisant de montrer

$$\sum_{i=m}^{n} \sum_{j=i+1}^{n} \left| \tilde{\mu}_1(\phi_j \circ G^{j-i}\phi_i) - \tilde{\mu}_1(\phi_j)\tilde{\mu}_1(\phi_i) \right| \leq C \sum_{i=m}^{n} \tilde{\mu}_1(\phi_i).$$

Mais $|\tilde{\mu}_1(\phi_j \circ G^{j-i}\phi_i) - \tilde{\mu}_1(\phi_j)\tilde{\mu}_1(\phi_i)| \leq C\kappa^{j-i}\|\phi_i\|_{L^1(\tilde{\mu}_1)}$ ce qui donne le résultat puisque $\sum_{j>i} \kappa^{j-i}$ est sommable en i. □

Afin de démontrer le théorème 4.2.7, on utilise la théorie du renouvellement pour les opérateurs, dans l'esprit de Sarig [128] et Gouëzel [61], bien que dans ce contexte où la queue des temps de retour décroit exponentiellement vite, la démonstration sera grandement simplifiée, voir par exemple la remarque 2.3.7 de [60].

Proposition 4.2.9. *Soit $\mathcal{B}$ un espace de Banach et (R_n) une suite d'opérateurs bornés sur $\mathcal{B}$. On suppose que*

1. *$\|R_n\| = \mathcal{O}(\theta^n)$ pour un certain $0 < \theta < 1$, et donc $R(z) = \sum_n R_n z^n$ et $R'(z) = \sum_n nR_n z^{n-1}$ sont bien définis pour z dans le disque unité fermé $\bar{\mathcal{D}}$,*
2. *1 est une valeur propre simple isolée de $R(1)$ et le projecteur propre Π associé vérifie $\Pi R'(1)\Pi = \gamma\Pi$ pour un certain $\gamma \neq 0$,*
3. *$I - R(z)$ est inversible pour tout $z \in \bar{\mathcal{D}} \setminus \{1\}$.*

Soit

$$V_n = \sum_{l=1}^{\infty} \sum_{k_1+\ldots+k_l=n} R_{k_l} \ldots R_{k_1}.$$

Alors V_n est un opérateur borné sur $\mathcal{B}$ et

$$\left\| V_n - \frac{1}{\gamma}\Pi \right\| = \mathcal{O}(\kappa^n)$$

pour un certain $0 < \kappa < 1$.

Démonstration. Comme $\|R_n\|$ décroit exponentiellement vite vers 0, la fonction $R(z) = \sum_n R_n z^n$ est bien définie et analytique sur un disque centré en 0 et de rayon $r > 1$. Comme $R(1)$ admet 1 comme valeur propre simple isolée, par la théorie des perturbations pour les opérateurs, pour tout z dans un petit voisinage de 1, $R(z)$ admet une valeur propre simple isolée dominante $\lambda(z)$, de projecteur associé $\Pi(z)$, et $\lambda(z)$ et $\Pi(z)$ dépendent analytiquement de z. On définit $S(z) = \frac{I-R(z)}{1-z}$ et on souhaite montrer que $S(z)^{-1}$ admet une extension analytique sur un disque de centre 0 et de rayon $r' > 1$. Comme $I - R(z)$ est inversible pour tout $z \in \bar{\mathcal{D}} \setminus \{1\}$, le seul problème est l'extension au voisinage de 1. Par hypothèse, $S(z)$ est inversible pour tout $z \in \bar{\mathcal{D}} \setminus \{1\}$, et on peut donc écrire pour z assez proche de 1

$$S(z)^{-1} = \frac{1-z}{1-\lambda(z)}\Pi(z) + (1-z)(I - R(z)Q(z))^{-1}Q(z),$$

avec $Q(z) = I - \Pi(z)$.
Puisque $\lambda(z)$ est analytique et non identiquement constante, il existe un disque $D(1,\epsilon)$ tel que $\lambda(z) \neq 1$ pour tout $z \in D(1,\epsilon) \setminus \{1\}$. Ainsi, le premier terme de la somme admet une extension analytique sur un disque épointé autour de 1. D'autre part, $I - R(1)Q(1)$ est inversible, donc $(I - R(z)Q(z))^{-1}$ est bien définie et dépend analytiquement de z dans $D(1,\epsilon)$ avec éventuellement un ϵ plus petit. Donc le second terme de la somme admet une extension analytique sur $D(1,\epsilon)$, et par conséquent $S(z)^{-1}$ se prolonge analytiquement sur $D(1,\epsilon) \setminus \{1\}$. D'après le théorème 2.2.5 de [60], $S(z)^{-1}$ peut être prolongée continument, et donc analytiquement, en 1, avec $S(1)^{-1} = \frac{1}{\gamma}\Pi$. Puisque $I - R(z)$ est inversible pour tout $z \in \bar{\mathcal{D}} \setminus \{1\}$ et que $\bar{\mathcal{D}}$ est compact, on en déduit que $S(z)^{-1}$ admet une extension analytique sur un disque centré en 0 de rayon $r' > 1$. Ainsi, $\frac{S(z)^{-1}-S(1)^{-1}}{1-z} = (I - R(z))^{-1} - \frac{\Pi}{\gamma}\frac{1}{1-z}$ est analytique sur le même disque. Par la formule intégrale de Cauchy, les coefficients du développement de Taylor de cette fonction en 0 décroissent exponentiellement vite, mais ces coefficients sont donnés par $V_n - \frac{\Pi}{\gamma}$, d'où le résultat. □

Soit P_1 l'opérateur de transfert de G par rapport à la mesure de Lebesgue m (ou plus précisément, la mesure $d\pi \times m$, bien qu'on fera toujours l'identification entre les deux), défini pour $\phi \in L^1(m)$ par

$$P_1\phi(x) = \sum_{Gy=x} \frac{\phi(y)}{\mathrm{Jac}(G)(y)}.$$

Soit $C_n = \{\tilde{n}_2 = n\}$ l'ensemble des points de B qui ont un temps de retour égal à n. C_n est constitué de l'union de plusieurs éléments $\{B_i\}_{i\in\mathcal{I}_n}$ de la partition $\mathcal{Q}$. On définit alors les opérateurs $R_n\phi = \mathbb{1}_B P_1^n(\mathbb{1}_{C_n}\phi) = \mathbb{1}_B P(\mathbb{1}_{C_n}\phi)$ et $V_n\phi = \mathbb{1}_B P_1^n(\mathbb{1}_B\phi)$, qui agissent sur $L^1_m(B)$. L'opérateur R_n correspond au premier retour dans B au temps n tandis que V_n correspond aux points qui retournent dans B au temps n, qu'il s'agisse ou non d'un premier retour.
On a l'équation du renouvellement suivante, voir par exemple la proposition 1 de [128]

$$V_n = \sum_{l=1}^{\infty} \sum_{k_1+\ldots+k_l=n} R_{k_l} \ldots R_{k_1}.$$

On va montrer que ces opérateurs vérifient les trois conditions requises pour appliquer la proposition 4.2.9.
Pour la première condition, on aura besoin du lemme technique suivant. On rappelle que pour $i \in \mathcal{I}_n$, $\mathcal{T}_2$, et donc G^n est une bijection de B_i sur B. Soit $I_i = I_{1,B_i} : B \to B_i$ la branche inverse de G^n restreinte à B_i.

Lemme 4.2.10. *La constante de Lipschitz $c_i = c_{1,B_i}$ de $I_i : B \to B_i$ décroit exponentiellement vite avec n : il existe $0 < \gamma < 1$ et $C > 0$ telles que $c_i \leq C\gamma^n$ pour tout $n \geq 1$ et tout $i \in \mathcal{I}_n$.*

Démonstration. D'après Avila et Bufetov [9], $m(C_n)$ décroit exponentiellement vite. L'application $I_i : B \to B_i$ est la composition d'une application linéaire $\lambda \mapsto A\lambda$ [1], suivie de $A\lambda \mapsto \frac{A\lambda}{|A\lambda|}$. A est une matrice positive, et $\frac{\lambda_k}{\lambda_l}$ est borné pour tout $\lambda = (\lambda_1, \ldots, \lambda_d)$ dans B. Ainsi $1 \geq \frac{|A\lambda|}{\|A\|} \geq C > 0$ pour tout $\lambda \in B$, d'après une observation de Avila-Bufetov [9] page 9. De plus, $\frac{|A\lambda'|}{|A\lambda|} \leq C$ pour tout $\lambda, \lambda' \in B$ d'après le lemme 4.1.2. Ainsi, la décroissance exponentielle du volume $m(B_i) \leq m(C_n)$ entraine qu'au moins une direction est contractée exponentiellement vite par I_i, et donc toutes le sont, d'où le lemme. □

Lemme 4.2.11. *Il existe $0 < \theta < 1$ et $C > 0$ telles que $\|R_n\| \leq C\theta^n$ pour tout n.*

Démonstration. On a $R_n\phi = \sum_{i\in\mathcal{I}_n} \frac{\phi(I_i x)}{\mathrm{Jac}(\mathcal{T}_2)(I_i x)}$, donc $R_n = \sum_{i\in\mathcal{I}_n} M_{1,B_i}$. Ainsi, par le lemme 4.1.6, on a

$$\|R_n\phi\|_{L^1_m} \leq \sum_{i\in\mathcal{I}_n} \|M_{1,B_i}\phi\|_{L^1_m} = \sum_{i\in\mathcal{I}_n} \int_{B_i} |\phi|\, dm = \int_{C_n} |\phi|\, dm \leq m(C_n)\|\phi\|_{L^\infty_m} \leq Cm(C_n)\|\phi\|_\alpha,$$

et

$$\begin{aligned}\int \mathrm{osc}(R_n\phi, B_\epsilon(x))dx &\leq \sum_{i\in\mathcal{I}_n} \int \mathrm{osc}(M_{1,B_i}\phi, B_\epsilon(x))dx \\ &\leq C\left(\sum_{i\in\mathcal{I}_n} \int_{B_i} \mathrm{osc}(\phi, B_{c_i\epsilon}(x))dx + \epsilon \sum_{i\in\mathcal{I}_n} \int_{B_i} |\phi|\, dm\right) \\ &\leq C\int_{C_n} \mathrm{osc}(\phi, B_{c^{(n)}\epsilon}(x))dx + C\epsilon \int_{C_n} |\phi|\, dm,\end{aligned}$$

où $c^{(n)} = \sup_{i\in\mathcal{I}_n} c_i$.
On a

$$\int_{C_n} \mathrm{osc}(\phi, B_{c^{(n)}\epsilon}(x))dx \leq \int_B \mathrm{osc}(\phi, B_{c^{(n)}\epsilon}(x))dx \leq (c^{(n)})^\alpha \epsilon^\alpha |\phi|_\alpha \leq (c^{(n)})^\alpha \epsilon^\alpha \|\phi\|_\alpha$$

et $\int_{C_n} |\phi|\, dm \leq m(C_n)\|\phi\|_{L^\infty_m} \leq Cm(C_n)\|\phi\|_\alpha$, d'où

$$|R_n\phi|_\alpha \leq C((c^{(n)})^\alpha + m(B_n))\|\phi\|_\alpha$$

1. Comme B est inclus dans un élément de la forme $\{\pi\} \times \Delta_{\pi,\epsilon}$, on identifie ses éléments $(\pi, \lambda) \in B$ avec λ.

et la même inégalité pour $\|R_n\phi\|_\alpha$. Comme $c^{(n)}$ et $m(B_n)$ décroissent exponentiellement vite par le lemme 4.2.10 et Avila-Bufetov [9] respectivement, on obtient bien que $\|R_n\| = \mathcal{O}(\theta^n)$ en tant qu'opérateur sur V_α, pour un certain $0 < \theta < 1$. □

Lemme 4.2.12. *$R(1)$ admet 1 pour valeur propre simple et isolée, et le projecteur correspondant est donné par*

$$\Pi\phi = \frac{1}{\tilde{\mu}_1(B)} \left(\int_B \phi \, dm \right) h_B,$$

où h_B est la restriction à B (non renormalisée) de la densité de probabilité de μ_1. De plus, $\Pi R'(1)\Pi = \frac{\Pi}{\tilde{\mu}_1(B)}$, de telle sorte que γ dans la proposition 4.2.9 est donné par $\gamma = \frac{1}{\tilde{\mu}_1(B)}$.

Démonstration. On remarque que $R(1) = P$ est l'opérateur de transfert de $\mathcal{T}_2$. Par conséquent, il admet bien 1 comme valeur propre simple isolée, et le projecteur correspondant est donné par la formule de l'énoncé du lemme. D'autre part,

$$\Pi R'(1)\Pi\phi = \left(\frac{\int_B R'(1)h_B \, dm}{\tilde{\mu}_1(B)} \right) \left(\frac{\int_B \phi \, dm}{\tilde{\mu}_1(B)} \right) h_B,$$

et donc $\gamma = \frac{\int_B R'(1)h_B \, dm}{\tilde{\mu}_1(B)}$. Mais

$$\int_B R'(1)h_B \, dm = \sum_n n \int_B P(\mathbb{1}_{C_n} h_B) \, dm = \sum_n n \int_{C_n} h_B \, dm = \sum_n n\tilde{\mu}_1(C_n) = \int_B \tilde{n}_2 \, d\tilde{\mu} = 1$$

par le lemme de Kac, d'où $\gamma = \frac{1}{\tilde{\mu}_1(B)}$. □

Il reste à démontrer la condition d'apériodicité :

Lemme 4.2.13. *Pour tout $z \in \bar{\mathcal{D}} \setminus \{1\}$, $I - R(z)$ est inversible.*

Démonstration. On démontre d'abord une inégalité de Lasota-Yorke pour $R(z)$. On remarque que

$$R(z)^k = \sum_{n_1,\ldots,n_k \geq 1} z^{n_1+\ldots+n_k} R_{n_k} \ldots R_{n_1},$$

et que

$$R_{n_k} \ldots R_{n_1} = \sum_{i_1 \in \mathcal{I}_{n_1},\ldots,i_k \in \mathcal{I}_{n_k}} M_{k,Q_{i_1,\ldots,i_k}},$$

où $Q_{i_1,\ldots,i_k} \in \mathcal{Q}_k$ est défini par $Q_{i_1,\ldots,i_k} = B_{i_1} \cap \mathcal{T}_2^{-1} B_{i_2} \cap \ldots \cap \mathcal{T}_2^{-(k-1)} B_{i_k}$.
Comme $|z| \leq 1$ et $n_1 + \ldots + n_k \geq k$, on obtient en sommant les relations fournies par le lemme 4.1.6 : $\|R(z)^k\phi\|_{L^1_m} \leq C|z|^k \|\phi\|_{L^1_m}$ et $|R(z)^k\phi|_\alpha \leq C|z|^k \left(\theta^{-\alpha k} |\phi|_\alpha + \|\phi\|_{L^1_m} \right)$, en raisonnant comme pour le lemme 4.1.5.
Ceci montre que le rayon spectral de $R(z)$ est plus petit que $|z|$, et que le rayon spectral essentiel de $R(z)$ est strictement plus petit que 1 lorsque $|z| = 1$, par le théorème de Hennion

[75]. Par conséquent, le problème se réduit à montrer que la relation $R(z)\phi = \phi$, avec $|z| = 1$ et $\phi \in V_\alpha$ entraine $z = 1$ ou $\phi = 0$.
Soit donc $\phi \in V_\alpha$ non nul vérifiant $R(z)\phi = \phi$, c'est-à-dire $P(z^{\tilde{n}_2}\phi) = \phi$, où on le rapelle, P est l'opérateur de transfert de $\mathcal{T}_2 = G^{\tilde{n}_2}$ par rapport à m, et $\tilde{n}_2$ est le premier temps de retour de G dans B. D'après la proposition 1.1 de Morita [117], on a $\left(\frac{\phi}{h_B}\right) \circ \mathcal{T}_2 = z^{\tilde{n}_2}\frac{\phi}{h_B}$. Comme $(G, \tilde{\mu}_1)$ est mélangeant, et donc faiblement mélangeant, on trouve $z = 1$ d'après la proposition 4.2.18, ce qui termine la démonstration. □

Démonstration du théorème 4.2.7. Grâce aux lemmes 4.2.11, 4.2.12 et 4.2.13, on est en mesure d'appliquer la proposition 4.2.9 et d'obtenir $\|V_n - \tilde{\mu}_1(B)\Pi\| \leq C\kappa^n$, c'est-à-dire

$$\left\| V_n\phi - \left(\int_B \phi\, dm\right) h_B \right\| \leq C\kappa^n \|\phi\|_\alpha$$

pour $0 < \kappa < 1$.
Soit $\phi \in V_\alpha$ et $\psi \in L^1(\tilde{\mu}_1)$ à support dans B. On a

$$\int \phi\psi \circ G^n\, dm = \int \mathbb{1}_B P_1^n(\mathbb{1}_B\phi)\psi\, dm = \int (V_n\phi)\psi\, dm,$$

d'où

$$\begin{aligned}
\left| \int V_n(\phi)\psi dm - \left(\int \phi\ dm\right) \int \psi\ d\tilde{\mu}_1 \right| &= \left| \int \left[V_n\phi - \left(\int \phi\ dm\right) h_B \right] \psi\ dm \right| \\
&\leq \left\| V_n\phi - \left(\int_B \phi\, dm\right) h_B \right\|_\alpha \int |\psi|\ dm \\
&\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_m}.
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
\left| \int \phi\psi \circ G^n dm - \left(\int \phi\ dm\right) \int \psi\ d\tilde{\mu}_1 \right| &\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_m} \\
&\leq C\kappa^n \|\phi\|_\alpha \|\psi\|_{L^1_{\tilde{\mu}_1}},
\end{aligned}$$

puisque $\|\psi\|_{L^1_m} \leq \|h_B^{-1}\|_{L^\infty_m} \|\psi\|_{L^1_{\tilde{\mu}_1}} \leq C\|\psi\|_{L^1_{\tilde{\mu}_1}}$, la densité h_B de $\tilde{\mu}_1$ étant bornée inférieurement sur B. Le théorème en découle, en prenant ϕh_B pour ϕ, et en utilisant l'inégalité $\|\phi h_B\|_\alpha \leq \|\phi\|_\alpha \|h_B\|_\alpha \leq C\|\phi\|_\alpha$. □

4.2.4 Appendice : apériodicité et mélange faible

Soit (X, T, μ) un système dynamique où μ est une mesure de probabilité T-invariante.

Définition 4.2.14. *(X, T, μ) est dit faiblement mélangeant si $f \circ T = e^{it} f$ pour $f \in L^2(\mu)$ non nul et $t \in [0, 2\pi[$ entraine $t = 0$ et f constante.*

Remarque 4.2.15. *Cette définition est équivalente à la définition classique, qui exige que*

$$\frac{1}{n}\sum_{k=0}^{n-1}\Big|\mu(T^{-k}(A)\cap B)-\mu(A)\mu(B)\Big|\to 0$$

pour tous sous-ensemble mesurables A et B, voir [142, Théorème 1.26] lorsque (X,T,μ) est inversible, et [86, Théorème 664] ou [52, Théorème 2.36] pour une preuve de l'équivalence valable dans tous les cas.

Soit $Y\subset X$ un sous-ensemble mesurable de mesure positive $\mu(Y)>0$. On note $\tau(y)$ le premier temps de retour de $y\in Y$ dans Y :

$$\tau(y)=\min\{n\geq 1\,:\,T^n y\in Y\}.$$

On définit l'application de premier retour $\hat{T}:Y\to Y$ par $\hat{T}=T^\tau$. Elle préserve la mesure μ_Y qui est la restriction de μ à Y, renormalisée de telle sorte que $\mu_Y(Y)=1$, et est ergodique.

Définition 4.2.16. *On dira que le premier temps de retour τ est apériodique si $f\circ\hat{T}=e^{it\tau}f$ pour $f\in L^2(\mu_Y)$ non nul et $t\in[0,2\pi[$ entraine que $t=0$ et f est constante.*

Remarque 4.2.17. *Par la proposition 1.1 de Morita [117], la relation $f\circ\hat{T}=e^{it\tau}f$ est équivalente à $\mathcal{L}(e^{it\tau}f)=f$, où $\mathcal{L}$ est l'opérateur de transfert de $\hat{T}$ par rapport à la mesure μ_Y.*

Proposition 4.2.18. *Le premier temps de retour τ est apériodique si et seulement si (X,T,μ) est faiblement mélangeant.*

Démonstration. On suppose d'abord que le premier temps de retour est apériodique. Soit $f\in L^2(\mu)$ non nul et $t\in[0,2\pi[$ tels que $f\circ T=e^{it}f$. On vérifie aisément que la restriction f_Y de f à Y satisfait $f_Y\circ\hat{T}=e^{it\tau}f_Y$:

$$f_Y(\hat{T}y)=f(T^{\tau(y)}y)=e^{it\tau(y)}f(y)=e^{it\tau(y)}f_Y(y).$$

f_Y est aussi non identiquement nulle : autrement, f s'annulerait sur l'ensemble $\cup_{n\geq 0}T^{-n}Y$ qui par ergodicité est égal à X modulo μ. Par apériodicité, on obtient bien que $t=0$, ce qui entraine que $f\circ T=f$. L'ergodicité montre que f est constante.
Réciproquement, supposons que (X,T,μ) est faiblement mélangeant, et soit $f\in L^2(\mu_Y)$ non nul tel que $f\circ\hat{T}=e^{it\tau}f$ pour $t\in[0,2\pi[$. On prolonge d'abord τ sur l'espace entier comme étant le premier temps d'entrée : pour $x\in X$, $\tau(x)=\inf\{n\geq 1\,:\,T^n x\in Y\}$, qui est bien défini par ergodicité. On définit ensuite $\tilde{f}\in L^2(\mu)$ par $\tilde{f}=e^{-it\tau}f\circ T^\tau$. Comme $T^{\tau(x)}x$ appartient à Y pour μ-p.t. $x\in X$ par définition, $\tilde{f}$ est bien définie. L'hypothèse faite sur f entraine que f et $\tilde{f}$ coïncident sur Y, et donc $\tilde{f}$ n'est pas identiquement nulle.
On vérifie maintenant que $\tilde{f}\circ T=e^{it}\tilde{f}$. Soit $x\in X$ avec $\tau(x)>1$. Comme τ est le premier temps de retour, on a $\tau(Tx)=\tau(x)-1$. Ainsi, $\tilde{f}(Tx)=e^{-it\tau(Tx)}f(T^{\tau(Tx)}Tx)=e^{it}e^{-it\tau(x)}f(T^{\tau(x)}x)=e^{it}\tilde{f}(x)$. Si $\tau(x)=1$, c'est-à-dire $Tx\in Y$, alors par définition de $\tilde{f}$, on a $\tilde{f}(x)=e^{-it}f(Tx)=e^{-it}\tilde{f}(Tx)$.
Le mélange faible entraine que $t=0$ et $\tilde{f}$ est constante. Comme f est la restriction de $\tilde{f}$ à Y, ceci montre que f est constante, et termine la démonstration. □

4.3 Statistiques de récurrence pour l'induction de Rauzy-Veech-Zorich

Ayant exprimé $\mathcal{T}_2$ comme une application multidimensionnelle uniformément dilatante et ayant une décroissance exponentielle des corrélations pour les observables Quasi-Hölder contre L^1, on est en mesure d'appliquer les résultats sur les statistiques de récurrence à $\mathcal{T}_2$ et $\mathcal{T}_1$.

4.3.1 Théorie des valeurs extrêmes

Soit (X_n) un processus stochastique stationnaire. On note F la fonction de répartition de X_0, c'est-à-dire $F(x) = \mathbb{P}(X_0 \leq x)$. Pour toute fonction de répartition F, on posera $\bar{F} = 1 - F$ et $u_F = \sup\{x \,:\, F(x) < 1\}$.
On définit une nouvelle suite de variables aléatoires $M_1, M_2, \ldots$ par

$$M_n = \max\{X_0, \ldots, X_{n-1}\}.$$

Définition 4.3.1. *On dira qu'on a une loi des valeurs extrêmes pour (M_n) si il existe une fonction de répartition $H : \mathbb{R} \to [0,1]$ non-dégénérée avec $H(0) = 0$, et pour tout $\tau > 0$, il existe une suite $u_n = u_n(\tau)$ telle que*

$$n\mathbb{P}(X_0 > u_n) \to \tau, \;\; lorsque \;\; n \to \infty,$$

telles que

$$\mathbb{P}(M_n \leq u_n) \to \bar{H}(\tau), \;\; lorsque \;\; n \to \infty,$$

en tout point de continuité de H.

Le cas le plus favorable correspond au cas d'une distribution limite exponentielle avec $\bar{H}(\tau) = e^{-\theta\tau}$ pour tout $\tau > 0$, où $0 \leq \theta \leq 1$ est appelé l'indice extrémal.
Pour un exposé de la théorie dans le cadre des variables i.i.d., on pourra consulter par exemple [101] ou [126].
Dans le cas où (X, T, μ) est un système dynamique ergodique sur un espace métrique (X, d), avec μ une mesure de probabilité T-invariante, on considèrera des processus de la forme

$$X_n = \phi \circ T^n,$$

où $\phi : X \to \mathbb{R} \cup \{+\infty\}$ est un observable, ayant un maximal global strict en un point $p_0 \in X$ et suffisamment régulier pour que pour tout u assez grand, l'ensemble $\{x \in X \,:\, \phi(x) > u\}$ corresponde à une boule topologique centrée en p_0. Pour des applications récentes de la théorie des valeurs extrêmes aux systèmes dynamiques, on renvoie par exemple à [43, 55, 80], ou au survol [54].
On considère tout d'abord $(B, \mathcal{T}_2, \mu_2)$. Soit $\phi : B \to \mathbb{R} \cup \{+\infty\}$ qui satisfait les conditions décrites auparavant, et maximisé strictement au point $p_0 \in B$. On supposera aussi que

pour tout u assez grand, les fonctions caractéristiques $\mathbb{1}_{\phi>u}$ appartiennent à V_α et y forment une famille bornée, ce qui est par exemple le cas si ϕ est C^2 au voisinage de p_0 avec une différentielle seconde qui y est non-dégénérée. On rappelle que $u_n = u_n(\tau) = u_n^{\mathcal{T}_2}(\tau)$ est choisi de telle sorte que $n\mu_2(\phi > u_n) \to \tau$. Une application directe de la proposition 3.3 de Aytaç, Freitas et Vaienti [12] donne :

Proposition 4.3.2. *Soit $p_0 \in B$.*

1. *Si p_0 n'est pas un point périodique de $\mathcal{T}_2$, alors $\mu_2(M_n \leq u_n^{\mathcal{T}_2}(\tau)) \to e^{-\tau}$ pour tout $\tau > 0$.*
2. *Si p_0 est un point périodique répulsif de $\mathcal{T}_2$ de période première k, alors $\mu_2(M_n \leq u_n^{\mathcal{T}_2}(\tau)) \to e^{-\theta\tau}$ pour tout $\tau > 0$, avec $\theta = 1 - |\mathrm{Jac}(\mathcal{T}_2^{-k})(p_0)|$.*

On peut aussi considérer les processus ponctuels associés à des extrêmes multiples. Pour $A \subset \mathbb{R}$ et $u > 0$, on définit

$$\mathcal{N}_u(A) = \sum_{i \in A \cap \mathbb{N}} \mathbb{1}_{X_i > u}.$$

On observe que pour $A = [0, n[$, $\mathcal{N}_u(A)$ compte le nombre d'occurrences de l'événement $\{X_i > u\}$ avant le temps n. De plus, $\{\mathcal{N}_u(A) = 0\} = \{M_n \leq u\}$.
Convenablement renormalisé, on peut montrer que ce processus ponctuel converge vers un processus de Poisson composé. On introduit d'abord les définitions, pour lesquelles on renvoie à [87, 126] pour plus de détails. Soit $\mathfrak{R}$ l'algèbre des parties de $\mathbb{R}^+$ engendrée par la semi-algèbre des intervalles de la forme $[a, b[$, de telle manière qu'un élément $J \in \mathfrak{R}$ est de la forme $J = \cup_{i=1}^n [a_i, b_i[$. Pour $\gamma > 0$, on pose $\gamma J = \cup_{i=1}^n [\gamma a_i, \gamma b_i[$.

Définition 4.3.3. *Soit $T_1, T_2, \ldots$ une suite de variables aléatoires i.i.d. suivant une loi exponentielle de paramètre $\frac{1}{\theta}$. Soit $D_1, D_2, \ldots$ une suite de variables i.i.d. et indépendante de $T_1, \ldots$, de fonction de répartition π. Pour $J \in \mathfrak{R}$, on pose*

$$N(J) = \int \mathbb{1}_J \, d\left(\sum_{i=1}^{\infty} D_i \delta_{T_1 + \ldots + T_i}\right),$$

où δ_t désigne la mesure de Dirac en $t > 0$. On dira alors que N est un processus de Poisson composé d'intensité θ et de multiplicité π.

Remarque 4.3.4.

1. Par la suite, les variables D_i seront toujours à valeurs entières, ce qui signifie que la multiplicité π est entièrement déterminée par les valeurs $\pi_k = \mathbb{P}(D_1 = k)$ pour $k \in \mathbb{N}$. Si $\theta = 1$ et $\pi_1 = 1$, alors N est le processus de Poisson standard, et la variable aléatoire $N([0, t[)$ suit une loi de Poisson d'espérance t.
2. Dans les applications qui vont suivre, π suivra toujours une loi géométrique de paramètre $\theta \in]0, 1]$: $\pi_k = \theta(1-\theta)^k$ pour tout $k \geq 0$. Dans ce cas, la variable aléatoire $N([0, t[)$ suit une loi de Pólya-Aeppli :

$$\mathbb{P}(N([0, t[) = k) = e^{-\theta t} \sum_{i=1}^{k} \theta^i (1-\theta)^{k-i} \frac{(\theta t)^i}{i!} \begin{pmatrix} k-1 \\ i-1 \end{pmatrix}.$$

On pose $v_n = v_n^{\mathcal{T}_2} = \mu_2(\phi > u_n^{\mathcal{T}_2}) \sim \frac{n}{\tau}$, on définit le processus ponctuel $N_n^{\mathcal{T}_2}$ pour $J \in \mathfrak{R}$ par

$$N_n^{\mathcal{T}_2}(J) = \mathcal{N}_{u_n}(v_n J) = \sum_{i \in v_n J \cap \mathbb{N}} \mathbb{1}_{\phi \circ \mathcal{T}_2^i > u_n}.$$

Par la proposition 3.3 de [12], on a :

Proposition 4.3.5. *Soit $p_0 \in B$.*

1. *Si p_0 n'est pas un point périodique de $\mathcal{T}_2$, alors le processus ponctuel $N_n^{\mathcal{T}_2}$ converge en loi vers un processus de Poisson standard d'intensité 1.*
2. *Si p_0 est un point périodique répulsif de $\mathcal{T}_2$ de période première k, alors le processus ponctuel $N_n^{\mathcal{T}_2}$ converge en loi vers un processus de Poisson composé d'intensité θ et de multiplicité une loi géométrique de paramètre θ, avec $\theta = 1 - |\mathrm{Jac}(\mathcal{T}_2^{-k})(p_0)|$.*

Comme $\mathcal{T}_2$ est l'application de premier retour de $\mathcal{T}_1$ dans B, on peut transposer les résultats obtenus pour $\mathcal{T}_2$ en résultats pour $\mathcal{T}_1$, grâce à [73, 57]. La deuxième partie sur les points périodiques est aussi traitée dans [56]. En effet, dès que p_0 satisfait la condition de Keane, on peut trouver un bon domaine d'induction B qui contient p_0.
Soit $\phi : \mathcal{R} \times \Delta_{\mathcal{A}} \to \mathbb{R} \cup \{+\infty\}$, maximisé strictement en p_0. On définit $u_n = u_n^{\mathcal{T}_1}(\tau)$ de telle sorte que $n\mu_1(\phi > u_n) \to \tau$, et $v_n = v_n^{\mathcal{T}_1}(\tau) = \mu_1(\phi > u_n^{\mathcal{T}_1}(\tau))^{-1}$. On définit le processus ponctuel $N_n^{\mathcal{T}_1}$ de façon analogue à $N_n^{\mathcal{T}_2}$.

Proposition 4.3.6. *On suppose que p_0 satisfait la condition de Keane.*

1. *Si p_0 n'est pas un point périodique de $\mathcal{T}_1$, alors $\mu_1(M_n \leq u_n^{\mathcal{T}_1}(\tau)) \to e^{-\tau}$ pour tout $\tau > 0$ et le processus ponctuel $N_n^{\mathcal{T}_1}$ converge en loi vers un processus de Poisson standard d'intensité 1.*
2. *Si p_0 est un point périodique répulsif de $\mathcal{T}_1$ de période première k, alors $\mu_1(M_n \leq u_n^{\mathcal{T}_1}(\tau)) \to e^{-\theta\tau}$ pour tout $\tau > 0$, avec $\theta = 1 - |\mathrm{Jac}(\mathcal{T}_1^{-k})(p_0)|$, et le processus ponctuel $N_n^{\mathcal{T}_1}$ converge en loi vers un processus de Poisson composé d'intensité θ et de multiplicité une loi géométrique de paramètre θ,*

4.3.2 Temps d'entrée et de retour

On présente dans cette section des résultats sur une notion de récurrence qui est d'un certain point de vue équivalente aux lois des valeurs extrêmes [55].
Soit (X, T, μ) un système dynamique ergodique, avec μ une mesure de probabilité. Pour un sous-ensemble $U \subset X$ avec $\mu(U) > 0$, on définit le temps d'entrée dans U par

$$\tau_U(x) = \min\{k \geq 1 \,:\, T^n x \in U\},$$

pour μ-p.t. $x \in X$.

Définition 4.3.7. *Soit (U_n) une suite décroissante d'ensembles telle que $\mu(U_n) \to 0$.*

1. *On dira que (U_n) a une statistique des temps d'entrée décrite par la fonction de répartition H si pour tout $t \geq 0$,*

$$\mu\left(\left\{x \in X \,:\, \tau_{U_n}(x) \leq \frac{t}{\mu(U_n)}\right\}\right) \to H(t).$$

2. *On dira que (U_n) a une statistique des temps de retour décrite par la fonction de répartition $\tilde{H}$ si pour tout $t \geq 0$,*

$$\frac{1}{\mu(U_n)}\mu\left(\left\{x \in U_n \,:\, \tau_{U_n}(x) \leq \frac{t}{\mu(U_n)}\right\}\right) \to \tilde{H}(t).$$

Il est démontré dans [69] que (U_n) admet une statistique des temps de retour si et seulement si elle admet une statistique des temps d'entrée, et les distributions H et $\tilde{H}$ sont liées par une relation intégrale dont l'unique point fixe est la distribution exponentielle $H(t) = 1 - e^{-t}$. On renvoie au survol [130] et aux référence contenues dans ce papier pour plus d'informations sur cette notion de récurrence quantitative.
La situation est parfaitement comprise pour $\mathcal{T}_2$ et $\mathcal{T}_1$ dans le cas où U_n est une suite de boules décroissante.

Proposition 4.3.8. *Soit U_n une suite décroissante de boules centrées en p_0. On suppose que p_0 satisfait la condition de Keane lorsque l'on considère $\mathcal{T}_1$.*

1. *Si p_0 n'est pas périodique pour $\mathcal{T}_2$ (resp. $\mathcal{T}_1$), alors (U_n) admet une statistique des temps d'entrée et de retour exponentielle, c'est-à-dire $H(t) = \tilde{H}(t) = 1 - e^{-t}$.*
2. *Si p_0 est périodique pour $\mathcal{T}_2$ (resp. $\mathcal{T}_1$), de période première k, alors (U_n) admet une statistique des temps d'entrée et de retour données respectivement par $H(t) = 1 - e^{-\theta t}$ et $\tilde{H}(t) = (1-\theta)+\theta(1-e^{-\theta t})$, où $\theta = 1 - |\mathrm{Jac}(\mathcal{T}_2^{-k})(p_0)|$ (resp. $\theta = 1 - |\mathrm{Jac}(\mathcal{T}_1^{-k})(p_0)|$).*

Le résultat pour $\mathcal{T}_2$ se déduit de [56] et [73]. Comme $\mathcal{T}_2$ est l'application de premier retour de $\mathcal{T}_1$ dans un ensemble B qui contient p_0, les mêmes lois limites restent vraies pour $\mathcal{T}_1$. Ceci a été démontré d'abord pour des points génériques dans [31, Théorème 2.1], pour les points périodiques dans [56, Théorème 5], et pour n'importe quel point dans [73].

4.4 Remarques et questions ouvertes

1. Beaucoup de résultats de ce chapitre peuvent sans doute être étendus à des systèmes multidimensionnels plus généraux. Par exemple, les résultats concernant $\mathcal{T}_2$ peuvent surement être généralisés à des systèmes uniformément dilatants et surjectifs par morceaux, tandis que les résultats concernant $\mathcal{T}_1$ pourraient s'étendre à des systèmes admettant un système induit du type précédent.

2. Peut-on démontrer des lemmes de Borel-Cantelli pour $\mathcal{T}_1$ lorsque la suite de boules rétrécit autour d'un point p qui ne satisfait pas la condition de Keane, ou se situe sur le bord du simplexe ? La méthode de ce chapitre est ici inapplicable car il n'est pas a priori possible de trouver un bon ensemble d'induction B qui contienne p.
3. Est-il possible d'étendre les résultats de ce chapitre aux versions inversibles de l'induction de Rauzy-Veech, comme introduites au paragraphe 2.7 de Viana [141] ?

Bibliographie

[1] J. Aaronson, M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn., **1**, 193-237, (2001)

[2] J. Aaronson, M. Denker, O. Sarig, R. Zweimuller, *Aperiodicity of cocycles and conditional local limit theorems*, Stoch. Dyn., **4**, 31-62, (2004)

[3] R. Aimino, M. Nicol, S. Vaienti, *Annealed and quenched limit theorems for random expanding dynamical systems*, paru en ligne sur Probability Theory and Related Fields, (2014), `http://dx.doi.org/10.1007/s00440-014-0571-y`

[4] R. Aimino, M. Nicol, M. Todd, *Recurrence statistics for the space of Interval Exchange maps and the Teichmüller flow on the space of translation surfaces*, preprint, (2014)

[5] R. Aimino, H. Hu, M. Nicol, A. Török, S. Vaienti, *Polynomial loss of memory for maps of the interval with a neutral fixed point*, à paraître sur Discrete and Continuous Dynamical Systems -A, **35**, (2015)

[6] R. Aimino, J. Rousseau, *Concentration inequalities for sequential dynamical systems of the unit interval*, preprint, (2014)

[7] J.F. Alves, J.M. Freitas, S. Luzzato, S. Vaienti, *From rates of mixing to recurrence times via large deviations*, Adv. in Maths., **228**, 2, 1203-1236, (2011)

[8] L. Arnold, *Random dynamical systems*, Berlin : Springer, (1998)

[9] A. Avila, A.Bufetov. *Exponential decay of correlations for the Rauzy-Veech-Zorich induction map*, Fields Institute Communications, **57**, (2007)

[10] A. Ayyer, M. Stenlund, *Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms*, Chaos, **17**, (2007)

[11] A. Ayyer, C. Liverani, M. Stenlund, *Quenched CLT for random toral automorphism*, Discrete and Continuous Dynamical Systems, **24**, 331-348, (2009)

[12] H. Aytaç, J.M. Freitas, S. Vaienti, *Laws of rare events for deterministic and random dynamical systems*, à paraitre sur Trans. Amer. Math. Soc., (2014)

[13] K. Azuma, *Weighted sums of certain dependent random variables*, Tôhoku Math. J., **19**, 357-367, (1967)

[14] V.I. Bakhtin, *Random processes generated by a hyperbolic sequence of mappings, I*, Rus. Ac. Sci. Izv. Math., **44**, 247-279, (1995)

[15] V.I. Bakhtin, *Random processes generated by a hyperbolic sequence of mappings, II*, Rus. Ac. Sci. Izv. Math., **44**, 617-627, (1995)

[16] V. Baladi, L.-S. Young, *On the spectra of randomly perturbed expanding maps*, Commun. Math. Phys., **156**, 355-385, (1993)

[17] V. Baladi, A. Kondah, B. Schmitt, *Random correlations for small perturbations of expanding maps*, Random and Computational Dynamics, **4**, 179-204, (1996)

[18] V. Baladi, *Correlation spectrum of quenched and annealed equilibrium states for random expanding maps*, Commun. Math. Phys., **186**, 671-700, (1997)

[19] V. Baladi, *Positive transfer operators and decay of correlations*, Vol. 16, World Scientific, (2000)

[20] V. Baladi, S. Gouëzel, *Good Banach spaces for piecewise hyperbolic maps via interpolation*, Ann. IHP Analyse non linéaire, **26**, 1453-1481, (2009)

[21] D. Berend, V. Bergelson, *Ergodic and mixing sequences of transformations*, Ergodic theory dynam. systems, **4**, 353-366, (1984)

[22] I. Berkes, *Results and problems related to the pointwise central limit theorem*, in *Asymptotic methods in probability and statistics (Ottawa, ON, 1997)*, ed B. Szyszkowicz (Amsterdam : North-Holland), 59-96, (1998)

[23] P. Billingsley, *The Lindeberg-Lévy theorem for martingales*, Proc. of the AMS, **12**, 788-792, (1961)

[24] M. Blank, G. Keller, C. Liverani, *Ruelle-Perron-Frobenius spectrum for Anosov maps*, Nonlinearity, **15**, 1905-1973, (2001)

[25] S. Boucheron, G. Lugosi, P. Massart, *Concentration inequalities : a nonasympototic theory of independence*, Oxford University Press, (2013)

[26] A. Boyarsky, P. Göra, *Absolutely continuous invariant measures for piecewise expanding C^2 transformations in R^N*, Israel J. Math., **67**, 272-286, (1989)

[27] A. Boyarsky, P. Göra, *Laws of chaos. Invariant measures and dynamical systems in one dimension*, Probability and its applications, Birkhauser, (1997)

[28] L. Breiman, *Probability*, Addison-Wesley, Reading, Mass, (1968)

[29] A. Broise, *Etudes spectrales d'opérateurs de transfert et applications*, Astérisque, **238**, (1996)

[30] B.M. Brown, *Martingale central limit theorems*, Annals of Math. Statistics, **42**, 59-66, (1971)

[31] H. Bruin, B. Saussol, S. Troubetzkoy, S. Vaienti, *Return time statistics via inducing*, Ergodic Theory Dynam. Systems **23**, 991-1013, (2003)

[32] A. Bufetov. *Decay of correlations for the Rauzy-Veech-Zorich induction map on the space of interval exchange transformations and the central limit theorem for the Teichmüller flow on the moduli space of abelian differentials*, J. Amer. Math. Soc. **19**, 579-562, (2006)

[33] D.L. Burkholder, *Distribution function inequalities for martingales*, Ann. Probability, **1**, 19-42, (1973)

[34] J. Buzzi, *Exponential decay of correlations for random Lasota-Yorke maps*, Comm. Math. Phys., **208**, 25-54, (1999)

[35] J.-R. Chazottes, P. Collet, F. Redig, E. Verbitskiy, *A concentration inequality for interval maps with an indifferent fixed point*, Ergodic theory dynam. systems, **29**, 1097-1117, (2009)

[36] J.-R. Chazottes, P. Collet, B. Schmitt, *Devroye inequality for a class of non-uniformly hyperbolic dynamical systems*, Nonlinearity, **18**, 2323-2340, (2005)

[37] J.-R. Chazottes, P. Collet, B. Schmitt, *Statistical consequences of the Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems*, Nonlinearity, **18**, 2341-2364, (2005)

[38] J.-R. Chazottes, S. Gouëzel, *On almost-sure versions of classical limit theorems for dynamical systems*, Probab. Theory Relat. Fields, **138**, 195-234, (2007)

[39] J.-R. Chazottes, S. Gouëzel, *Optimal concentration inequalities for dynamical systems*, Commun. Math. Phys., **316**, 843-889, (2012)

[40] G. Chen, Y. Shi, *Chaos of time-varying discrete dynamical systems*, Journal of Difference Equations and applications, **15**, 429-449, (2009)

[41] N. Chernov, D. Kleinbock, *Dynamical Borel-Cantelli lemmas for Gibbs measures*, Israel J. Math., **122**, 1-27, (2001)

[42] D.L. Cohn, *Measure Theory : Second edition*, Birkhaüser Advanced Texts, Birkhaüser, (2013)

[43] P. Collet. *Statistics of closest return for some non-uniformly hyperbolic systems*, Ergodic Theory Dynam. Systems, **21**, 401-420, (2001)

[44] P. Collet, S. Martinez, B. Schmitt, *Exponential inequalities for dynamical measures of expanding maps of the interval*, Probab. Theory Relat. Fields, **123**, 301-322, (2002)

[45] J.-P. Conze, A. Raugi, *Convergence of iterates of a transfer operator, application to dynamical systems and to Markov chains*, ESAIM Probability and Statistics, **7**, 115-146, (2003)

[46] J.-P. Conze, A. Raugi, *Limit theorems for sequential expanding dynamical systems on* $[0, 1]$, Ergodic theory and related fields, **89121**, Contemp. Math., 430, Amer. Math. Soc., Providence, RI, 2007

[47] G. Cristadoro, N. Haydn, P. Marie, S. Vaienti, *Statistical properties of intermittent maps with unbounded derivative*, Nonlinearity, **23**, 1071-1096, (2010)

[48] A. Dembo, O. Zeitouni, *Large Deviations, Techniques and Applications*, Applications of Mathematics **38**, Springer-Verlag, 2nd Edition, 1998

[49] W. de Melo, S. van Strien, *One-dimensional Dynamics*, Springer, Berlin, 1993

[50] D. Dolgopyat. *Limit theorems for partially hyperbolic systems*, Trans. of the AMS, **356**, 1637-1689, (2004)

[51] R. Durrett, *Probability : theory and examples*, third ed., Duxbury Press, Belmont, CA, (1996)

[52] M. Einsiedler, T. Ward, *Ergodic theory with a view towards number theory*, Graduate Texts in Mathematics, **259**, Springer, (2011)

[53] R.S. Ellis, *Entropy, Large Deviations and Statistical Mechanics*, Springer-Verlag, New-York, 1985

[54] J.M. Freitas, *Extremal behaviour of chaotic dynamics*, Dynamical Systems : an international journal, **28**, 302-332, (2013)

[55] A.C.M. Freitas, J.M. Freitas, M. Todd, *Hitting Time Statistics and Extreme Value Theory*, Probab. Theory Related Fields, **147**, 675-710, (2010)

[56] A.C.M. Freitas, J.M. Freitas, M. Todd, *The compound Poisson limit ruling periodic extreme behaviour of non-uniformly hyperbolic dynamics*, Comm. Math. Phys., **321**, 483-527, (2013)

[57] A.C.M. Freitas, J.M. Freitas, and M. Todd, *Dichotomies in extreme value laws for dynamical systems*, Manuscript

[58] A. Furman, Y. Shalom, *Sharp ergodic theorems for group actions and strong ergodicity*, Ergodic Theory and Dynamical Systems, **19**, 1037-1061, (1999)

[59] M.I. Gordin, *The central limit theorem for stationary processes*, Soviet. Math. Dokl., **10**, 1174-1176, (1969)

[60] S. Gouëzel, *Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes*, Thèse, École Norme Supérieure, (2004)

[61] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math., **139**, 29-65, (2004)

[62] S. Gouëzel, C. Liverani, *Banach spaces adapted to Anosov systems*, Ergodic Theory and Dynamical Systems, **26**, 189-217, (2006)

[63] S. Gouëzel, *A Borel-Cantelli lemma for intermittent interval maps*, Nonlinearity, **20**, 1491-1497, (2007)

[64] S. Gouëzel, *Almost sure invariance principle for dynamical systems by spectral methods*, Annals of Probability, **38**, 1639-167, (2010)

[65] S. Gouëzel, I. Melbourne, *Moment bounds and concentration inequalities for slowly mixing dynamical systems*, preprint, (2014)

[66] Y. Guivarc'h, J. Hardy, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*, Annales de l'I.H.P., section B, **24**, 73-98, (1988)

[67] C. Gupta, M. Nicol, W. Ott, *A Borel–Cantelli lemma for non-uniformly expanding dynamical systems*, Nonlinearity, **23**, 1991-2008, (2010)

[68] C. Gupta, W. Ott, A. Török, *Memory loss for time-dependent piecewise-expanding systems in higher dimension*, Math. Res. Lett., **20**, 141-161, (2013)

[69] N. Haydn, Y. Lacroix, S. Vaienti, *Hitting and return time statistics in ergodic dynamical systems*, Ann. Probab., **33**, 2043-2050, (2005)

[70] N. Haydn, M. Nicol, T. Persson, S. Vaienti, *A note on Borel-Cantelli lemmas for non-uniformly hyperbolic dynamical systems,* Ergodic Theory Dynam. Systems, **33**, 475-498, (2013)

[71] N. Haydn, M. Nicol, S. Vaienti, L. Zhang, *Central limit theorems for the shrinking target problem*, J. Stat. Phys., **153**, 864-887,(2013)

[72] N. Haydn, M. Nicol, A. Török, S. Vaienti, *Almost sure invariance principle for sequential and non-stationary dynamical systems*, preprint, (2014)

[73] N. Haydn, N. Winterberg, R. Zweimuller, *Return-time statistics, Hitting-time statistics and Inducing,* Preprint (`http://mat.univie.ac.at/~zweimueller/MyPub/WiZ1.pdf`), à paraitre dans Ergodic Theory, Open Dynamics, and Coherent Structures, Springer Proceedings in Mathematics & Statistics.

[74] L. Heinrich, *Mixing properties and central limit theorems for a class of non-identical piecewise monotonic C^2-transformations*, Math. Nachr., **181**, 185-214, (1996)

[75] H. Hennion, *Sur un théorème spectral et son application aux noyaux Lipschitziens*, Proceedings of the A.M.S., **118**, 627-634, (1993)

[76] H. Hennion, L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasicompactness*, Lect. Notes in Math., **1766**, (2001), Springer-Verlag

[77] H. Hennion, L.Hervé, *Central limit theorems for iterated random Lipschitz mappings*, Ann. Probab., **32**, 1934-1984, (2004)

[78] W. Hoeffding, *Probability inequality for sums of bounded random variables*, J. Amer. Statist. Assoc., **58**, 13-30, (1963)

[79] F. Hofbauer, G. Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Mathematische Zeitschrift, **180**, 119-140, (1982)

[80] M.P. Holland, M. Nicol, A. Török, *Extreme value distributions for non-uniformly hyperbolic dynamical systems*, Trans. Amer. Math. Soc., **364**, 661-688, (2012)

[81] L.-Y. S. Hsieh, *Ergodic theory of multidimensional random dynamical systems*, Ph.D. Thesis, University of Victoria, (2008)

[82] H. Hu, *Decay of correlations for piecewise smooth maps with indifferent fixed points*, Ergodic Theory and Dynamical Systems, **24**, 495-524, (2004)

[83] C.T. Ionescu-Tulcea, G. Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Ann. Math., **52**, 140-147, (1950)

[84] H. Ishitani, *Central limit theorems for the random iterations of 1-dimensional transformations (dynamics of complex systems)*, RIMS Kokyuroku Univ., **1404**, 21-31, (2004)

[85] F. Jonsson, *Almost sure central limit theory*, Uppsala University : Department of Mathematics, Project report, (2007)

[86] S. Kalikow, R. McCutcheon, *An outline of ergodic theory*, Cambridge studies in advanced mathematics, **122**, Cambridge University Press, (2010)

[87] O. Kallenberg, *Random measures, 4th edition*, Akademie-Verlag, Berlin, (1986)

[88] T. Kato, *Perturbation theory for linear operators*, Springer-Verlag, (1967)

[89] C. Kawan, *Metric entropy of nonautonomous dynamical systems*, Nonautonomous dynamical systems, **1**, 1-71, (2014)

[90] G. Keller, *Generalized bounded variation and applications to piecewise monotonic transformations*, Z. Wahr. verw. Geb., **69**, 461-478, (1985)

[91] Y. Kifer, *Ergodic theory for random transformations*, Boston : Birkhauser, (1986)

[92] Y. Kifer, *Random perturbations of dynamical systems*, Boston : Birkhauser, (1988)

[93] Y. Kifer, *Limit theorems for random transformations and processes in random environments*, Trans. Amer. Math. Soc., **350**, 1481-1518, (1998)

[94] Y. Kifer, *Thermodynamic formalism for random transformations revisited*, Stoch. Dyn., **08**, 77-102, (2008)

[95] D. Kim, *The dynamical Borel-Cantelli lemma for interval maps*, Discrete Contin. Dyn. Syst., **17**, 891-900, (2007)

[96] D. Kleinbock, G. Margulis, *Logarithm laws for flows on homogeneous spaces,* Inv. Math., **138**, 451-494, (1999)

[97] E. Kobre, L.-S. Young, *Extended systems with deterministic local dynamics and random jumps*, Commun. Math. Phys., **275**, 709-720, (2007)

[98] S. Kolyada, L. Snoha, *Topological entropy of nonautonomous dynamical systems*, Random Comput. Dynam., **4**, 205-223, (1996)

[99] A. Lasota, J.-A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc., **186**, 481-488, (1973)

[100] M. Ledoux, *The concentration of measure phenomenon*, Mathematical surveys and monographs, **89**, AMS, (2005)

[101] G. Lindgren, M.R. Leadbetter, H. Rootzén, *Extremes and related properties of random sequences and processes*, Springer Series in Statistics, Springer-Verlag, New York-Berlin, (1983)

[102] C. Liverani, *Decay of correlations for piecewise expanding maps*, Journal of Stat. Phys., **78**, 1111-1129, (1995)

[103] C. Liverani, *Central limit theorem for deterministic systems*, Pitman Research Notes in Mathematics Series, **362**, 56-75, (1996)

[104] C. Liverani, B. Saussol, S. Vaienti, *A probabilistic approach to intermittency*, Ergodic theory and dynamical systems, **19**, 671-685, (1999)

[105] C. Liverani, *Multidimensional expanding maps with singularities : a pedestrian approach*, Ergodic Theory and Dynamical Systems, **33**, 168-182, (2013)

[106] C. Maldonado, *Fluctuation bounds for chaos plus noise in dynamical systems*, Journal of Stat. Phys., **148**, 548-564, (2012)

[107] H. Masur, *Interval exchange transformations and measured foliations*, Ann. of Math., **115**, 169-200, (1982)

[108] V. Maume-Deschamps, *Correlation decay for Markov maps on a countable state space*, Ergodic Theory and Dynamical Systems, **21**, 165-196, (2001)

[109] V. Maume-Deschamps, *Projective metrics and mixing properties on towers*, Trans. Amer. Math. Soc., **353**, 3371-3389, (2001)

[110] C. McDiarmid, *On the method of bounded differences*, in *Surveys in Combinatorics*, London Math. Soc. Lecture Note Ser., **141**, Cambridge Univ. Press., 148-188, (1989)

[111] C. McDiarmid, *Concentration*, in *Probabilistic methods for algorithmic discrete mathematics*, Algorithms Combin., **16**. Springer, Berlin, 195-248, (1998)

[112] I. Melbourne, M. Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Probab., **37**, 478-505, (2009)

[113] I. Melbourne, A. Török, *Convergence of moments for Axiom A and nonuniformly hyperbolic flows*, Ergodic theory and dynamical systems, **32**, 1091-1100, (2012)

[114] A. Mohapatra, W. Ott, *Memory loss for nonequilibrium open dynamical systems*, à paraitre sur Disc. Cont. Dyn. Sys., (2014)

[115] T. Morita, *Random iteration of one-dimensional transformations*, Osaka J. Math., **22**, 489-518, (1985)

[116] T. Morita, *Deterministic version lemmas in ergodic theory of random dynamical systems*, Hiroshima Math. J., **18**, 15-29, (1988)

[117] T. Morita, *A generalized local limit theorem for Lasota-Yorke transformations*, Osaka J. Math., **26**, 579-595, (1989)

[118] T. Morita. *Renormalized Rauzy inductions*, Advanced studies in pure mathematics, **43**, 1-25, (2005)

[119] P. Nándori, D. Szász, T. Varjú, *A central limit theorem for time-dependent dynamical systems*, Journal of Stat. Phys., **146**, 1213-1220, (2012)

[120] J. Neveu, *Bases mathématiques du calcul des probabilités, deuxième édition*, Masson et Cie, (1970)

[121] W. Ott, M. Stenlund, L.-S. Young, *Memory loss for time-dependent dynamical systems*, Math. Res. Lett., **16**, 463-475, (2009)

[122] S. Pelikan, *Invariant densities for random maps of the interval*, Trans. Amer. Math. Soc., **281**, 813-825, (1984)

[123] W. Philipp, *Some metrical theorems in number theory*, Pacific J. Math., **20**, 109-127, (1967)

[124] M. Pollicott, *Statistical properties of the Rauzy-Veech-Zorich map*, Notes non publiées, (2009) (`http://homepages.warwick.ac.uk/~masdbl/teichmuller-asip.pdf`)

[125] A. Raugi, *Étude d'une transformation non uniformément hyperbolique de l'intervalle [0,1[*, Bull. Soc. Math. France, **132**, 81-103, (2004)

[126] S. Resnick, *Extreme values, regular variation, and point processes*, Springer Science & Business Media, Springer-Verlag, (2007)

[127] J. Rousseau-Egele, *Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux*, Annals of Probability, **11**, 772-788, (1983)

[128] O. Sarig. *Subexponential decay of correlations*, Invent. Math., **150**, 629-653, (2002)

[129] B. Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps*, Israel J. Math., **116**, 223-248, (2000)

[130] B. Saussol, *An introduction to quantitative Poincaré recurrence in dynamical systems*, Reviews in math. physics, **21**, 949-979, (2009)

[131] H. Schaefer, *Banach lattices and positive operators*, Springer-Verlag, **215**, (1974)

[132] V.G. Sprindzuk, *Metric theory of Diophantine approximations*, V. H. Winston and Sons, Washington, D.C., 1979, Translated from the Russian and edited by Richard A. Silverman, With a foreword by Donald J. Newman, Scripta Series in Mathematics. MR MR548467 (80k :10048).

[133] M. Stenlund, *Non-stationary compositions of Anosov diffeomorphisms*, Nonlinearity, **24**, 2991-3018, (2011)

[134] M. Stenlund, L.-S. Young, H. Zhang, *Dispersing billiards with moving scatterers*, Commun. Math. Phys., **322**, 909-955, (2013)

[135] M. Stenlund, H. Sulku, *A coupling approach to random circle maps expanding on the average*, à paraître sur Stochastics and Dynamics, (2014)

[136] M. Thaler, *Estimates of the invariant densities of endomorphisms with indifferent fixed points*, Israel J. Math., **37**, 303-314, (1980)

[137] D. Thomine, *A spectral gap for transfer operators of piecewise expanding maps*, Discrete Contin. Dyn. Syst., **30**, 917-944, (2011)

[138] M. Tsujii, *Absolutely continuous invariant measures for expanding piecewise linear maps*, Inventiones Mathematicae, **143**, 349-373, (2001)

[139] F. Tümel, *Random walks on a lattice with deterministic local dynamics*, Ph.D thesis, University of Houston, (2012)

[140] W. A. Veech, *Gauss measures for transformations on the space of interval exchange maps*, Ann. of Math., **115**, 201-242, (1982)

[141] M. Viana, *Dynamics of interval exchange maps and Teichmüller flows*, IMPA 2008 (`http://w3.impa.br/~viana/out/ietf.pdf`).

[142] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, **79**, Springer, (1982)

[143] L.-S. Young, *Statistical properties of dynamical systems with some hyperbolicity*, **147**, 585–650, (1998)

[144] L.-S. Young, *Recurrence times and rates of mixing*, Israel J. Math., **110**, 153-188, (1999)

[145] A. Zorich, *Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents*, Ann. Inst. Fourier (Grenoble), **46**, 325-370, (1996)

Vitesse de mélange et théorèmes limites pour les systèmes dynamiques aléatoires et non-autonomes

Résumé

Dans cette thèse, nous nous intéressons aux propriétés statistiques des systèmes dynamiques aléatoires et non-autonomes. Dans le premier chapitre, consacré aux systèmes aléatoires, nous établissons un cadre fonctionnel abstrait, couvrant une large classe de systèmes dilatants en dimension 1 et supérieure, permettant de démontrer de nombreux théorèmes limites annealed. Nous donnons aussi une condition nécessaire et suffisante pour que la version quenched du théorème de la limite centrale soit valide en dimension 1. Dans le chapitre deux, après avoir introduit la notion de système non-autonome, nous étudions un système composé d'applications en dimension 1 ayant un point fixe neutre commun, et nous montrons que celui-ci admet une vitesse de perte de mémoire polynomiale. Le chapitre trois est consacré aux inégalités de concentration. Nous établissons de telles inégalités pour des systèmes dynamiques aléatoires et non-autonomes, et nous étudions diverses applications. Dans le chapitre quatre, nous nous intéressons aux lemmes dynamiques de Borel-Cantelli pour l'induction de Rauzy-Veech-Zorich, et présentons quelques résultats liés aux statistiques de récurrence pour cette application.

Rates of mixing and limit theorems for random and non-autonomous dynamical systems

Abstract

In this thesis, we study statistical properties of random and non-autonomous dynamical systems. In the first chapter, devoted to random systems, we establish an abstract functional framework, including a large class of expanding systems in dimension 1 and higher, under which we can prove annealed limit theorems. We also give a necessary and sufficient condition for the quenched central limit theorem to hold in dimension 1. In chapter 2, after an introduction to the notion of non-autonomous system, we study an example consisting of a family of maps of the unit interval with a common neutral fixed point, and we show that this system admits a polynomial loss of memory. The chapter 3 is devoted to concentration inequalities. We establish such inequalities for random and non-autonomous dynamical systems in dimension 1, and we study some of their applications. In chapter 4, we study dynamical Borel-Cantelli lemmas for the Rauzy-Veech-Zorich induction, and we present some results concerning statistics of recurrence for this map.