ANNÉE 2013

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Mathématiques

Ecole doctorale Matisse

présentée par

Damien THOMINE

préparée à l'IRMAR
Institut de Recherche Mathématique de Rennes
UFR de Mathématiques - Université de Rennes 1

Théorèmes limites pour les sommes de Birkhoff de fonctions d'intégrale nulle en théorie ergodique en mesure infinie

Thèse soutenue à Rennes
le 10 décembre 2013

devant le jury composé de :

Mme Viviane Baladi
Directeur de recherche CNRS / examinateur

M. Jean-Pierre Conze
Professeur, Université de Rennes 1 / examinateur

M. Sébastien Gouëzel
Chargé de recherche CNRS / directeur de thèse

M. Marc Peigné
Professeur, Université de Tours / rapporteur

Mme Françoise Pène
Maître de conférence, UBO / examinateur

M. Sandro Vaienti
Professeur, Université du Sud - Toulon / examinateur

suivant un rapport de :

M. Ian Melbourne
Professeur, Université de Warwick / rapporteur

Table des matières

Introduction

Cette thèse porte sur certains aspects de théorie ergodique en mesure infinie. Les objets d'étude de la théorie ergodique sont des systèmes dynamiques préservant la mesure, c'est-à-dire des triplets (Ω, μ, T) formés d'un espace polonais, d'une mesure μ positive et σ-finie, et d'une transformation mesurable $T : \Omega \to \Omega$. Ici, on supposera que T préserve la mesure μ, et l'on s'intéresse en particulier au comportement des processus $(f \circ T^n)_{n \geq 0}$, où $f : \Omega \to \mathbb{R}$ est une observable. Des problèmes similaires se posent pour des systèmes dynamiques en temps continu.

La majeure partie de la discipline se concentre sur des systèmes finis, pour lesquels μ est une mesure de probabilité.

Théorie ergodique en mesure finie

Rappelons quelques résultats de théorie des probabilités. Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes, identiquement distribuées et à valeurs réelles. Si X_0 est intégrable, alors le comportement moyen de cette suite est capturé par la loi forte des grands nombres :

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} X_k = \mathbb{E}(X_0) \text{ presque sûrement.}$$

Cependant, si l'espérance de X_0 est nulle, ce théorème annonce que $(X_n)_{n \geq 0}$ converge en moyenne de Césarò vers 0. Cela revient à dire que l'on ne regarde pas la suite $(X_n)_{n \geq 0}$ à la bonne échelle, et que l'on ne voit pas les comportement intéressants. Le second résultat fondamental qui vient à notre aide est le théorème central limite. Si X_0 est d'espérance nulle et de carré intégrable, alors :

$$\lim_{n \to +\infty} \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} X_k = \|X_0\|_{\mathbb{L}^2} \mathcal{N} \text{ en loi,}$$

où $\mathcal{N}$ suit une loi normale, centrée en 0 et de variance 1. Quitte à changer de renormalisation et d'objet limite, il existe des résultats similaires quand X_0 n'est pas de moyenne nulle, mais est dans le bassin d'attraction d'une loi stable de Lévy.

Revenons aux systèmes dynamiques mesurés. Le comportement moyen du processus $(f \circ T^n)_{n \geq 0}$ est décrit par un théorème fondamental de la théorie ergodique, dû à G.D. Birkhoff [12] :

Théorème 0.0.1 (Théorème ergodique de Birkhoff)**.**
Soit (Ω, μ, T) un système dynamique muni d'une mesure de probabilité invariante et ergodique. Soit $f \in \mathbb{L}^1(X, \mu)$. Alors :

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = \int_\Omega f \, \mathrm{d}\mu \ \mu \textit{ presque sûrement.}$$

Sous ces hypothèses, les moyennes temporelle et spatiale des observables sont presque sûrement égales. On peut voir ce résultat comme une loi forte des grands nombres pour la suite $(f \circ T^n)_{n \geq 0}$ pour toute fonction intégrable f ; en fait, on peut déduire la loi forte des grands nombres en appliquant ce résultat à des systèmes dynamiques bien choisis.

Si f est d'intégrale nulle, les choses se compliquent. En théorie des probabilités, on peut obtenir un théorème central limite pour une suite $(X_n)_{n \geq 0}$ sous des hypothèses bien plus générales qu'être une suite de variables indépendantes et identiquement distribuées. Par exemple, on peut obtenir un théorème central limite pour des martingales [44], ou pour des processus ayant une propriété de mélange forte (ϕ-mélange, α-mélange) ou faible (β-mélange), pourvu que ce mélange soit suffisamment rapide [40, théorèmes 1.5, 1.6 et 1.7]. De façon générale, un théorème central limite demande néanmoins une certaine forme de décorrélation. La situation est la même en théorie ergodique : il faut que le processus $(f \circ T^n)_{n \geq 0}$ vérifie une certaine forme de décroissance des corrélations. Pour cela, le système dynamique sous-jacent (Ω, μ, T) doit être chaotique, et la fonction f doit être suffisamment régulière.

Ceci dit, la situation est assez bien comprise, et on sait montrer pour de nombreux systèmes chaotiques et des classes d'observables suffisamment larges que $n^{-1/2} \sum_{k=0}^{n-1} f \circ T^k$ converge en loi vers une gaussienne, dont on peut même calculer la variance. C'est le cas, par exemple, si (Ω, μ, T) est une application uniformément dilatante d'un intervalle et f est à variation bornée. Il existe plusieurs méthodes pour démontrer un théorème central limite dans ce contexte. Par exemple, on peut utiliser des méthodes de martingales, et montrer que la suite $(f \circ T^n)_{n \geq 0}$ est, à l'ajout d'un cobord près, une suite inversée de différences de martingales pour une filtration bien choisie. On dispose aussi de critères spectraux sur l'opérateur de transfert qui assurent l'existence d'un théorème central limite. Les résultats connus ne se limitent pas au théorème central limite, mais incluent aussi des principes des grandes déviations, des principes d'invariance presque sûre, des résultats similaires pour des observables de variance infinie, etc.

Théorie ergodique en mesure infinie

Les systèmes ergodiques en mesure finie sont récurrents positifs. La situation est cependant très différente si la mesure considérée est σ-finie, mais infinie : dans ce cas, on est face à un phénomène de type récurrence nulle. Si on se donne un ensemble mesurable A tel que $\mu(A) > 0$, les itérés $T^n x$ vont revenir (presque sûrement) un nombre infini de fois dans l'ensemble A, mais le temps de retour moyen en A est infini.

Le problème se rapproche de l'étude du temps local en un point pour des marches aléatoires récurrentes : ce sont bien des systèmes récurrents nuls, et le temps local en 0 est la somme de Birkhoff de la fonction valant 1 en 0, et 0 ailleurs. Une adaptation naïve du théorème de Birkhoff ne fonctionne pas :

Proposition 0.0.2.

Soit (X, μ, T) un système dynamique conservatif muni d'une mesure infinie, invariante et ergodique. Soit f une fonction de $\mathbb{L}^1(X, \mu)$. Alors :

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k = 0 \ \mu \text{ presque sûrement.} \tag{0.0.1}$$

Les sommes des observations successives croissent donc à une vitesse sous-linéaire. On peut chercher une suite $(a_n)_{n \geq 0}$ telle que $a_n^{-1} \sum_{k=0}^{n-1} f \circ T^k$ converge en un certain sens. Cependant, d'après un résultat de J. Aaronson [1], même avec des hypothèses très faibles sur la suite $(a_n)_{n \geq 0}$, on ne peut pas avoir de bonne convergence presque partout. Pour aller plus loin, il va falloir faire des hypothèses supplémentaires sur les systèmes étudiés, qui devront être en un certain sens chaotiques. On va alors découvrir un phénomène un peu plus subtil, où la convergence se fait d'une certaine façon en loi. Par exemple, pour la marche aléatoire simple sur $\mathbb{Z}$:

Proposition 0.0.3.

Soit $(S_n)_{n \geq 0}$ une marche aléatoire simple sur $\mathbb{Z}$, dont la loi du point de départ est $\nu \in \mathbb{P}(\mathbb{Z})$. Soit $f \in \ell^1(\mathbb{Z})$. Alors :

$$\lim_{n \to +\infty} \sqrt{\frac{\pi}{2n}} \sum_{k=0}^{n-1} f(S_k) = \sum_{k \in \mathbb{Z}} f(k) \cdot |\mathcal{N}| \text{ en loi,} \tag{0.0.2}$$

où $\mathcal{N}$ suit une loi normale centrée en 0 et de variance $\pi/2$.

Ce type de proposition existe pour des classes assez larges d'applications en mesure infinie, qui incluent des marches aléatoires sur $\mathbb{Z}$ ou $\mathbb{Z}^2$ plus générales, le flot géodésique sur des surfaces hyperboliques périodiques, des billards périodiques, ou dans un autre registre des applications hyperboliques avec un point fixe neutre, comme les applications de Pomeau-Manneville. Les renormalisations dépendent du système, de même que la loi limite, qui dans tous ces exemples est une loi de Mittag-Leffler. Afin d'étudier ces systèmes munis d'une mesure infinie, on induit sur un sous-système de mesure finie (par exemple, pour une marche aléatoire, on induit en l'origine). Ces propriétés de convergence en loi se démontrent par le biais d'une théorie du renouvellement, dont les versions les plus récentes sont formulées en termes d'opérateurs de transfert.

La situation en mesure infinie est donc relativement bien comprise, bien que plus parcellaire qu'en mesure finie ; il existe par exemple des systèmes dynamiques non hyperboliques, qui vérifient des propriétés de convergence en loi similaires, mais vers des lois qui ne sont pas de Mittag-Leffler.

De même qu'en mesure finie, ces résultats sont précis tant que l'intégrale de l'observable est non nulle. Si l'observable est d'intégrale nulle, on va devoir trouver une nouvelle renormalisation, et un nouvel objet limite. Il n'y a malheureusement presque aucun résultat précis sur ce problème. Citons cependant un résultat dû à R.L. Dobrushin [24, théorème 2] :

Théorème 0.0.4 (Dobrushin, 1955).
Soit ν une mesure de probabilité sur $\mathbb{Z}$. Soit $(S_n)_{n\geq 0}$ la marche aléatoire simple sur $\mathbb{Z}$ telle que la loi de S_0 soit ν. Soit f une fonction définie sur $\mathbb{Z}$, à valeurs réelles, à support fini, et de somme nulle. Alors :

$$\lim_{n\to+\infty} \frac{1}{n^{\frac{1}{4}}} \sum_{k=0}^{n-1} f(S_k) = d(f) \left(\frac{2}{\pi}\right)^{\frac{1}{4}} \sqrt{|\mathcal{N}'|}\mathcal{N} \text{ en loi,}$$

où $\mathcal{N}$ sont $\mathcal{N}'$ indépendantes, où $\mathcal{N}$ suit un loi normale standard, où $\mathcal{N}'$ suit une loi normale centrée de variance $\pi/2$ (de telle sorte que $\mathbb{E}(|\mathcal{N}'|) = 1$), et où :

$$d(f)^2 = 4\sum_{k\in\mathbb{Z}} kf^2(k) + 8 \sum_{\substack{(k,\ell)\in\mathbb{Z}^2 \\ k<\ell}} kf(k)f(\ell) - \sum_{k\in\mathbb{Z}} f^2(k).$$

Le cas des marches aléatoires est en fait bien compris, notamment grâce à des travaux récents de E. Csáki et A. Földes [22] [23], qui sont néanmoins moins explicites.

L'objectif de cette thèse est d'étendre ces résultats à différentes classes de systèmes dynamiques ergodiques en mesure infinie, et à des ensembles de fonctions suffisamment larges sur ces systèmes dynamiques. On montrera par exemple, vers la fin de cette thèse, une proposition que l'on peut formuler ainsi :

Proposition 0.0.5.
Soit N un $\mathbb{Z}$-revêtement connexe d'une variété riemannienne compacte et de courbure sectionnelle strictement négative. Soit μ_N la mesure de Liouville sur T^1N.

Soit f une fonction définie sur T^1N, à valeurs réelles, à support compact et höldérienne. Supposons que $\int_{T^1N} f \, \mathrm{d}\mu_N = 0$. Alors il existe un réel positif $K(f)$ tel que, pour toute mesure de probabilité $\nu \ll \mu_N$,

$$\lim_{t\to+\infty} \frac{1}{t^{\frac{1}{4}}} \int_0^t f \circ g_s(x,v) \, \mathrm{d}s = K(f)\sqrt{|\mathcal{N}'|}\mathcal{N} \text{ en loi dans } (T^1N, \mu_N),$$

où $\mathcal{N}$ sont $\mathcal{N}'$ indépendantes, où $\mathcal{N}$ suit un loi normale standard, et où $\mathcal{N}'$ suit une loi normale centrée de variance $\pi/2$ (de telle sorte que $\mathbb{E}(|\mathcal{N}'|) = 1$).

De plus, $K(f) = 0$ si et seulement si f est un cobord.

Délimitation du problème

Comme son titre l'indique, le sujet de la thèse est l'étude des propriétés asymptotiques de sommes de Birkhoff d'observables d'intégrale nulle définies sur des systèmes dynamiques munis d'une mesure infinie. On se donne donc un système dynamique (Ω, μ, T) qui soit ergodique, conservatif, et qui préserve la mesure, ainsi qu'une fonction f définie sur Ω d'intégrale nulle, et on cherche à connaître le comportement limite des sommes partielles $\sum_{k=0}^{n-1} f \circ T^k$. Nous allons faire trois importantes restrictions au sujet de cette étude.

Premièrement, les systèmes dynamiques étudiés sont, en un certain sens, *hyperboliques*. Plus précisément, on supposera qu'en induisant sur un sous-ensemble de mesure finie (ou sur une section bien choisie) A, on peut se ramener à un système dynamique (A, μ_A, T_A) qui vérifie une condition très contraignante : celle d'être *Gibbs-Markov*. En particulier, le système induit est chaotique. Cette hypothèse très forte, ou certaines propriétés proches, est vérifiée par une variété de systèmes intéressants, tels que les applications de Pomeau-Manneville, le flot géodésique sur certaines surfaces, le modèle du gaz de Lorentz... Ces systèmes sont les mieux compris en théorie ergodique en mesure infinie. Cette restriction exclut cependant d'autres classes de systèmes. Citons en particulier les travaux récents réalisés par O. Sarig sur des cocycles au-dessus de systèmes d'entropie nulle : étude du flot horocyclique sur des surfaces périodique avec F. Ledrappier [50] [51] [52], limite distributionnelle pour des cocycles au-dessus d'odomètres avec J. Aaronson [5] (pour lesquels la limite en distribution du temps local est notoirement différente du cas hyperbolique), récurrence de cocycles au-dessus d'une rotation avec A. Avila, D. Dolgopyat et E. Duriev [7]...

Deuxièmement, la plupart de nos résultats porteront sur une certaine forme de convergence en loi (la *convergence forte en loi*) des sommes de Birkhoff. Nous établirons au passage d'autres types de propriétés limites, telles que des bornes sur la taille des sommes de Birkhoff valables presque partout, mais ce n'est pas l'objet central de cette thèse.

Troisièmement, les observables dont on étudiera les sommes de Birkhoff seront d'intégrale nulle au sens le plus fort. Les observables f étudiées seront non seulement dans $\mathbb{L}^1$, mais de plus satisferont des conditions d'intégrabilité plus contraignantes (telles que l'intégrabilité $\mathbb{L}^p$ de l'observable induite par $|f|$ sur un sous-système). Ces observables devront aussi être régulières. En pratique, les classes de fonctions les plus explicites que nous parviendrons à étudier sont des fonctions höldériennes à support compact. Ceci exclut notamment des fonctions qui seraient de moyenne nulle en un sens plus faible. Citons par exemple, pour une marche aléatoire sur $\mathbb{Z}$, la fonction signe sur $\mathbb{Z}$ (d'intégrale nulle au sens où $\sum_{k=-n}^{n} f(k) = 0$ pour tout n), pour laquelle le comportement asymptotique des sommes de Birkhoff est donnée par la loi de l'arcsinus [59, théorème 9.11]. On pourra penser aussi à des travaux récents sur les marches aléatoires en paysage aléatoire, pour lesquelles la suite $(f(n))_{n\in\mathbb{Z}}$ peut être une suite de variables indépendantes et identiquement distribuées de moyenne nulle (d'intégrale nulle au sens où $\mathbb{E}(f(n)) = 0$ pour tout n, et où $\sum_{k=-n}^{n} f(k) = o(n^{1/2+\varepsilon})$ pour tout $\varepsilon > 0$) [18] [19] [20]. La renormalisation que l'on doit appliquer aux sommes de Birkhoff dépend fortement du cas considéré : pour la fonction signe, on renormalise par la suite $a(n) = n$; pour les marches aléatoires en environnement aléatoire, la suite de renormalisation est typiquement comprise entre $n^{1/2}$ et n. La loi limite dépend elle aussi du type d'observable considéré. On peut imaginer d'autre types d'observables, pour lesquelles les suites de renormalisation, les lois limites et les méthodes utilisées diffèrent, d'où l'intérêt de cette restriction.

Composition de la thèse

Le contenu de cette thèse est, pour l'essentiel, compris dans trois articles [69] [70] [71]. La thèse est divisée en trois chapitres, mais le découpage choisi est différent du découpage correspondant à ces trois articles.

Le chapitre 1 rassemble l'ensemble du contexte mathématique et des outils nécessaires par la suite. Il sert d'introduction condensée à la théorie des applications Gibbs-Markov, aux méthodes d'induction et temps de retour, aux fonctions à variation régulière, et à la théorie ergodique en mesure infinie (en particulier appliquée à des systèmes hyperboliques). Un certain nombre d'exemples sont aussi présentés. Une grande partie de ce matériel est commune aux trois articles sus-cités.

Le chapitre 2 traite des systèmes non inversibles en temps discret. Il regroupe l'ensemble de l'article [69], ainsi que la première partie de l'article [70], qui offre un certain nombre de généralisations des résultats principaux de [69] telles que l'étude des observables de faible régularité.

Le chapitre 3 traite des systèmes en temps continu : flots de suspension au-dessus d'un système Gibbs-Markov, $\mathbb{Z}^d$ extensions de tels flots, et extensions naturelles de celles-ci. Il se finit sur quelques résultats sur le flot géodésique sur des variétés hyperboliques périodiques. L'essentiel du contenu de ce chapitre vient de la deuxième partie de l'article [70], excepté la partie 3.2, qui contient l'ensemble du troisième article [71]. Notons que ce troisième article, et par là même la partie 3.2, contient un certain nombre de résultats plus généraux, et ne porte pas que sur les systèmes induisant une application Gibbs-Markov. Il est aussi de nature assez différente : les méthodes employées dans

[71] sont uniquement des méthodes spectrales, alors que le reste de la thèse contient surtout des méthodes de nature plus probabiliste.

Chapitre 1

Outils employés

Sommaire

Ce premier chapitre regroupe des notions et outils qui sont communs aux trois articles publiés, et seront utilisés dans le reste de la thèse. Nous rappelons d'abord quelques concepts de théorie ergodique, avant d'aborder des sujets plus spécialisés. Nos travaux se concentrent sur les systèmes dynamiques qui induisent une application Gibbs-Markov, et dont le temps de premier retour a des queues à variation régulière.

La première partie présente les applications Gibbs-Markov, ainsi qu'un certain nombre de résultats utiles à leur propos. Les techniques de perturbation de l'opérateur de transfert ne sont cependant pas abordées : elle ne seront utilisées que dans la partie 3.2, et les techniques relatives seront énoncées à ce moment-là.

La seconde partie traite des techniques d'induction et de construction de tour. Elle permet de définir ce qu'est un système dynamique induisant une application Gibbs-Markov. Les systèmes dynamiques en temps continu sont aussi discutés, *via* la notion de (semi-)flot de suspension. Nous en profitons pour présenter la technique de saut de Schweiger, qui sera utile pour le traitement de systèmes non mélangeants, et pour les marches aléatoires.

La troisième partie porte sur les fonctions à variation régulière, et regroupe certaines de leurs propriétés les plus importantes. Nous utiliserons des propositions élémentaires, telles que le théorème de convergence uniforme ou le théorème de Karamata, mais aussi des théorèmes taubériens. Quelques applications des fonctions à variation régulière en théorie des probabilités, en particulier les lois stables de Lévy, sont aussi présentées.

La quatrième partie exploite les notions définies dans les trois premières parties pour exposer quelques résultats de théorie ergodique en mesure infinie. Une attention particulière est portée aux théorèmes donnant une limite distributionnelle pour les sommes de Birkhoff d'observables intégrables.

Finalement, la cinquième partie expose quelques exemples, qui motivent la direction que nous avons choisie pour cette thèse. Ces exemples sont classés en deux catégories, les applications avec un point fixe neutre (ou systèmes similaires), et les marches aléatoires (ou systèmes similaires).

Quelques rappels de théorie ergodique générale

Avant d'entrer dans le vif du sujet, nous allons rappeler quelques définitions fondamentales de théorie ergodique (voir par exemple [35, chapitre 4]). Celle-ci est, en bref, l'étude des systèmes dynamiques mesurés. On distinguera deux situations, les systèmes en temps discret et les systèmes en temps continu. Le premier cas correspond à des actions du semi-groupe $\mathbb{N}$ (pour les systèmes non-inversibles) ou du groupe $\mathbb{Z}$ (pour les systèmes inversibles). Le second cas correspond à des actions du semi-groupe $\mathbb{R}_+$ ou du groupe $\mathbb{R}$. Pour simplifier, nous allons commencer par présenter diverses définitions dans le cadre des systèmes en temps discret.

Un *système dynamique mesuré* en temps discret est la donnée d'un triplet (Ω, μ, T), où :

- Ω est un espace polonais, c'est-à-dire un espace topologique métrisable tel qu'il existe une distance le rendant complet ;
- μ est une mesure σ-finie, positive et non nulle sur Ω ;
- T est une application mesurable de Ω dans lui-même, définie partout ou éventuellement μ-presque partout, et non singulière.

On notera $T_*\mu$ la *mesure image* de μ par T, c'est-à-dire la mesure définie sur Ω par $T_*\mu(B) = \mu(T^{-1}B)$ pour tout borélien B. Un système *préserve la mesure* si $T_*\mu = \mu$. Un sous-ensemble borélien est dit *errant* si les ensembles $(T^{-n}B)_{n\geq 0}$ sont deux à deux disjoints, et un système est dit *conservatif* si tout ensemble errant est de mesure nulle. Un système est *ergodique* si tout ensemble borélien B tel que $T^{-1}B = B$ modulo μ vérifie $\mu(B) = 0$ ou $\mu(\Omega - B) = 0$, autrement dit si tout ensemble invariant par T est trivial pour la mesure μ (remarquons que nous ne supposons pas qu'un système ergodique préserve la mesure).

Finalement, une fonction f définie sur Ω et à valeurs réelles est un *cobord* s'il existe une fonction u, mesurable et à valeurs réelles sur Ω, telle que :

$$f = u \circ T - u.$$

Si la mesure μ est finie, on peut la multiplier par une constante pour qu'elle soit de masse 1 ; ainsi, on peut supposer sans perte de généralité que μ est une mesure de probabilité. On dit alors qu'un tel système qui préserve la mesure est *mélangeant* si, pour tous boréliens A et B,

$$\lim_{n\to+\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B).$$

La propriété de mélange s'exprime aussi d'un point de vue fonctionnel : un système est mélangeant si et seulement si, pour toutes fonctions f et g dans $\mathbb{L}^2(\Omega, \mu)$,

$$\lim_{n\to+\infty} \int_\Omega f \cdot g \circ T^n \, \mathrm{d}\mu = \int_\Omega f \, \mathrm{d}\mu \cdot \int_\Omega g \, \mathrm{d}\mu,$$

autrement dit, les observables f et $g \circ T^n$ sont asymptotiquement de covariance nulle. En appliquant cette propriété à $h \circ f$ et $h' \circ g$, où h et h' sont des fonctions continues bornées sur $\mathbb{R}$, on s'aperçoit que f et $g \circ T^n$ sont en fait asymptotiquement indépendantes. En prenant $A = B$ dans la définition du mélange, on peut vérifier que le mélange implique l'ergodicité.

En mesure finie, la notion de conservativité est souvent redondante :

Théorème 1.0.6 (Théorème de récurrence de Poincaré).

Considérons un système dynamique mesuré en temps discret. Supposons que ce système préserve une mesure de probabilité. Alors ce système est conservatif.

Remarque 1.0.7.

Le théorème 1.0.6 n'est plus vrai si l'on ne suppose pas que la mesure μ est finie : par exemple, la transformation $n \mapsto n+1$ définie sur $\mathbb{Z}$ muni de la mesure de comptage préserve la mesure et est ergodique. Cependant, ce système n'est pas conservatif.

La raison de cette différence entre théorie ergodique en mesure finie et théorie ergodique en mesure infinie est qu'il y a en fait deux définitions de l'ergodicité. Par l'absence de "sous-ensemble invariant non trivial", on peut entendre l'absence de sous-ensemble B non trivial tel que $B = T^{-1}B$, mais aussi l'absence de sous-ensemble B non trivial tel que $B \subset T^{-1}B$. Pour un système dynamique préservant une mesure de probabilité, il n'y a pas de différence entre ces deux définitions (voir [2, proposition 1.2.2], [2, proposition 1.1.3] et le théorème de Birkhoff). Cependant, en mesure infinie, ces deux définitions diffèrent sensiblement.

Le théorème 1.0.6 reste vrai en mesure infinie si l'on utilise la seconde définition de l'ergodicité. Cependant, c'est la première définition qui est utilisée en pratique [2], et donc c'est celle-ci que nous adoptons. Ceci qui explique que nous parlerons souvent de systèmes "ergodiques et conservatifs" ou bien parfois "récurrents", au lieu de se contenter de systèmes ergodiques.

Enfin, si la majeure partie de cette thèse est dédiée à des systèmes non inversibles, nous traiterons quelques systèmes inversibles. Afin de passer de la première situation à la seconde, nous utiliserons la notion d'*extension naturelle*, définie par exemple dans [36]. Pour cela, donnons-nous un système dynamique mesurable surjectif (A, d, T_A) sur un espace métrique polonais borné. Soit $\widehat{A}$ la limite projective de la suite $(A_n)_{n\in\mathbb{N}}$, où $A_n = A$ pour tout n et $T_A^n \ : \ A_{n+m} \to A_n$. En d'autre termes,

$$\widehat{A} := \{(x_n)_{n\in\mathbb{N}} \in A^{\mathbb{N}} \ : \ T_A x_{n+1} = x_n \forall n \geq 0\}.$$

L'ensemble $\widehat{A}$ est métrisable, par exemple en posant $\hat{d}((x_n),(y_n)) := \sum_n 2^{-n} d(x_n, y_n)$; en pratique, nous choisirons d'autres distances ayant de meilleures propriétés. Enfin, nous disposons d'une transformation $\widehat{T}_A$ sur $\widehat{A}$ définie par $\widehat{T}_A(x_n)_{n\in\mathbb{N}} = (T_A x_0, x_0, x_1, \ldots)$. Le système dynamique $(\widehat{A}, \widehat{T}_A)$ est alors une extension inversible de (A, T_A). Elle vérifie une propriété universelle : toute extension inversible de (A, T_A) est une extension de $(\widehat{A}, \widehat{T}_A)$. De plus, si (A, T_A) préserve une mesure de probabilité μ_A, alors par le théorème d'extension de Kolmogorov, il existe une mesure de probabilité $\hat{\mu}_A$ sur $\widehat{A}$ préservée par $\widehat{T}_A$ et dont la projection sur A est μ_A.

Passons maintenant aux systèmes dynamiques en temps continu, qu'ils soient inversibles (on parle de *flots*) ou non-inversibles (on parle alors de *semi-flots*). Un système dynamique mesuré inversible en temps continu est la donnée d'un triplet $(\Omega, \mu, (g_t)_{t\in\mathbb{R}})$, où :

- Ω est un espace polonais, c'est-à-dire un espace topologique métrisable tel qu'il existe une distance le rendant complet ;
- μ est une mesure σ-finie, positive et non nulle sur Ω ;
- $(g_t)_{t\in\mathbb{R}}$ est un groupe de transformations mesurables et non singulières de Ω.

Pour obtenir la définition des semi-flots, il suffit de remplacer $(g_t)_{t\in\mathbb{R}}$ par $(g_t)_{t\in\mathbb{R}_+}$, et "groupe" par "semi-groupe".

Les définitions des propriétés de préservation de la mesure, d'ergodicité et de mélange peuvent être modifiées de façon élémentaire pour s'adapter au contexte des (semi-)flots. La notion de cobord est un peu plus délicate ; pour éviter de faire appel à une sorte de "générateur infinitésimal" du flot (comme par exemple un champ de vecteurs sur une variété), le plus simple est d'adopter une définition intégrale. En temps discret, une fonction f définie sur Ω et à valeurs réelles est un cobord s'il existe une fonction u, mesurable et à valeurs réelles sur Ω, telle que, pour tout n :

$$\sum_{k=0}^{n-1} f \circ T^k = u \circ T^n - u.$$

En s'inspirant de cette définition, en temps continu, on dit qu'une fonction f mesurable bornée sur Ω et à valeurs réelles est un cobord s'il existe une fonction u, mesurable et à valeurs réelles sur Ω, telle que, pour tout t :

$$\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_t - u.$$

Suivant les situations, l'hypothèse selon laquelle f doit être bornée peut être affaiblie ; être localement intégrable suffira le plus souvent. Ce type de définition s'adapte à des actions de groupes plus généraux.

1.1 Applications Gibbs-Markov

Les applications markoviennes avec un espace d'états fini et une condition de régularité (condition d'Adler ou condition de distortion bornée) ont été étudiées en détail. Elles se retrouvent par exemple quand on étudie le flot géodésique sur des surfaces compactes de courbure négative. Le fait qu'un système dynamique ait un sous-système induit (ou une section) qui est markovien apporte donc beaucoup d'informations sur le système initial : décroissance des corrélations, mesures de Gibbs, etc.

Cependant, si le système dynamique peut passer un temps arbitrairement long au voisinage d'un point fixe ou au voisinage de l'infini, on ne peut pas se ramener à une telle application markovienne - ou alors, le temps de premier retour ne serait pas régulier. Afin de contourner ce problème, on peut travailler avec des applications markoviennes avec un espace d'états dénombrable. Cependant, le fait qu'il existe une infinité d'états possibles peut engendrer des phénomènes pathologiques. On rajoute alors une hypothèse supplémentaire sur le système, la propriété de grande image. La famille d'applications ainsi obtenue est celle des applications Gibbs-Markov. Nous renvoyons le lecteur au livre de J. Aaronson [2, chapitre 4] pour une présentation détaillée de ces applications.

1.1.1 Définition et propriétés élémentaires

Avant toutes choses, voici la définition d'une application Gibbs-Markov.

Définition 1.1.1 (Applications Gibbs-Markov).

Soient (A, d) un espace métrique polonais et borné, $\mathcal{B}$ la tribu borélienne associée, une partition dénombrable π de A, un mesure de probabilité μ_A sur A et une application mesurable T_A de A dans A. Le système dynamique $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ est Markov si :

- *Pour tout élément a de π, l'image de a par T_A est une réunion d'éléments de π (à un ensemble de mesure nulle près) ;*
- *Pour tout élément a de π, l'application T_A est un isomorphisme de a sur son image ;*
- *Le complété de la tribu $\bigvee_{n\in\mathbb{N}} T_A^{-n}\pi$ pour μ_A est la tribu borélienne $\mathcal{B}$;*
- *L'application T_A préserve la mesure μ_A.*

Ce système dynamique est Gibbs-Markov si, en plus, il vérifie les propriétés suivantes :

- *Propriété de la grande image : $\inf_{a\in\pi} \mu_A(T_A a) > 0$;*
- *Dilatation locale uniforme : il existe un $\lambda > 1$ tel que, pour tout $a \in \pi$ et pour presque tous x, y dans a, l'on ait $d(T_A x, T_A y) \geq \lambda d(x, y)$;*
- *Distorsion lipschitzienne : il existe une constante C telle que, pour tout $a \in \pi$, pour presque tous x, y dans a :*

$$\left| \frac{\mathrm{d}\mu_A}{\mathrm{d}\mu_A \circ T}(x) - \frac{\mathrm{d}\mu_A}{\mathrm{d}\mu_A \circ T}(y) \right| \leq C d(Tx, Ty) \frac{\mathrm{d}\mu_A}{\mathrm{d}\mu_A \circ T}(x). \tag{1.1.1}$$

D'après cette définition, un système dynamique Gibbs-Markov est la donnée des six objets $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$. Cependant, par abus de notation, on omettra le plus souvent une partie des données quand celles-ci ne prêtent pas à confusion.

Dans toute cette partie, (A, π, d, μ_A, T_A) désigne une application Gibbs-Markov.

Propriété de distorsion

Si T est une application de classe $\mathcal{C}^2$ d'un intervalle I, alors la propriété de distorsion lipschitzienne est impliquée par la propriété suivante [2, chapitre 4.3] :

$$\sup_I \frac{|T''|}{|T'|^2} < +\infty.$$

Cette condition, appelée *condition de Rényi* ou *condition d'Adler*, apparaît fréquemment dans des travaux sur les transformations dilatantes de classe $\mathcal{C}^2$ de l'intervalle, par exemple [16]. La condition de distorsion lipschitzienne peut être vue comme une extension de la condition de Rényi quand l'espace Ω n'est pas une variété ; en effet, nous aurons souvent à changer la distance et la topologie sur Ω, qui pourra avoir une structure d'ensemble de Cantor (en d'autres termes, être parfait et totalement discontinu).

Pour presque tout x dans A, on notera $g(x) := \frac{\mathrm{d}\mu}{\mathrm{d}\mu \circ T_A}(x)$ l'inverse du jacobien de T_A en x, et par $g^{(n)}(x) := \prod_{i=0}^{n-1} g(T_A^i x)$ l'inverse du jacobien de T_A^n. La propriété de distorsion lipschitzienne peut aussi être reformulée en : il existe une constante $C > 0$ telle que, pour tout $a \in \pi$, pour presque tous x, y dans a,

$$|\ln g(x) - \ln g(y)| \leq C d(Tx, Ty).$$

Cylindres et temps de séparation

Un *cylindre* de longueur n est un sous-ensemble $\overline{a} = [a_0, \cdots, a_{n-1}]$ de A tel que les a_i soient des élément des π, et :

$$\overline{a} = \bigcap_{i=0}^{n-1} T_A^{-i} a_i.$$

Soit $\lambda > 1$ un facteur d'expansion compatible avec l'application Gibbs-Markov. Pour tous points x et y dans A, on appelle *temps de séparation* de x et y pour la partition π et l'application T_A la quantité :

$$s(x,y) := \inf\left\{n \geq 0 \ : \ T_A^n x \text{ et } T_A^n y \text{ n'appartiennent pas au même élément de } \pi\right\}.$$

En d'autres termes, $s(x, y)$ est la longueur du plus long cylindre qui contienne à la fois x et y. La propriété de distorsion lipschitzienne peut être renforcée de la façon suivante :

Lemme 1.1.2 (Lemme de distorsion)**.**
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov. Il existe une constante C telle que, pour presque tous x, y dans A, pour tout $n \leq s(x, y)$,

$$\left|g^{(n)}(x) - g^{(n)}(y)\right| \leq C d(T_A^n x, T_A^n y) g^{(n)}(x), \tag{1.1.2}$$

et, pour tout cylindre $\overline{a}$ de longueur n et pour tout $x \in \overline{a}$:

$$g^{(n)}(x) \leq C\mu(\overline{a}). \tag{1.1.3}$$

De plus, pour tout $\kappa > 1$, on peut définir une distance ultramétrique d_κ sur A par :

$$d_\kappa(x, y) := \kappa^{-s(x,y)}.$$

Si $\kappa \in (1, \lambda]$, alors l'application $(A, \pi, d_\kappa, \mu_A, T_A)$ est encore Gibbs-Markov. On supposera parfois que la distance sur A est de la forme d_κ par un certain paramètre κ, auquel cas on dira que $(A, \pi, \kappa, \mu_A, T_A)$ est Gibbs-Markov.

1.1.2 Propriétés spectrales des applications Gibbs-Markov

La théorie spectrale de l'opérateur de transfert (ou d'opérateurs de transfert généralisés) joue un rôle central dans l'étude des systèmes dynamiques hyperboliques. Tout d'abord, définissons l'opérateur de transfert et montrons quelques-unes de ses propriétés élémentaires.

Propriétés de l'opérateur de transfert

L'*opérateur de transfert*, ou *opérateur de Perron-Frobenius*, que l'on notera $\mathcal{L}$, est défini sur $\mathbb{L}^1(A, \mu_A)$ comme le dual de l'opérateur de composition par T_A :

$$\int_A h \cdot \mathcal{L}f \, \mathrm{d}\mu_A = \int_A h \circ T_A \cdot f \, \mathrm{d}\mu_A \quad \forall \ f \in \mathbb{L}^1(A, \mu_A), \ \forall \ h \in \mathbb{L}^\infty(A, \mu_A).$$

Grâce à la formule du changement de variable, l'opérateur de transfert est aussi défini par :

$$\mathcal{L}f(x) = \sum_{\{y \ : \ T_A y = x\}} g(y) f(y) \quad \forall \ f \in \mathbb{L}^1(A, \mu_A). \tag{1.1.4}$$

Cette dernière formule se généralise aux itérés de l'opérateur de transfert. Pour tout $n \geq 0$, pour toute fonction $f \in \mathbb{L}^1(A, \mu_A)$ et pour presque tout $x \in A$,

$$\mathcal{L}^n f(x) = \sum_{\{y \ : \ T_A^n y = x\}} g^{(n)}(y) f(y).$$

La mesure μ_A étant T_A-invariante, 1 est une valeur propre de $\mathcal{L}$ et le sous-espace propre correspondant contient les fonctions constantes. On en déduit par l'inégalité de Jensen que l'opérateur $\mathcal{L}$ est une contraction sur $\mathbb{L}^p(A, \mu_A)$ pour tout $p \in [1, \infty]$, et que son rayon spectral est de 1.

Quasi-compacité

L'un des caractères remarquables de cette théorie est que les opérateurs de transfert, agissant sur des espaces fonctionnels bien choisis, sont parfois *quasi-compacts*. C'est notamment le cas pour les applications Gibbs-Markov.

Définition 1.1.3 (Quasi-compacité)**.**
Soient $\mathcal{B}$ un espace de Banach complexe, et $\mathcal{L}$ un opérateur agissant continûment sur $\mathcal{B}$.

Le rayon spectral essentiel de $\mathcal{L}$ est le plus petit réel $r \geq 0$ tel que, pour tout $\varepsilon > 0$, le spectre de $\mathcal{L}$ hors de la boule centrée en zéro et de rayon $r + \varepsilon$ est constitué d'un nombre fini de points, qui sont des valeurs propres de multiplicité finie. En d'autres termes, pour tout ρ dans le spectre de $\mathcal{L}$ et hors de la boule centrée en zéro et de rayon $r+\varepsilon$, l'opérateur $\mathcal{L} - \rho \, \mathrm{Id}$ est inversible sur un sous-espace de codimension finie.

L'opérateur $\mathcal{L}$ agissant sur $\mathcal{B}$ est dit quasi-compact si son rayon spectral essentiel est strictement inférieur à son rayon spectral.

Remarquons qu'un opérateur compact a un rayon spectral essentiel de 0, et donc est quasi-compact si son spectre n'est pas réduit à $\{0\}$. Si l'action d'un opérateur $\mathcal{L}$ sur un espace de Banach complexe $\mathcal{B}$ est quasi-compacte, alors on peut trouver un sous-espace vectoriel propre V de dimension finie tel qu'au premier ordre, pour tout vecteur $f \in \mathcal{B} - V$, le comportement asymptotique de la suite $(\mathcal{L}^n f)$ ne dépend que de la projection canonique de f sur V. Autrement dit, si l'on cherche à étudier le comportement en temps grand de $(\mathcal{L}^n f)$, alors tout se passe comme en dimension finie.

Pour démontrer qu'un opérateur est quasi-compact, l'un des outils les plus employés consiste à démontrer tout d'abord une inégalité dite de Lasota-Yorke, aussi appelée inégalité de Döblin-Fortet. Pour cela, on fait agir un opérateur $\mathcal{L}$ sur deux espaces de Banach complexes $\mathcal{B}$ et $\mathcal{B}'$ tels que :

- $\mathcal{L} \ : \ \mathcal{B} \to \mathcal{B}'$ est compact ;
- Il existe deux suites de réels positifs $(r_n)_{n \in \mathbb{N}}$ et $(R_n)_{n \in \mathbb{N}}$ telles que, pour tout n et pour tout $f \in \mathcal{B}$,
$$\|\mathcal{L}^n f\|_{\mathcal{B}} \leq r_n \|f\|_{\mathcal{B}} + R_n \|f\|_{\mathcal{B}'} .$$

D'après une résultat de H. Hennion [37, corollaire 1], le rayon spectral essentiel de $\mathcal{L}$ agissant sur $\mathcal{B}$ vaut alors au plus $r := \liminf_{n \to +\infty} r_n^{1/n}$. Il suffit alors de montrer que r est strictement plus petit que le rayon spectral de $\mathcal{L}$. Si $\mathcal{L}$ est l'opérateur de transfert et si les fonctions constantes appartiennent à $\mathcal{B}$, cela revient à montrer que $r < 1$. De plus, en pratique, la première hypothèse est remplacée par "$\mathcal{L}$ agit continûment sur $\mathcal{B}$ et l'inclusion $\mathcal{B} \to \mathcal{B}'$ est compacte".

Le résultat de H. Hennion est une amélioration d'un théorème de C.T. Ionescu-Tulcea et G. Marinescu [42], qui rajoutaient l'hypothèse $\sup_n \|\mathcal{L}^n\|_{\mathcal{B}' \to \mathcal{B}'} < +\infty$ et ne concluait qu'à la quasi-compacité de l'opérateur, sans donner de borne supérieure sur le rayon spectral essentiel.

Dans le cadre des applications Gibbs-Markov, on peut trouver des espaces de Banach $\mathcal{B}$ et $\mathcal{B}'$ convenables, puis démontrer une inégalité de Lasota-Yorke [2, chapitre 4.7].

Fonctions localement lipschitziennes

Pour tout sous-ensemble $\omega \subset A$, on note $|\cdot|_{\mathrm{Lip}_d(\omega)}$ la semi-norme des fonctions lipschitziennes sur ω : pour toute fonction sur ω à valeurs dans un espace métrique (E, d') :

$$|f|_{\mathrm{Lip}_d(\omega)} := \mathrm{Esup} \left\{ \frac{d'(f(x), f(y))}{d(x,y)} \ : \ (x,y) \in \omega^2, \ x \neq y \right\}.$$

On définit de plus une application :

$$D_{\pi,d} := \left\{ \begin{array}{lll} \mathcal{C}(A, E) & \to & \mathcal{C}(A, \mathbb{R}_+ \cup \{+\infty\}) \\ f & \mapsto & \sum_{a \in \pi} |f|_{\mathrm{Lip}_d(a)} 1_a \end{array} \right. .$$

Par exemple, $\|D_{\pi,d}(f)\|_\infty = \sup_{a\in\pi} |f|_{\mathrm{Lip}_d(a)}$ et $\mathbb{E}_\mu(D_{\pi,d}(f)) = \sum_{a\in\pi} \mu(a) |f|_{\mathrm{Lip}_d(a)}$.

Pour toute fonction f à valeurs dans un espace métrique, la fonction $D_{\pi,d}(f)$ joue le rôle de la valeur absolue de la dérivée de f. Comme d'habitude en analyse, les contraintes sur la régularité de f seront exprimées en tant que contraintes sur l'intégrabilité de $D(f)$. L'emploi de l'application $D_{\pi,d}$, au lieu de l'emploi direct de quantités telles que $\sum_{a\in\pi} \mu(a) |f|_{\mathrm{Lip}_d(a)}$, nous donne plus de flexibilité dans la manipulation des contraintes de régularité. Ce sera utile dans les parties 2.5 (étude de conditions faibles de régularité) et 3.3 (étude de $\mathbb{Z}^d$-extensions d'applications Gibbs-Markov).

Enfin, on note Lip_d^∞ l'espace des fonctions à valeurs réelles sur A telles que la norme $\|\cdot\|_{\mathrm{Lip}_d^\infty} := \|\cdot\|_{\mathbb{L}^\infty} + \|D_{\pi,d}(\cdot)\|_\infty$ est finie. L'injection canonique $\mathrm{Lip}_d^\infty \hookrightarrow \mathbb{L}^\infty$ est compacte.

Soit $h \in (0,1]$. Comme il existe une constante C telle que $d \leq C d_\lambda$, on a $d^h \leq C^h d_\lambda^h = C^h d_{\lambda^h}$. Par conséquent, pour tout $\kappa \in (1, \lambda^h]$, toute fonction h-höldérienne pour la distance d est lipschitzienne pour la distance d_κ. Tout résultat exprimé pour des fonctions lipschitziennes est en fait valable pour des fonctions höldériennes.

De même que l'on peut noter une application Gibbs-Markov avec son facteur de dilatation $(A, \pi, \lambda, \mu_A, T_A)$, ou omettre une partie des données, en l'absence d'ambiguïté, on pourra noter Lip_λ ou simplement Lip au lieu de Lip_d. De même, on notera $D_{\pi,\lambda}$ ou simplement D au lieu de $D_{\pi,d}$, ainsi que $\mathrm{Lip}_\lambda^\infty$ ou Lip^∞ au lieu de Lip_d^∞.

Action de l'opérateur de transfert sur Lip^∞

Finalement, on fait agir l'opérateur de transfert sur l'espace des fonctions globalement lipschitziennes Lip_d^∞.

Lemme 1.1.4.
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov, et soit λ son facteur de dilatation. Alors le rayon spectral essentiel de l'opérateur de transfert $\mathcal{L}$ agissant sur Lip_d^∞ est d'au plus $1/\lambda$. En particulier, cette action est quasi-compacte.

Les propriétés élémentaires du système dynamique (A, π, d, μ_A, T_A) se lisent sur le spectre de $\mathcal{L}$. Pour tout entier $p \geq 1$, notons $\mathbb{U}_p$ l'ensemble des racines p-ièmes de l'unité.

Corollaire 1.1.5.
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov. Alors il existe un entier $N \geq 1$ et un N-uplet d'entiers strictement positifs $(p_1, \cdots, p_N)$ tels que l'ensemble des valeurs propres de module 1 de $\mathcal{L}$ agissant sur Lip_d^∞ est $\bigcup_{k=1}^N \mathbb{U}_{p_k}$, et pour tout ρ dans cet ensemble, la multiplicité (algébrique et géométrique) de ρ est $\mathrm{Card}\{k \ : \ \rho \in \mathbb{U}_{p_k}\}$.

Le système est ergodique si et seulement si $N = 1$. Il est mélangeant si et seulement si $N = 1$ et $p_1 = 1$.

Si une application Gibbs-Markov est ergodique mais pas mélangeante, alors l'un de ses itérés aura une décomposition en composantes ergodiques mélangeantes. Plus précisément, si l'on a une application Gibbs-Markov ergodique $(A, \pi, \lambda, \mu_A, T_A)$, alors il existe un unique entier $M \geq 1$ et une partition $(A_k)_{k\in\mathbb{Z}/M\mathbb{Z}}$ de A en M sous-ensembles tels que :

- La partition $(A_k)_{k\in\mathbb{Z}/M\mathbb{Z}}$ est moins fine que π ;
- Les sous-ensembles A_k sont invariants par T_A^M à un ensemble de mesure nulle près ;
- $T_A(A_k) = A_{k+1}$ pour tout $k \in \mathbb{Z}/M\mathbb{Z}$;
- Les M systèmes dynamiques (A_k, d, T_A^M) sont topologiquement mélangeants.

L'entier M est égal à l'entier p_1 du corollaire 1.1.5.

Propriétés de décorrélation

Si l'application Gibbs-Markov (A, π, d, μ_A, T_A) est mélangeante, alors l'opérateur $\mathcal{L}$ agissant sur Lip^∞ a un trou spectral. En effet, la valeur propre 1 est de multiplicité 1, et correspond à la projection $f \mapsto \int_A f \, \mathrm{d}\mu_A$ sur l'espace des fonctions constantes. Le reste du spectre de $\mathcal{L}$ est inclus dans un disque centré en 0 et de rayon strictement plus petit que 1. Par conséquent, il existe deux constantes $C > 0$ et $\rho \in [0, 1)$ telles que, pour toute fonction $f \in \mathrm{Lip}_d^\infty$, pour tout entier $n \geq 0$,

$$\left\| \mathcal{L}^n f - \int_A f \, \mathrm{d}\mu_A \right\|_{\mathrm{Lip}_d^\infty} \leq C\rho^n \left\| f \right\|_{\mathrm{Lip}_d^\infty} .$$

Le trou spectral peut être exploité pour donner des versions quantitatives de la propriété de mélange, et montrer que le système dynamique possède des caractéristiques chaotiques. En particulier, pourvu que la fonction f ait de bonnes propriétés d'intégrabilité et de régularité, la suite $(f, f \circ T_A, f \circ T_A^2, \cdots)$ ressemble à une suite de variables aléatoires indépendantes et identiquement distribuées.

Nous citons ici quelques conséquences plus avancées de ce trou spectral, en commençant par la décroissance des corrélations à vitesse exponentielle.

Proposition 1.1.6 (Décroissance exponentielle des corrélations).
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Il existe deux constantes $C > 0$ et $\rho \in [0, 1)$ telles que, pour toutes fonctions $f \in \mathrm{Lip}^\infty$ et $h \in \mathbb{L}^1(A, \mu_A)$, pour tout entier $n \geq 0$,

$$\left| \int_A f \cdot h \circ T_A^n \, \mathrm{d}\mu_A - \int_A f \, \mathrm{d}\mu_A \cdot \int_A h \, \mathrm{d}\mu_A \right| \leq C\rho^n \left\| f \right\|_{\mathrm{Lip}^\infty} \left\| h \right\|_{\mathbb{L}^1} . \tag{1.1.5}$$

La deuxième proposition que nous rappelons est un théorème central limite pour des fonctions régulières, qui peut être prouvé aussi bien par des méthodes de perturbation de l'opérateur de transfert ([31, Theorem 3.7], des techniques proches sont utilisées dans la partie 3.2) que par des méthodes de martingales (voire les parties 2.5 et 2.6).

Proposition 1.1.7 (Théorème central limite).
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soit $f \in \mathbb{L}^2(A, \mu_A)$ une fonction telle que $\int_A f \, \mathrm{d}\mu_A = 0$ et $\mathbb{E}(D(f)) < +\infty$. Alors :

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} f \circ T_A^i \to \sigma(f)\mathcal{N},$$

où la convergence est en loi, $\mathcal{N}$ est une variable aléatoire gaussienne standard, et :

$$\sigma(f)^2 = \int_A f^2 \, \mathrm{d}\mu_A + 2 \sum_{n=1}^{+\infty} \int_A f \cdot f \circ T_A^n \, \mathrm{d}\mu_A.$$

De plus, $\sigma(f) = 0$ si et seulement si f est un cobord.

Avec des hypothèses un peu plus fortes, on dispose de l'inégalité de Burkholder-Rosenthal, qui se démontre par des méthodes de martingales. Des versions proches de celle que nous proposons sont nombreuses dans la littérature sur le sujet (voir par exemple [33, Proposition 4.1] pour une inégalité presque aussi forte, et en faite suffisamment forte pour toutes nos applications). Nous en démontrerons une généralisation dans la partie 2.5.

Proposition 1.1.8 (Inégalité de Burkholder-Rosenthal).
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov. Soient $p > 2$, et f une fonction de $\mathbb{L}^p(A, \mu_A)$ telle que $\int_A f \, d\mu_A = 0$ et $\mathbb{E}(D(f)) < +\infty$. Alors il existe une constante $C > 0$ telle que, pour tout entier positif n,

$$\left\| \sum_{i=0}^{n-1} f \circ T_A^i \right\|_{\mathbb{L}^p(A,\mu_A)} \leq C n^{\frac{1}{2}}.$$

Le processus stochastique $(f \circ T_A^n)_{n \geq 0}$ étant stationnaire, l'inégalité de Burkholder-Rosenthal peut être utilisée en conjonction avec un théorème de R.J. Serfling [62, théorème B] pour montrer que, sous les mêmes hypothèses, il existe une constante C telle que, pour tout entier positif n,

$$\left\| \sup_{0 \leq k \leq n} \left| \sum_{i=0}^{k-1} f \circ T_A^i \right| \right\|_{\mathbb{L}^p(A,\mu_A)} \leq C n^{\frac{1}{2}}.$$

1.2 Induction, tours, et accélération

Nous décrivons dans cette section un ensemble de techniques permettant de manipuler l'écoulement du temps dans un système dynamique.

1.2.1 Induction

Soit (Ω, μ, T) un système dynamique mesuré ergodique, conservatif et préservant la mesure μ. Dans le contexte de la théorie ergodique en mesure infinie, on s'intéresse principalement au cas où μ est une mesure infinie et σ-finie. L'outil le plus puissant dont nous disposons pour étudier de tels système est l'induction. Soit A une partie mesurable de Ω de mesure strictement positive et finie. Si la mesure μ est infinie, on peut la multiplier par une constante ; ainsi, on peut supposer sans perte de généralité que $\mu(A) = 1$.

On définit alors le *temps de premier retour en* A, et on note φ_A, l'application définie pour tout $x \in \Omega$ par :

$$\varphi_A(x) := \inf\{n > 0 \ : \ T^n x \in A\}$$

Cette application est à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Le système dynamique étant supposé ergodique et conservatif, φ_A est finie presque partout.

L'*application induite en* A, que l'on notera T_A, est l'application définie pour tout $x \in A$ par :

$$T_A x := T^{\varphi_A(x)} x$$

Cette application est bien définie pour presque tout $x \in A$. On notera de plus μ_A la mesure μ restreinte à A. Le résultat suivant (voir [2, proposition 1.5.3]) est fondamental pour les techniques d'induction.

Théorème 1.2.1.
Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant la mesure μ. Soit A un sous-ensemble de mesure strictement positive de Ω. Alors (A, μ_A, T_A) est un système dynamique préservant la mesure μ_A.

De plus, si (Ω, μ, T) est ergodique alors (A, μ_A, T_A) l'est aussi.

De plus, la formule de Kac [45] relie les mesures μ_A et μ, et l'application φ_A.

Théorème 1.2.2 (Formule de Kac).
Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant la mesure μ. Soit A un sous-ensemble de mesure strictement positive de Ω. Alors :

$$\mu(\Omega) = \int_A \varphi_A \, \mathrm{d}\mu_A.$$

En particulier, si Ω est de mesure infinie, alors φ_A n'est pas intégrable. Les conditions que l'on imposera sur φ_A utiliseront la notion de variation régulière, qui sera introduite en partie 1.3.

Soit (Ω, μ, T) un système dynamique mesuré ergodique, conservatif et préservant la mesure μ. On dit que (Ω, μ, T) *induit un système Gibbs-Markov* s'il existe un borélien A de mesure 1 tel que le système dynamique induit (A, μ_A, T_A) puisse être muni d'une partition π et d'une distance d telles que :

- Le système (A, π, d, μ_A, T_A) est Gibbs-Markov ;
- La restriction du temps de premier retour φ_A à A est $\sigma(\pi)$-mesurable.

Quitte à renormaliser la mesure μ, on peut adapter cette définition au cas où A est de mesure strictement positive et finie.

Finalement, étant donnés un système dynamique mesuré (Ω, μ, T) ergodique, conservatif et préservant la mesure, un borélien A de mesure strictement positive et une observable f définie presque partout sur Ω et à valeurs dans un groupe abélien, on appelle *observable induite en A par f*, et l'on note X_f, la fonction définie pour presque tout $x \in A$ par :

$$X_f(x) = \sum_{k=0}^{\varphi_A(x)-1} f \circ T^k x.$$

L'observable induite X_f est un cobord pour l'application T_A si et seulement si f est un cobord pour l'application T. Le sens indirect est facile à montrer : si $f = u \circ T - u$, alors $X_f = u \circ T_A - u$. Pour montrer le sens direct, supposons que $X_f = \tilde{u} \circ T_A - \tilde{u}$ presque partout sur A. On pose :

$$u(x) := \tilde{u}(T^{\varphi_A(x)} x) - \sum_{k=0}^{\varphi_A(x)-1} f(T^k x),$$

et alors $f = u \circ T - u$.

Dans cette thèse, on s'intéresse aux systèmes dynamiques munis d'une mesure infinie, mais tels que l'application induite en un sous-ensemble bien choisi puisse être munie d'une structure la rendant Gibbs-Markov, et pour lesquels la queue du temps de premier retour est à variation régulière. On renvoie le lecteur à la partie 1.5 pour des exemples de tels systèmes.

1.2.2 Tours de Rokhlin

L'opération inverse de l'induction en un sous-ensemble est la construction de *tours de Rokhlin* au-dessus d'un système dynamique. On se donne maintenant un système dynamique préservant la mesure (A, μ_A, T_A), et une fonction mesurable φ_A définie presque partout sur A et à valeurs dans $\mathbb{N}^*$.

Posons $\Omega := \{(x, n) \in A \times \mathbb{N} \ : \ n < \varphi_A(x)\}$; un point de Ω est décrit par ses coordonnées $(x, n) \in A \times \mathbb{N}$. Définissons alors une application T sur Ω par :

$$T(x, n) := \begin{cases} (x, n+1) & \text{si} \quad n < \varphi_A(x) - 1 \\ (T_A x, 0) & \text{si} \quad n = \varphi_A(x) - 1 \end{cases}$$

Finalement, définissons une mesure μ sur Ω par $\mu(\mathcal{B} \times \{n\}) := \mu_A(\mathcal{B})$ pour tout $n \geq 0$ et pour tout $\mathcal{B} \subset \{\varphi_A > n\}$ mesurable. Le système dynamique (Ω, μ, T) est alors ergodique, conservatif, et préserve la mesure μ. Le système (A, μ_A, T_A) est la *base* de la tour, et la fonction φ_A est la *hauteur* de la tour. Informellement, l'application T fait monter d'un étage tant que c'est possible, et sinon renvoie au pied de la tour et applique T_A.

La construction de tours de Rokhlin commute avec l'induction en la base : si (Ω, μ, T) est la tour de Rokhlin de base (A, μ_A, T_A) et de hauteur φ_A, alors $(A, \mu_A, T_A) \simeq (A \times \{0\}, T_{A\times\{0\}}, \mu_{A\times\{0\}})$, et les fonctions φ_A et $\varphi_{A\times\{0\}}$ sont égales une fois que l'identification canonique est faite. La technique d'induction accélère le temps : pour presque tout $x \in A$, en une seule itération, l'application T_A effectue la même chose que l'application T itérée $\varphi_A(x)$ fois. A l'inverse, la technique de construction de tours ralentit le temps : pour presque tout $x \in A$, la suite de points $(T^n x)$ reste un temps $\varphi_A(x)$ à la verticale de x avant que l'on applique T_A pour la première fois.

De façon générale, si (Ω, μ, T) est un système dynamique mesuré ergodique, conservatif et préservant la mesure μ et si (A, μ_A, T_A) est le système induit sur un sous-ensemble mesurable non trivial A, alors (Ω, μ, T) est isomorphe à la tour de Rokhlin de base (A, μ_A, T_A) et de hauteur φ_A. L'étude des systèmes dynamiques induisant une application Gibbs-Markov est donc équivalente à l'étude de tours au-dessus d'applications Gibbs-Markov. Pour des systèmes dynamiques en temps discret, on utilisera les deux formulations. Pour des systèmes dynamiques en temps continu (c'est-à-dire des flots et semi-flots), les deux approches existent, mais on favorisera l'approche correspondant à la construction de tours, qui est plus facile à formuler. C'est l'objet de la prochaine sous-partie.

1.2.3 Flots de suspension

Les deux sous-sections précédentes portaient sur des systèmes dynamique en temps discret, c'est-à-dire du semi-groupe $\mathbb{N}$ ou du groupe $\mathbb{Z}$. Des constructions équivalentes sont possibles en temps continu, c'est-à-dire pour des actions du semi-groupe $\mathbb{R}_+$ ou du groupe $\mathbb{R}$. La notion correspondant à celle de *système induit* est celle de *section*, et la notion correspondant à celle de *tour de Rokhlin* est celle de *(semi-)flot de suspension*.

De même qu'en temps discret, les notions de section d'un flot et de flot de suspension sont deux points de vue différents d'un même objet. Par simplicité, on ne traitera que des semi-flots de suspension, même si la notion de section apparaîtra brièvement dans certains exemples.

Informellement, pour définir un (semi-)flot de suspension, on se donne une fonction donnant la hauteur du flot ; le flot monte alors à vitesse constante, jusqu'à ce qu'il atteigne le haut de la tour. Il est alors renvoyé sur la base.

Plus rigoureusement, on se donne un système dynamique préservant la mesure (A, T_A, μ_A), et une fonction mesurable φ_A définie presque partout sur A et à valeurs dans $\mathbb{R}_+^*$. Le *semi-flot de suspension* de *base* (A, T_A, μ_A) et de *hauteur* φ_A est le système dynamique $(\Omega, \mu, (g_t)_{t\geq 0})$, où :

- Ω est l'espace quotient de l'espace topologique $A \times \mathbb{R}_+$ par la relation d'équivalence $\sim$ définie par $(x, \varphi_A(x) + t) \sim (T_A(x), t)$ pour presque tous $(x, t) \in A \times \mathbb{R}_+$;
- le semi-flot $(g_t)_{t\geq 0}$ agit sur $A \times \mathbb{R}_+$ par translation sur la seconde coordonnée, et sur Ω par la projection canonique de cette action (qui est bien définie) ;
- $\mu = \mu_A \otimes \text{Leb}$ sur $A \times \mathbb{R}_+$, et est définie sur Ω par restriction au domaine fondamental $\{(x, t) \in A \times \mathbb{R}_+ : \ t \geq \varphi_A(x)\}$.

Le système dynamique $(\Omega, (g_t)_{t\geq 0})$ préserve alors la mesure μ. De plus, un semi-flot de suspension est ergodique si et seulement si sa base l'est.

Une autre façon de définir un semi-flot de suspension est de partir du domaine fondamental. Posons $A \times [0, \varphi_A] := \{(x, t) \in A \times \mathbb{R}_+ \ : \ t \in [0, \varphi_A(x)]\}$, et définissons une relation d'équivalence $\sim$ sur $A \times [0, \varphi_A]$ par $(x, \varphi_A(x)) \sim (T_A(x), 0)$ pour tout x dans A. L'espace Ω peut alors être vu comme l'espace quotient de $A \times [0, \varphi_A]$ par $\sim$. Le flot est défini par translation verticale à vitesse unitaire le long de la seconde coordonnée, en recollant $(x, \varphi_A(x))$ et $(T_A(x), 0)$ quand il le faut.

On identifiera parfois l'ensemble A avec le sous-ensemble $A \times \{0\}$ de Ω, et la mesure μ_A avec la mesure définie sur Ω par $\mu_A(B) := \mu_A(x \in A : \ x \times \{0\} \in B)$ pour tout sous-ensemble borélien B de Ω. La fonction hauteur φ_A peut être étendue à Ω par $\varphi_A(x, t) := \inf\{s > 0 \ : \ g_s(x, t) \in A\}$; en d'autres termes, pour tout $x \in A$ et pour tout $t < \varphi_A(x)$, on a $\varphi_A(x, t) = \varphi_A(x) - t$.

Quand le système (A, T_A, μ_A) peut être muni d'une structure Gibbs-Markov, on parlera de *semi-flot Gibbs-Markov*. De tels semi-flots apparaissent quand l'on étudie le flot géodésique sur des $\mathbb{Z}$ ou $\mathbb{Z}^2$ revêtements de variété compactes de courbure sectionnelle négative, ou des billards hyperboliques qui sont $\mathbb{Z}$ ou $\mathbb{Z}^2$-périodiques.

Finalement, pour toute fonction mesurable f définie sur Ω, à valeurs réelles et telle que l'intégrale $\int_0^{\varphi_A(x)} |f(x,t)| \, \mathrm{d}t$ soit finie pour presque tout $x \in A$, on appelle *observable induite en A par f* la fonction X_f définie presque partout sur A par :

$$X_f \ : \ \left\{ \begin{array}{lll} A & \to & \mathbb{R} \\ x & \mapsto & \int_0^{\varphi_A(x)} f(x,t) \, \mathrm{d}t \end{array} \right. .$$

En particulier, si f appartient à $\mathbb{L}^1(\Omega, \mu)$, alors par le théorème de Fubini la fonction X_f est bien définie sur A et appartient à $\mathbb{L}^1(A, \mu_A)$. L'observable induite X_f est un cobord pour l'application T_A si et seulement si f est un cobord pour le flot de suspension. Une fois encore, le sens indirect est facile à montrer. Pour montrer le sens direct, supposons que $X_f = \tilde{u} \circ T_A - \tilde{u}$ presque partout sur A. On pose pour presque tout $(x,t) \in \Omega$:

$$u(x,t) := \tilde{u}(T_A x) - \int_t^{\varphi_A(x)} f(x,s) \, \mathrm{d}s.$$

Alors, pour tout $t \geq 0$,

$$\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_t - u.$$

1.2.4 Temps local

Aux trois types de systèmes présentés ci-dessus - les systèmes induits, les tours de Rokhlin et les flots de suspension - on peut associer deux fonctions à deux paramètres, le temps local et les temps de retour. Une fois encore, nous commençons par présenter ces définitions en temps discret, dans le cadre des systèmes induits.

Définition 1.2.3 (Temps local et temps de retour).
Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant la mesure. Soit $A \subset \Omega$ un borélien de mesure strictement positive, et soit (A, μ_A, T_A) le système induit en A.

Le temps local en A est la fonction :

$$\xi \ : \ \left\{ \begin{array}{lll} \Omega \times \mathbb{N} & \to & \mathbb{N} \\ (x,N) & \mapsto & \xi_N(x) := \sum_{i=1}^N 1_A \circ T^i x = \mathrm{Card}\{1 \leq i \leq N \ : \ T^i x \in A\} \end{array} \right. .$$

Le temps de passage en A est la fonction :

$$\tau \ : \ \left\{ \begin{array}{lll} A \times \mathbb{N} & \to & \mathbb{N} \\ (x,N) & \mapsto & \tau_N(x) := \sum_{i=0}^{N-1} \varphi_A \circ T_A^i x \end{array} \right. .$$

La fonction "temps de passage" est souvent vue comme une suite de fonctions de N-ième passage $(\tau_0, \tau_1, \ldots)$.

La fonction "temps local en A" évaluée en (x,N) compte le nombre de fois où le système initial (Ω, μ, T), partant de x, passe par l'ensemble A. La fonction "temps de passage en A" évaluée en (x,N) renvoie le temps auquel le système initial (Ω, μ, T), partant de x, revient en A. Pour presque tout $x \in A$, les fonctions $N \mapsto \xi_N(x)$ et $N \mapsto \tau_N(x)$ sont en quelque sorte l'inverse l'un de l'autre. En particulier, pour presque tout $x \in A$, pour tout $N \in \mathbb{N}$, on a :

- $\xi_{\tau_N} = N$;
- $\tau_{\xi_N} \leq N < \tau_{\xi_N+1}$.

Ces définitions et propriétés élémentaires se transposent aisément aux flots de suspension.

Définition 1.2.4 (Temps local et temps de retour en temps continu).
Soit $(\Omega, \mu, (g_t)_{t \geq 0})$ un semi-flot de suspension de base (A, μ_A, T_A) et de fonction hauteur φ_A.

Le temps local en A est la fonction :

$$\xi \ : \ \left\{ \begin{array}{lll} \Omega \times \mathbb{R}_+ & \to & \mathbb{N} \\ (x,t) & \mapsto & \xi_t(x) = \mathrm{Card}\{s \in (0,t] \ : \ g_s x \in A\} \end{array} \right. .$$

Le temps de passage en A est la fonction :

$$\tau \ : \ \left\{ \begin{array}{lcl} A \times \mathbb{N} & \to & \mathbb{R}_+ \\ (x,N) & \mapsto & \tau_N(x) := \sum_{i=0}^{N-1} \varphi_A \circ T_A^i x \end{array} \right. .$$

Une fois encore, ces fonctions sont presque inverses l'une de l'autre au sens suivant : pour presque tout $x \in A$, pour tout $N \in \mathbb{N}$ et pour tout $t \geq 0$, on a :
- $\xi_{\tau_N} = N$;
- $\tau_{\xi_t} \leq t < \tau_{\xi_t+1}$.

Au passage, ces relations permettent de montrer que le temps local est bien défini pour presque tout $x \in \Omega$ et pour tout $t \geq 0$. En effet, sans perte de généralité, on peut supposer que $x \in A$. Le risque est que la fonction $t \mapsto \xi_t(x)$ explose, c'est-à-dire tende vers l'infini en temps fini. Cependant, la fonction $t \mapsto \xi_t(x)$ est croissante, et $\xi_{\tau_N} = N$ pour tout $N \geq 0$. Le temps local explose donc si et seulement si la fonction $N \mapsto \tau_N(x)$ est bornée. D'après le théorème de Birkhoff, c'est un évènement de probabilité nulle.

La notion de temps local définie ci-dessus est une généralisation de la notion de temps local utilisée pour les marches aléatoires. Nous expliciterons en sous-partie 1.5.3 la manière de retrouver le temps local probabiliste à partir de cette définition.

1.2.5 Accélération

Quand nous travaillerons sur des $\mathbb{Z}^d$-extensions d'applications Gibbs-Markov au chapitre 3, il sera utile d'avoir un autre outil pour manipuler les itérées T_A^φ d'une application Gibbs-Markov, où φ est une fonction potentiellement non constante : les temps d'arrêt. Les résultats que nous présentons ici seront aussi utilisés pour passer de théorèmes demandant des applications Gibbs-Markov mélangeantes à des théorèmes demandant des applications Gibbs-Markov ergodiques en partie 2.4.

Le type d'accélération que nous allons définir est aussi appelé *saut de Schweiger*. Nous renvoyons le lecteur à l'article [61] pour la construction originelle, et à [2, Section 4.6] pour une autre présentation de cette méthode dans le cadre des applications Gibbs-Markov.

Définition 1.2.5 (Temps d'arrêt).
Soit $(A, \mathcal{B}, (\mathcal{F}_n)_{n\geq 0})$ un espace mesurable filtré. Une fonction $\varphi : A \to \mathbb{N} \cup \{+\infty\}$ est un temps d'arrêt si $\{\varphi \leq n\} \in \mathcal{F}_n$ pour tout entier positif n. La tribu du passé jusqu'à l'instant φ, que l'on notera $\mathcal{F}_\varphi$, est définie par :

$$\mathcal{F}_\varphi := \{A \in \mathcal{B} : \ A \cap \{\varphi \leq n\} \in \mathcal{F}_n \ \forall n \geq 0\} .$$

La filtration naturelle pour une application Gibbs-Markov $(A, \pi, \lambda, \mu_A, T_A)$ est la suivante :

$$\mathcal{F}_n := \sigma\left(\bigvee_{k=0}^{n-1} T_A^{-k}\pi\right).$$

Pour cette filtration, $\mathcal{F}_0$ est la tribu triviale. Par conséquent, tout temps d'arrêt φ est ou bien nul presque sûrement, ou bien strictement positif presque sûrement. La première possibilité étant inintéressante, on supposera dans la suite que $\varphi > 0$ presque sûrement. De plus, comme nous voulons travailler avec l'application T_A^φ, nous devons aussi supposer que le temps d'arrêt φ est fini presque sûrement.

Si φ est presque sûrement fini, nous désignerons par π_φ la partition dénombrable de A telle que $\mathcal{F}_\varphi = \sigma(\pi_\varphi)$, par s_φ le temps de séparation de points de A pour la partition π_φ et l'application T_A^φ, par d_φ la distance λ^{-s_φ}, et par D_φ l'application $D_{\pi_\varphi, d_\varphi}$.

Grâce à cette construction, l'application accélérée est potentiellement Gibbs-Markov :

Lemme 1.2.6.
Soit $(A, \pi, \lambda, \mu_A, T_A)$ une application Gibbs-Markov. Soit φ un temps d'arrêt pour la filtration naturelle. Supposons que φ est presque sûrement strictement positif et fini, et que T_A^φ préserve la mesure μ_A. Alors $(A, \pi_\varphi, \lambda, \mu_A, T_A^\varphi)$ est une application Gibbs-Markov.

Démonstration.

Étant donné que $\varphi \geq 1$, la partition π_φ est plus fine que π, et *a fortiori* la partition $\bigvee_{i=0}^{n-1}(T_A^\varphi)^{-i}\pi_\varphi$ est plus fine que $\bigvee_{i=0}^{n-1}(T_A)^{-i}\pi$. Par conséquent, la partition π_φ engendre la tribu borélienne sur A sous l'action de T_A^φ.

Soit $\overline{a}$ un élément de π_φ. Alors il existe un entier $n \geq 1$ et un n-uplet $(a_0, \cdots, a_{n-1})$ d'éléments de π tels que $\overline{a} = [a_0, \cdots, a_{n-1}]$ et que $\varphi = n$ sur $\overline{a}$. Par récurrence, on montre que $T_A^i\overline{a} = [a_i, \cdots, a_{n-1}]$ pour tout entier $i < n$; ainsi, T_A^i est un isomorphisme de $\overline{a}$ sur $[a_i, \cdots, a_{n-1}]$. En prenant $i = n-1$, on obtient $T_A^{n-1}\overline{a} = a_{n-1} \in \pi$. On applique T_A une fois de plus, ce qui montre que $T_A^n\overline{a} = T_A^\varphi\overline{a}$ est une union d'éléments de π, que T_A^n est un isomorphisme de $\overline{a}$ dans son image, et que $(A, \pi_\varphi, \mu_A, T_A^\varphi)$ a la propriété de grande image.

La nouvelle distance sur A est $d_\varphi = \lambda^{-s_\varphi}$, donc $d_\varphi(T_A^\varphi x, T_A^\varphi y) = \lambda d_\varphi(x, y)$ dès que x et y sont dans le même élément de π_φ. Finalement, la propriété de distorsion lipschitzienne pour la nouvelle application découle du lemme de distorsion (lemme 1.1.2) appliqué au système initial $(A, \pi, \lambda, \mu_A, T_A)$, en s'aidant du fait que $d_\varphi \geq d$. $\square$

1.3 Fonctions à variation régulière

Il est très fréquent en probabilités et en théorie ergodique en mesure infinie de travailler avec des suites ou des fonctions à variation régulière, qui se comportent bien pour diverses opérations : renormalisation, intégration, transformées de Laplace ou de Fourier, inverse généralisé... La référence la plus complète sur le sujet est le livre de N.H. Bingham, C.M. Goldie et J.L. Teugels [11]. Nous présentons ici leur définition, et quelques-unes de leurs propriétés les plus importantes, qui sont abordées avec plus de détails dans le livre sus-cité.

1.3.1 Définition et propriétés élémentaires

Pour tout nombre réel β, on définit des fonctions à variation régulière d'indice β, en 0 ou en l'infini, qui se comportent d'un certain point de vu comme des fonctions de la forme Cx^β au voisinage du point choisi.

Définition 1.3.1 (Fonction à variation régulière).

Soit β un nombre réel. Une fonction ψ définie sur un voisinage de $+\infty$ dans $\mathbb{R}_+$, à valeurs strictement positives et mesurable est dite à variation régulière d'indice β à l'infini si, pour tout $x > 0$,

$$\lim_{y \to +\infty} \frac{\psi(xy)}{\psi(y)} = x^\beta.$$

Une fonction ψ définie sur un voisinage de 0 dans $\mathbb{R}_+$, à valeurs strictement positives et mesurable est dite à variation régulière d'indice β en zéro si, pour tout $x > 0$,

$$\lim_{y \to 0^+} \frac{\psi(xy)}{\psi(y)} = x^\beta,$$

ou, de façon équivalente, si $\psi(1/x)$ est à variation régulière d'indice $-\beta$ à l'infini, .

Une fonction à variation régulière d'indice 0 est aussi dite à variation lente.

En l'absence d'ambiguïté, on omettra le plus souvent de préciser si une fonction est à variation régulière en l'infini ou en zéro.

Si l'on omet l'hypothèse de mesurabilité, il existe des fonctions à variation régulière exhibant un comportement pathologique ; en particulier, aucun des résultats présentés dans cette partie ne sont valables. L'hypothèse de mesurabilité permet de renforcer de beaucoup la qualité de la convergence dans la définition des fonctions à variation régulière [11, théorème 1.5.2].

Théorème 1.3.2 (Théorème de convergence uniforme).

Soit β un nombre réel. Soit ψ une fonction à variation régulière d'indice β en $+\infty$. Alors :

$$\lim_{y \to +\infty} \frac{\psi(xy)}{\psi(y)} = x^\beta,$$

où la convergence est uniforme en x sur tout compact de $(0, +\infty)$.

Le théorème de Potter [11, théorème 1.5.6] en découle :

Théorème 1.3.3 (Théorème de Potter)**.**
Soit β un nombre réel. Soit ψ une fonction à variation régulière d'indice β en $+\infty$. Alors pour tout $A > 1$, pour tout $\varepsilon > 0$, pour tous x, y suffisamment grands :

$$\frac{\psi(y)}{\psi(x)} \leq A \max\left\{ \left(\frac{y}{x}\right)^{\beta+\varepsilon}, \left(\frac{y}{x}\right)^{\beta-\varepsilon} \right\}.$$

Les fonctions à variation lente en l'infini les plus fréquemment rencontrées sont les fonctions constantes, la fonction logarithme, et parfois la fonction logarithme itéré. Une fonction ψ à variation régulière d'indice β en l'infini peut toujours s'écrire sous la forme :

$$\psi(x) = x^\beta \ell(x),$$

où ℓ est une fonction à variation lente en l'infini. En particulier, l'archétype de la fonction à variation régulière d'indice β en l'infini est la fonction x^β. Ceci explique la forme des primitives de fonctions à variation régulière qui suit.

Primitives

Soit $\beta > -1$. Une primitive de la fonction x^β sur $\mathbb{R}_+$ est la fonction $(1+\beta)^{-1}x^{1+\beta}$; plus généralement, toutes les primitives de x^β sont équivalentes en l'infini, et à variation régulière d'indice $1+\beta$. Le résultat suivant [11, Proposition 1.5.8] étend cette propriété à l'ensemble des fonctions à variation régulière.

Théorème 1.3.4 (Théorème de Karamata)**.**
Soit $\beta > -1$. Soit ψ une fonction à variation régulière d'indice β en l'infini, et soit $M \geq 0$ tel que ψ soit localement intégrable sur $[M, +\infty)$. Alors, au voisinage de l'infini :

$$\int_M^x \psi(y) \, \mathrm{d}y \sim \frac{x\psi(x)}{1+\beta}.$$

Ce théorème échoue pour $\beta < -1$ car la fonction ψ est alors intégrable au voisinage de l'infini par le théorème de Potter, et échoue pour $\beta = 1$ pour la même raison que les primitives de $1/x$ ne sont pas des fonctions constantes. Par changement de variable, on dispose d'un théorème similaire pour les fonctions à variation régulière de paramètre $\beta < -1$ en 0.

1.3.2 Théorèmes taubériens

L'une des forces de la théorie des fonctions à variation régulière est qu'elle forme un contexte confortable pour des théorèmes taubériens. En particulier, on peut sous certaines conditions relier le comportement asymptotique d'une fonction à variation régulière en l'infini avec le comportement asymptotique de sa transformée de Laplace en zéro. C'est l'objet, entre autres, des théorèmes taubériens de Wiener-Ikehara [48, théorème 2.2], de Hardy-Littlewood [48, théorème 15.1] et de Karamata [11, théorème 1.7.1]. Ces théorèmes ont de nombreuses applications, que ce soit en théorie des nombres (par exemple pour obtenir le théorème des nombres premiers), en analyse harmonique et en probabilités.

Le théorème taubérien de Karamata se formule de la façon suivante :

Théorème 1.3.5 (Théorème taubérien de Karamata)**.**
Soit $\beta \geq 0$. Soit ψ une fonction positive, croissante sur $\mathbb{R}_+$ et nulle en 0. Alors ψ est à variation régulière d'indice β en l'infini si et seulement si la fonction $\lambda \mapsto \int_{\mathbb{R}_+} e^{-st} \, \mathrm{d}\psi(t)$ est à variation régulière d'indice $-\beta$ en zéro. De plus, dans ce cas,

$$\int_{\mathbb{R}_+} e^{-st} \, \mathrm{d}\psi(t) \sim \Gamma(1+\beta)\psi(1/s) \text{ en } 0.$$

Classiquement, le théorème taubérien de Karamata est exprimé à l'aide d'une intégrale de Stieltjes. Au cours de notre travail, il sera plus aisé de travailler avec les mesures $\psi^{-1}(t)\,\mathrm{d}t$, ce qui revient à intégrer ψ^{-1}. Le contrôle de l'intégrale de ψ^{-1} étant donné ψ^{-1} se fait grâce au théorème de Karamata. Le chemin inverse, afin de contrôler ψ^{-1}, étant donnée l'intégrale de ψ^{-1}, se fait à l'aide du théorème de densité monotone [11, théorème 1.7.2]. Cependant, le théorème de Karamata, qui permet de contrôler l'intégration, ne permet pas de traiter le cas $\beta = 0$ (ou $\beta = 1$ ci-dessous), que nous laissons à part.

Théorème 1.3.6 (Théorème taubérien de Karamata, $\beta < 1$).
Soit $\beta < 1$. Soit ψ une fonction strictement positive et croissante sur $\mathbb{R}_+$. Alors les propositions suivantes sont équivalentes :
- *ψ est à variation régulière d'indice β en l'infini ;*
- *$\int_{\mathbb{R}_+} \psi(t)^{-1} e^{-st}\,\mathrm{d}t$ est à variation régulière d'indice $\beta - 1$ en 0.*

De plus, si ces propositions sont vérifiées, alors :

$$\int_{\mathbb{R}_+} \frac{1}{\psi(t)} e^{-st}\,\mathrm{d}t \sim \frac{\Gamma(1-\beta)}{s\psi(1/s)} \text{ en } 0.$$

Théorème 1.3.7 (Théorème taubérien de Karamata, $\beta = 1$).
Soit ψ une fonction strictement positive, croissante sur $\mathbb{R}_+$ et à variation régulière d'indice 1 en l'infini. Alors :

$$\int_{\mathbb{R}_+} \frac{1}{\psi(t)} e^{-st}\,\mathrm{d}t \sim \int_0^{1/s} \frac{1}{\psi(t)}\,\mathrm{d}t \text{ en } 0,$$

et le membre de droite est à variation lente en 0.

Il existe plusieurs méthodes pour évaluer la vitesse à laquelle l'aire sous la courbe de la fonction ψ croît en l'infini. La méthode la plus simple consiste à évaluer l'aire entre 0 et x, et à faire tendre x vers l'infini, c'est-à-dire à regarder la limite en l'infini de $\int_{\mathbb{R}_+} \psi^{-1}(t) 1_{[0,x]}(t)\,\mathrm{d}t$. Une autre méthode consiste à “lisser” les fonctions de la forme $1_{[0,x]}$, en prenant à la place des fonctions exponentielles e^{-sx} (qui elles aussi convergent vers 1 uniformément sur tout compact quand s tend vers 0). Les théorèmes ci-dessus affirment que, pour des fonctions à variation régulière, ces deux méthodes donnent le même résultat.

Ces théorèmes s'obtiennent directement pour des fonctions de la forme x^β, grâce à un changement de variable. En effet, pour tout $\beta < 1$, pour tout $s > 0$,

$$\int_{\mathbb{R}_+} \frac{1}{t^\beta} e^{-st}\,\mathrm{d}t = \frac{1}{s^{\beta+1}} \int_{\mathbb{R}_+} \frac{1}{(st)^\beta} e^{-st}\,\mathrm{d}(st) = \frac{1}{s^{1-\beta}} \int_{\mathbb{R}_+} t^{(1-\beta)-1} e^{-t}\,\mathrm{d}t = \frac{\Gamma(1-\beta)}{s^{1-\beta}}.$$

On se servira de ces théorèmes taubériens en partie 3.2. On déduira de relations algébriques entre les transformées de Laplace de différentes fonctions, des relations entre leur comportements asymptotiques.

1.3.3 Inverse généralisé

Les bijections monotones à variation régulière se comportent bien quand on prend leur inverse. Par exemple, pour tout $\beta > 0$, l'inverse de la fonction x^β est la fonction $x^{1/\beta}$, qui est à variation régulière d'indice $1/\beta$. Si les fonctions avec lesquelles on travaille ne sont pas bijectives, on va devoir utiliser la notion d'inverse généralisé.

Définition 1.3.8 (Inverse généralisé d'une fonction càdlàg).
Soit ψ une fonction positive, définie sur un voisinage de $+\infty$ dans $\mathbb{R}_+$, continue à droite avec des limites à gauche (ou càdlàg), croissante et non bornée. On appelle inverse généralisé de ψ, et on note ψ^*, la fonction définie sur $\mathbb{R}_+$ par :

$$\psi^*(y) := \inf\{x \geq 0 : \ \psi(x) > y\}.$$

Avec les conditions que l'on a imposées à la fonction ψ, son inverse généralisé est lui aussi une fonction positive, définie sur un voisinage de $+\infty$ dans $\mathbb{R}_+$, càdlàg, croissante et non bornée. Si ψ est de plus à variation régulière d'indice $\beta > 0$, alors ψ^* est à variation régulière d'indice $1/\beta$ [11, Theorem 1.5.12].

Étant données ces conventions, les inégalités suivantes sont toujours satisfaites, pourvues qu'elles aient un sens :
- $\psi^* \circ \psi(x) \geq x$;
- $\psi \circ \psi^*(y) \geq y$.

Les inégalités réciproques ne sont en général pas vraies. On utilisera à la place les propriétés suivantes :
- Si $x < \psi^*(y)$, alors $\psi(x) \leq y$;
- Si ψ est à variation régulière, alors pour tout $C > 1$, pour tout y suffisamment grand, $\psi \circ \psi^*(y) \leq Cy$.

Si ψ est à variation régulière d'indice $\beta > 0$, cette dernière propriété vient du fait que, la constante $C > 1$ étant fixée, pour tout y suffisamment grand,

$$\psi^*(y) < C^{\frac{1}{2\beta}} \psi^*(y) < \psi^*(Cy).$$

Dans le cas $\beta = 0$, l'inverse généralisé ψ^* est une fonction à variation rapide ("à variation régulière d'indice $+\infty$ "), et cette propriété est encore satisfaite.

Gestion des termes polynomiaux

Soit ψ une fonction positive, définie sur un voisinage de $+\infty$ dans $\mathbb{R}_+$, càdlàg, croissante, non bornée et à variation régulière. On verra souvent apparaître dans cette thèse des expressions du type $C\psi^*(x^\gamma)$ ou $x^\kappa \psi^*(x^\gamma)$. Afin de les manipuler plus aisément, on voudrait absorber le terme constant ou polynomial dans la fonction ψ^*.

Un terme constant se gère facilement. Soit β l'indice de variation de ψ. La fonction ψ^* est à variation régulière d'indice strictement positif (si $\beta > 0$) ou à variation rapide (si $\beta = 0$). Par le théorème de convergence uniforme (Théorème 1.3.2), pour tout $C > 1$, pour tout $\varepsilon > 0$, pour tous $\gamma^* > \gamma > 0$, pour tout y suffisamment grand,

$$C\psi^*(y^\gamma) \leq \psi^*(C^{\beta+\varepsilon} y^\gamma) \leq \psi^*(y^{\gamma^*}). \tag{1.3.1}$$

Les termes d'ordre polynomial sont plus délicats à gérer. Ils sont l'objet du lemme technique suivant.

Lemme 1.3.9.
Soit ψ une fonction positive, définie sur un voisinage de $+\infty$ dans $\mathbb{R}_+$, càdlàg, croissante, non bornée et à variation régulière d'indice $\beta \geq 0$. Soient $\gamma^* > \gamma > 0$ et $\kappa > 0$. Alors, pour tout y suffisamment grand,

$$y^\kappa \psi^*(y^\gamma) \leq \psi^*(y^{\gamma^* + \kappa\beta}).$$

Démonstration.
Soit $\gamma' \in (\gamma, \gamma^*)$. On pose $\varepsilon := \frac{\gamma^* - \gamma'}{\kappa}$. Alors, par le théorème de Potter (Théorème 1.3.3), pour tout y suffisamment grand,

$$\psi(y^\kappa \psi^*(y^\gamma)) \leq y^{\kappa(\beta+\varepsilon)} \psi(\psi^*(y^\gamma)).$$

Si y est suffisamment grand, alors $\psi^*(y^\gamma) < \psi^*(y^{\gamma'})$, et donc $\psi(\psi^*(y^\gamma)) \leq y^{\gamma'}$. Ainsi, on obtient :

$$\psi(y^\kappa \psi^*(y^\gamma)) \leq y^{\gamma^* + \kappa\beta}.$$

On applique l'inverse généralisé ψ^* des deux côtés de l'inégalité :

$$y^\kappa \psi^*(y^\gamma) \leq \psi^*(\psi(y^\kappa \psi^*(y^\gamma))) \leq \psi^*(y^{\gamma^* + \kappa\beta}).$$

□

1.3.4 Quelques applications en théorie des probabilités

Présentons maintenant quelques applications de la théorie des fonctions à variation régulière en théorie des probabilités. La première concerne les lois stables de Lévy, et une généralisation du théorème central limite. Le contenu de cette sous-partie peut être trouvé dans le livre de I.A. Ibragimov et Y.V Linnik [41].

Lois stables de Lévy

Les lois stables de Lévy sont une famille à quatre paramètres de lois de probabilités, dont certaines propriétés sont réminiscentes des lois normales (qui sont des lois stables de Lévy pour des paramètres bien choisis). Elles sont définies par leur transformée de Fourier.

Définition 1.3.10 (Lois stables de Lévy).
Soient $\alpha \in (0,2]$, $\beta \in [-1,1]$, $\gamma \in \mathbb{R}$ et $c \geq 0$. Soit $\operatorname{sgn}(t)$ *la fonction signe sur $\mathbb{R}$, avec* $\operatorname{sgn}(0) = 0$. *Posons, pour tout $t \in \mathbb{R}$:*

$$\omega(\alpha, t) := \begin{cases} \operatorname{sgn}(t) \tan\left(\frac{\pi\alpha}{2}\right) & si \quad \alpha \neq 1; \\ \frac{2}{\pi} \operatorname{sgn}(t) \ln|t| & si \quad \alpha = 1. \end{cases}$$

On dit qu'une variable aléatoire à valeurs réelles X suit une loi stable de Lévy de paramètres $(\alpha, \beta, \gamma, c)$ si, pour tout $t \in \mathbb{R}$:

$$\mathbb{E}(e^{itX}) = e^{i\gamma t - c|t|^\alpha(1 - i\beta\omega(\alpha,t))}.$$

Si $\alpha = 2$, la fonction $\omega(\alpha, \cdot)$ est identiquement nulle. Une loi stable de Lévy de paramètres $(2, 0, \gamma, c)$ est une loi normale d'espérance γ et de variance $2c$. Une loi stable de Lévy de paramètres $(1, 0, \gamma, c)$ est une loi de Cauchy centrée de γ. Les lois de Lévy sont absolument continues par rapport à la mesure de Lebesgue ; hors de quelques paramètres, leur densité ne s'exprime pas avec des fonctions usuelles, mais est analytique et entière si $\alpha > 1$.

Le paramètre γ décrit le point où la variable aléatoire est centrée ; en particulier, si $\alpha \in (1,2]$, le paramètre γ est l'espérance de la variable aléatoire. Le paramètre c décrit la dispersion de la variable aléatoire, et le paramètre β son asymétrie.

Les lois stables de Lévy sont *infiniment divisibles* et stables, au sens suivant. Soit X une variable aléatoire qui suit une loi stable de Lévy de paramètres $(\alpha, \beta, \gamma, c)$. Supposons que $\alpha \neq 1$ ou que $\alpha = 1$ et $\beta = 0$. Pour tout entier $N \geq 1$, soient X_0, ..., X_{N-1} des variables aléatoires indépendantes, identiquement distribuées et de même loi que X. Alors $\sum_{k=0}^{N-1} X_k$ suit une loi stable de Lévy de paramètres $(\alpha, \beta, N\gamma, N^{1/\alpha}c)$. Si $\alpha = 1$ et $\beta \neq 0$, alors $\sum_{k=0}^{N-1} X_k$ suit une loi stable de Lévy de paramètres $(\alpha, \beta, N(\gamma + 2\beta c \ln(N)/\pi), Nc)$.

Une des propriétés les plus importantes des lois stables de Lévy est qu'elles permettent de formuler un équivalent du théorème central limite pour des variables aléatoires de variance infinie. Chaque loi stable a un certain bassin d'attraction, et la suite des sommes de variables aléatoires indépendantes et identiquement distribuées dans le bassin d'attraction d'une loi stable, correctement renormalisée, converge vers cette loi stable.

Définition 1.3.11 (Bassin d'attraction d'une loi stable de Lévy).
Soient $\beta \in [-1,1]$, $\gamma \in \mathbb{R}$ et $c \geq 0$.

Si $\alpha \in (0,2)$, on dit qu'une variable aléatoire X est dans le bassin d'attraction d'une loi stable de Lévy de paramètre α s'il existe des réels positifs C_1, C_2 non tous deux nuls et une fonction à variation régulière ψ d'indice α en l'infini tels que :

$$\left.\begin{array}{lcl} \mathbb{P}(X \geq x) & \sim & \frac{C_1}{\psi(x)}; \\ \mathbb{P}(X \leq -x) & \sim & \frac{C_2}{\psi(x)} \end{array}\right. \quad \textit{en l'infini,}$$

avec la convention $\mathbb{P}(X \geq \pm x) \sim 0$ si et seulement si $\mathbb{P}(X \geq \pm x) = o(\psi(x)^{-1})$.

Une variable aléatoire X est dans le bassin d'attraction d'une loi normale si elle est de variance finie, ou s'il existe une fonction à variation régulière ψ d'indice 2 en l'infini telle que :

$$\mathbb{P}(X \geq x) + \mathbb{P}(X \leq -x) \sim \frac{1}{\psi(x)} \text{ en l'infini.}$$

Une telle variable aléatoire est dite dans le bassin d'attraction non standard de la loi normale si elle est dans le bassin d'attraction de la loi normale, et est de variance infinie.

La principale application du bassin d'attraction non standard de la loi normale est l'étude du modèle du gaz de Lorentz à horizon infini. Les variables aléatoires dans ce bassin ne satisfont pas le théorème central limite habituel (il n'est satisfait que par les variables aléatoires de variance finie), mais vérifient un théorème central limite avec une renormalisation non standard.

D'après [41, théorème 2.6.5], si $\alpha \in (1,2)$ et X est une variable aléatoire d'espérance nulle, alors X est dans le bassin d'attraction d'une loi stable de paramètre α si et seulement s'il existe $\beta \in [-1,1]$ et une fonction $\tilde{\psi}$, à variation régulière de paramètre α en 0, telle que $\mathbb{E}(e^{itX}) \sim 1 - \tilde{\psi}(|t|)(1 - i\beta\omega(\alpha,t))$. De plus, la fonction $\tilde{\psi}$ et la constante β sont données par :

$$\begin{cases} \tilde{\psi}(t) & \sim \quad (C_1 + C_2)\Gamma(1-\alpha)\cos\left(\frac{\pi\alpha}{2}\right)\psi\left(\frac{1}{t}\right); \\ \beta & = \quad \frac{C_1 - C_2}{C_1 + C_2}. \end{cases}$$

Pour $\alpha = 2$, si X est une variable aléatoire d'espérance nulle dans le bassin d'attraction non standard de la loi normale, alors il existe une fonction $\tilde{\psi}$, à variation régulière de paramètre 2 en 0, telle que $\mathbb{E}(e^{itX}) \sim 1 - \tilde{\psi}(|t|)$. De plus, la fonction $\tilde{\psi}$ est donnée par :

$$\tilde{\psi}(t) \sim \frac{t^2}{2}\mathbb{E}(X^2 1_{|X| \leq 1/t}).$$

Dans les deux cas, on peut supposer sans perte de généralité que $\tilde{\psi}$ est au voisinage de 0 une injection continue. Soient $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes et identiquement distribuées de même loi que X. Alors la suite de variables aléatoires $\tilde{\psi}^*(1/n) \sum_{k=0}^{n-1} X_k$ converge en loi vers une loi stable de Lévy de paramètres $(\alpha, \beta, 0, 1)$. C'est une généralisation du théorème central limite, qui se montre elle aussi en manipulant les transformées de Fourier des variables aléatoires considérées.

Variables aléatoires à queue épaisse

Les fonctions à variation régulière apparaissent aussi en théorie des probabilités via les variables aléatoires à queue épaisse, qui seront omniprésentes dans cette thèse. Comme nous l'avons remarqué en sous-partie 1.2.1, à cause de la formule de Kac, le temps de premier retour en un sous-ensemble mesurable est non intégrable si le système initial est muni d'une mesure infinie et le système induit d'une mesure finie. La notion de variation régulière va entre autres nous servir à formuler des hypothèses sur le temps de premier retour.

Soit φ une variable aléatoire positive. La première condition dont on se servira est la suivante :

$$\forall x > 0, \ \mathbb{P}(\varphi > x) \leq 1/\psi(x), \tag{1.3.2}$$

où la fonction ψ est croissante, non bornée, càdlàg, et à variation régulière d'indice $\beta \in [0,1]$ à l'infini.

On aura aussi besoin d'une condition plus forte sur les queues de φ, comme par exemple :

$$\forall x > 0, \ \mathbb{P}(\varphi > x) = 1/\psi(x), \tag{1.3.3}$$

où la fonction ψ est à variation régulière d'indice $\beta \in [0,1]$ à l'infini (par construction, ψ est automatiquement croissante, non bornée et càdlàg).

1.4 Théorèmes limite en mesure infinie

Nous allons maintenant présenter quelques éléments de théorie ergodique en mesure infinie ; pour simplifier, nous n'allons travailler ici qu'avec des systèmes en temps discret. Notre but dans cette partie est de présenter le comportement limite des sommes de Birkhoff de fonctions d'intégrale non nulle. L'essentiel du matériel est issu du livre de J. Aaronson [2].

Remarquons pour commencer que le théorème de Birkhoff suppose que le système dynamique que l'on étudie soit de mesure finie. Il n'est donc pas utilisable directement quand les mesures considérées sont infinies. En revanche, on dispose d'une généralisation du théorème de Birkhoff qui reste valable dans ce contexte : le théorème ergodique de Hopf [39, $ 14, Individueller Ergodensatz für Abbildungen] (voir [39, $ 14, Individueller Ergodensatz bei Strömungen] pour un résultat similaire en temps continu).

Théorème 1.4.1 (Théorème ergodique de Hopf)**.**
Soit (Ω, μ, T) *un système dynamique ergodique, conservatif, préservant la mesure* σ*-finie* μ*. Alors, pour toutes fonctions* f *et* g *dans* $\mathbb{L}^1(\Omega, \mu)$ *telles que* $\int_\Omega g \,\mathrm{d}\mu \neq 0$*, presque partout pour la mesure* μ*,*

$$\lim_{n\to+\infty} \frac{\sum_{k=0}^{n-1} f\circ T^k}{\sum_{k=0}^{n-1} g\circ T^k} = \frac{\int_\Omega f \,\mathrm{d}\mu}{\int_\Omega g \,\mathrm{d}\mu}.$$

Pour retrouver le théorème de Birkhoff à partir du théorème de Hopf, il suffit de choisir $g \equiv 1$. Le théorème ergodique de Hopf implique qu'en mesure infinie, la renormalisation utilisée dans le théorème de Birkhoff produit un résultat trivial.

Corollaire 1.4.2.
Soit (Ω, μ, T) *un système dynamique ergodique, conservatif, préservant la mesure* σ*-finie et infinie* μ*. Alors, pour toute fonction* $f \in \mathbb{L}^1(\Omega, \mu)$*, presque partout pour la mesure* μ*,*

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k = 0.$$

Démonstration.
Soit $M > 0$. Soit A un borélien tel que $\mu(A) > M$; la mesure μ étant infinie, un tel ensemble existe toujours. Alors, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$, pour tout $n > 0$ tel que les expressions ci-dessous soient bien définies,

$$\left|\frac{1}{n}\sum_{k=0}^{n-1} f\circ T^k\right| \leq \frac{\sum_{k=0}^{n-1} |f|\circ T^k}{\sum_{k=0}^{n-1} 1\circ T^k} \leq \frac{\sum_{k=0}^{n-1} |f|\circ T^k}{\sum_{k=0}^{n-1} 1_A\circ T^k} \to \frac{1}{\mu(A)}\int_\Omega |f| \,\mathrm{d}\mu \leq \frac{1}{M}\int_\Omega |f| \,\mathrm{d}\mu.$$

Cette majoration est vraie pour tout $M > 0$, d'où le résultat. □

L'étape suivante consiste à chercher une suite de renormalisation (a_n) mieux adaptée au processus, c'est-à-dire telle que $a_n^{-1}\sum_{k=0}^{n-1} f\circ T^k$ converge en un certain sens vers un objet non trivial. Une première possibilité est de choisir une fonction positive et intégrable g, et de poser $a_n = \sum_{k=0}^{n-1} g\circ T^k(x)$ pour un point x générique. Cependant, le choix du point x n'est pas évident, et une suite (a_n) ainsi construite peut être vue comme aléatoire. Or, on préférerait travailler avec une suite (a_n) facile à calculer et dont le comportement est simple. Avec des hypothèses très faibles sur une telle suite (a_n), la suite $a_n^{-1}\sum_{k=0}^{n-1} f\circ T^k$ ne peut pas converger presque sûrement vers un objet non trivial [2, théorème 2.4.1]. Pour pouvoir aller plus loin, il faut définir de nouveaux objets limites, formuler d'autres modes de convergence vers ces objets, et renforcer les hypothèses sur le système pour démontrer ce nouveau type de convergence.

Objets limites

Dans le cadre qui nous intéresse, les objets limites font intervenir une famille de lois de probabilité : les lois de Mittag-Leffler. Elles sont caractérisées par leurs fonctions génératrices.

Définition 1.4.3 (Lois de Mittag-Leffler).

Pour tout $\beta \in [0,1]$, on dit qu'une variable aléatoire Y_β à valeurs positives suit une loi de Mittag-Leffler standard de paramètre β si :

$$\mathbb{E}(e^{zY_\beta}) = \sum_{n=0}^{+\infty} \frac{\Gamma(1+\beta)^n z^n}{\Gamma(1+\beta n)},$$

où l'égalité est vraie pour tout complexe z si $\beta \in (0,1]$, et pour tout z dans le disque unité ouvert centré en 0 si $\beta = 0$.

D'après [10, théorème 30.1], les lois de Mittag-Leffler sont caractérisées par leurs moments. Une telle loi de Mittag-Leffler est dite *standard* car elle est normalisée pour que son premier moment soit de 1. Pour $\beta \in [0,1)$, la loi de Mittag-Leffler standard de paramètre β a pour support $\mathbb{R}_+$, et est absolument continue par rapport à la mesure de Lebesgue. En général, la densité ne s'exprime pas avec des fonctions usuelles, mais trois cas particuliers sont remarquables :

- Une loi de Mittag-Leffler standard de paramètre 0 est une loi exponentielle de paramètre 1, et a pour densité la fonction $t \mapsto e^{-t}$;
- Une loi de Mittag-Leffler standard de paramètre $1/2$ correspond à la valeur absolue d'une variable aléatoire gaussienne de variance $\pi/2$, et a pour densité la fonction $t \mapsto 2e^{-t^2/\pi}/\pi$;
- Une loi de Mittag-Leffler standard de paramètre 1 est une mesure de Dirac en 1.

Remarque 1.4.4.

Les lois de Mittag-Leffler sont appelées ainsi parce que leur fonction caractéristique s'exprime à l'aide de la fonction de Mittag-Leffler E, définie par :

$$E(\beta, z) := \sum_{n=0}^{+\infty} \frac{z^n}{\Gamma(1+\beta n)}.$$

Pour tout $\beta \in (0,1]$, il est aussi possible de définir une mesure de probabilité sur $\mathbb{R}_+$ en imposant que sa fonction de répartition soit $1 - E(\beta, -x^\beta)$. Ces lois de probabilités ainsi définies sont aussi appelées, dans un autre contexte, lois de Mittag-Leffler. Nous enjoignons le lecteur de ne pas confondre ces deux familles homonymes.

Les queues des lois de Mittag-Leffler peuvent se calculer rapidement.

Proposition 1.4.5.

Soit $\beta \in (0,1)$. Soit Y_β une variable aléatoire qui suit une loi de Mittag-Leffler standard de paramètre β. Alors, pour tout $C > 1$, pour tout t suffisamment grand,

$$\mathbb{P}(Y_\beta > t) \leq \frac{C}{\beta} e^{-C_\beta t^{\frac{1}{1-\beta}}},$$

où :

$$C_\beta = -\frac{\beta^{\frac{\beta}{1-\beta}} - \beta^{\frac{1}{1-\beta}}}{\Gamma(1+\beta)^{\frac{1}{1-\beta}}}.$$

Démonstration.

Il suffit d'utiliser une borne de Chernov. Soient $t \geq 0$ et $\lambda > 0$. Alors :

$$\mathbb{P}(Y_\beta > t) = \mathbb{P}\left(e^{\lambda Y_\beta} > e^{\lambda t}\right) \leq e^{-\lambda t} \mathbb{E}\left(e^{\lambda Y_\beta}\right) = e^{-\lambda t} E\left(\beta, \Gamma(1+\beta)\lambda\right).$$

D'après [27], pour tout $\beta > 0$, on a $E(\beta, z) \sim \beta^{-1} e^{z^{1/\beta}}$. Soit $C > 1$. Alors, pour tout λ suffisamment grand,

$$\mathbb{P}(Y_\beta > t) \leq \frac{C}{\beta} e^{\Gamma(1+\beta)^{1/\beta} \lambda^{1/\beta} - \lambda t}.$$

Afin d'optimiser cette majoration, on choisit :

$$\lambda := (\beta t)^{\frac{\beta}{1-\beta}} \Gamma(1+\beta)^{-\frac{1}{1-\beta}},$$

ce qui donne la majoration souhaitée pour tout t suffisamment grand. $\square$

Mode de convergence

On ne peut pas utiliser immédiatement la convergence en loi quand le système est muni d'une mesure infinie, car il n'y a pas de mesure de probabilité naturelle sur l'espace sous-jacent. Une solution possible est de démontrer la convergence pour toute mesure de probabilité raisonnable ; c'est la notion de convergence forte en loi, aussi appelée convergence forte en distribution.

Définition 1.4.6 (Convergence forte en loi).
Soit (Ω, μ) un espace polonais mesuré, et soit E un espace polonais. Soit $(X_n)_{n\geq 0}$ une suite de fonctions mesurables à valeurs dans E. Soit X une variable aléatoire à valeurs dans E. On dit que la suite $(X_n)_{n\geq 0}$ converge fortement en loi vers X si, pour toute mesure de probabilité $\nu \ll \mu$, la suite $(X_n)_{n\geq 0}$ vue comme suite de variables aléatoires sur (Ω, ν) converge en loi vers X. Autrement dit, si p_X est la loi de X, alors pour toute mesure de probabilité $\nu \ll \mu$,

$$X_{n*}\nu \to p_X \text{ pour la topologie faible.}$$

Limite en loi du temps local

Dans ce qui suit, (Ω, μ, T) est un système dynamique ergodique, conservatif et préservant la mesure, et f est une fonction intégrable sur Ω. On cherche une suite (a_n) qui soit à variation régulière, et telle que $a_n^{-1} \sum_{k=0}^{n-1} f \circ T^k$ converge fortement en distribution vers une certaine variable aléatoire.

Comme il a été mentionné dans l'introduction, sans hypothèse supplémentaire sur le système dynamique, celui-ci peut très bien n'avoir aucune bonne renormalisation avec ces propriétés. C'est le cas quand le système n'a pas de *loi des grands nombres généralisée* (voir [2, proposition 3.6.3]). Ce problème apparaît par exemple pour le flot horocyclique sur une surface périodique de période compacte et de courbure négative constante, si la mesure μ est l'une des mesures invariantes ergodiques qui n'est pas la mesure de Liouville [50]. Une autre possibilité est que la loi limite ne soit pas une loi de Mittag-Leffler, comme par exemple dans [5].

Notons T^* l'opérateur sur $\mathbb{L}^1(\Omega, \mu)$ défini comme dual de l'opérateur de composition par T agissant sur $\mathbb{L}^\infty(\Omega, \mu)$. La proposition qui suit est entre autres [2, corollaire 3.7.3].

Proposition 1.4.7.
Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant une mesure infinie. Supposons qu'il existe une suite $(a_n)_{n\geq 0}$ à variation régulière d'indice $\beta \in [0, 1]$ telle que, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$,

$$\frac{1}{a_n} \sum_{k=0}^{n-1} (T^*)^k f \to \int_\Omega f \, \mathrm{d}\mu \ \mu\text{-presque partout sur } \Omega.$$

Alors, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$,

$$\frac{1}{a_n} \sum_{k=0}^{n-1} f \circ T^k \to \int_\Omega f \, \mathrm{d}\mu \cdot Y_\beta \text{ fortement en loi,}$$

où Y_β est une variable aléatoire qui suit une loi de Mittag-Leffler standard de paramètre β.

Remarquons au passage que, si $\beta = 1$, d'après cette proposition et [2, Théorème 2.4.1], le processus $(a_n^{-1} \sum_{k=0}^{n-1} f \circ T^k)$ sous une loi $\nu \ll \mu$ est un rare exemple “naturel” de processus convergeant en probabilité, mais pas presque sûrement.

Certains systèmes dynamiques qui induisent une application Gibbs-Markov vérifient les hypothèses de cette proposition. Ce ne sont pas les seuls ; citons comme autre exemple les applications dites AFN (nommées ainsi par R. Zweimüller [73] [74]), qui sont des applications induites par une transformation dilatantes de l'intervalle ayant un point fixe neutre, et vérifiant quelques hypothèses supplémentaires raisonnables (telles que la condition d'Adler). La convergence presque partout peut alors être améliorée en convergence uniforme sur tout compact de l'intervalle, pourvu que la fonction f satisfasse une condition de régularité minimale (être intégrable au sens de Riemann suffit) [55, théorème 1.1].

Ces hypothèses additionnelles peuvent aussi s'exprimer en termes d'existence d'un système induit ayant certaines propriétés. Le résultat suivant décrit le comportement asymptotique du temps local pour des transformations induisant un système Gibbs-Markov. Pour un paramètre $\beta \in [0,1)$ et un système induit mélangeant, ce résultat peut se déduire de diverses propositions du livre de J. Aaronson [2]. La version ci-dessous est un peu plus générale. On note $\mathrm{sinc}(x) := \sin(x)/x$ le sinus cardinal.

Proposition 1.4.8.

Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant la mesure, qui induit un système Gibbs-Markov sur un sous-ensemble borélien A. Supposons que la variable aléatoire φ_A remplit la condition (1.3.3) pour un indice $\beta \in [0,1]$ et une fonction auxiliaire ψ.

Si $\beta \in [0,1)$, alors, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$,

$$\frac{1}{\mathrm{sinc}(\beta\pi)\psi(N)} \sum_{k=0}^{N-1} f \circ T^k \to \int f \, \mathrm{d}\mu \cdot Y_\beta \text{ fortement en loi,} \tag{1.4.1}$$

où Y_β suit une loi de Mittag-Leffler standard de paramètre β.

Si $\beta = 1$, alors, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$,

$$\frac{1}{N} \left(\sum_{i=0}^{N-1} \frac{1}{\psi(i)} \right) \sum_{k=0}^{N-1} f \circ T^k \to \int f \, \mathrm{d}\mu \text{ fortement en loi.} \tag{1.4.2}$$

Démonstration.

D'après [30, lemme 6.5], la famille analytique d'opérateurs de renouvellement $T(z)(v) = \mathcal{L}(z^{\varphi_A} v)$ agissant sur un espace de Banach bien choisi se comporte bien sur le disque unité ouvert centré en 0. Par conséquent, les hypothèses $(H1')$ et $(H2)$ de [55] sont satisfaites. Les conclusions de [55, théorème 3.6] sont donc valides. Il reste à appliquer la proposition 1.4.7. Dans le cas $\beta = 1$, la variable aléatoire Y_1 n'apparaît pas pour la simple raison qu'elle vaut 1 presque sûrement. □

Cette proposition est aussi un corollaire du travail que nous mènerons en partie 3.2, et en particulier du corollaire 3.2.13 appliqué aux résultats de la table 3.2.4 pour $d = 0$.

La proposition 1.4.8 relie la suite (a_n) utilisée au début de cette partie avec la fonction ψ. Si $\beta \in [0,1)$, alors $a_n \sim \mathrm{sinc}(\beta\pi)\psi(n)$. La renormalisation dans le cas $\beta = 1$ est plus compliquée, pour la même raison que nous avons dû donner une version distincte du théorème taubérien de Karamata dans ce cas (théorème 1.3.7).

Au passage, la proposition 1.4.8 justifie le but de la présente thèse. Si la fonction f est d'intégrale nulle, alors on n'obtient qu'une convergence forte en loi vers 0 : la limite est triviale.

1.5 Exemples fondamentaux

Nous allons enfin présenter quelques exemples de systèmes que notre travail nous permettra, dans une certaine mesure, de traiter. Dans les deux premiers exemples, une structure de tour de Rokhlin avec une base Gibbs-Markov est facile à mettre en évidence ; en particulier, le calcul de la hauteur de la tour ainsi que le contrôle de l'intégrabilité des observables se font assez facilement. Nous donnerons les résultats correspondants à la fin du second chapitre. Les autres exemples sont du type "marche aléatoire", et demanderont plus de travail. Nous donnerons des résultats partiels, dans le cas des vraies marches aléatoires, à la fin du second chapitre, mais les systèmes plus complexes ne seront traités que dans le troisième chapitre.

1.5.1 Une chaîne de Markov

L'exemple le plus simple que nous pourrons manipuler est celui d'une chaîne de Markov de type "échelle". Il s'agit d'une processus défini sur les entiers positifs. Quand le processus est en zéro, il est renvoyé à hauteur aléatoire ; sinon, il descend d'une unité. Plus précisément, on fixe un paramètre

$\alpha > 1$, et on se donne une chaîne de Markov (M_n) sur $\mathbb{N}$ dont la matrice de transition est donnée par :

$$\begin{cases} P_{0,n} & = & \frac{1}{\zeta(\alpha)n^\alpha} & \forall\, n \geq 1 \,; \\ P_{n,n-1} & = & 1 & \forall\, n \geq 1 \,; \\ P_{n,m} & = & 0 & \text{sinon.} \end{cases}$$

Cette chaîne de Markov est ergodique. Elle admet donc à multiplication près une unique mesure stationnaire sur $\mathbb{N}$, que l'on notera p. Sans perte de généralité, on peut supposer que $p(0) = 1$. On calcule alors :

$$\begin{cases} p(1) & = & 1 \,; \\ p(n) & = & p(n+1) + \frac{1}{\zeta(\alpha)n^\alpha} \ \forall\, n \geq 1, \end{cases}$$

ce qui permet par récurrence de calculer la suite $p(n)$: pour tout $n \geq 1$,

$$p(n) = \frac{1}{\zeta(\alpha)} \sum_{k=n}^{+\infty} \frac{1}{k^\alpha} \sim \frac{1}{(\alpha-1)\zeta(\alpha)} \frac{1}{n^{\alpha-1}}.$$

En particulier, si $\alpha \in (1, 2]$, alors la suite $p(n)$ n'est pas sommable : l'unique mesure invariante pour cette chaîne de Markov est infinie.

Comme toutes les chaînes de Markov, un tel système est exprimable dans le vocabulaire de théorie ergodique. Posons $\Omega := \{(a_n) \in \mathbb{N}^\mathbb{N} \ : \ P_{a_n, a_{n+1}} \neq 0 \ \forall n \geq 0\}$, et utilisons la topologie sur Ω induite par la topologie produit (c'est-à-dire la topologie des cylindres) sur $\mathbb{N}^\mathbb{N}$. Soit T le décalage sur Ω, défini pour tout $a \in \Omega$ et tout $n \geq 0$ par :

$$(Ta)_n := a_{n+1}.$$

Pour tout $n \geq 0$, soit μ_n la mesure de probabilité sur Ω définie comme la loi de la chaîne de Markov (M_n) sachant que $M_0 = n$. Posons $\mu := \sum_{n=0}^{+\infty} p(n)\mu_n$. Alors le système dynamique (Ω, μ, T) est ergodique, conservatif, et préserve la mesure. De plus, la mesure μ est infinie si et seulement si $\alpha \in (1, 2]$.

L'ensemble sur lequel on a envie d'induire ce système est la base. Posons $A := \{a \in \Omega \ : \ a_0 = 0\}$. Alors A peut être vu comme l'ensemble des excursions à partir de la base. De plus, par la propriété de Markov forte, ces excursions sont indépendantes et identiquement distribuées. L'ensemble A est donc isomorphe à $\{(0, n, n-1, \cdots, 1) \ : \ n \geq 1\}^\mathbb{N}$, et l'application induite est le décalage de Bernoulli sur ce dernier ensemble. On partitionne A en sous-ensembles selon la première excursion (qui est ici caractérisée par sa longueur). Grâce à l'indépendance, le fait que cette application induite est Gibbs-Markov est évident ; en particulier, la distorsion est nulle, donc n'importe quelle constante de dilatation $\lambda > 1$ convient. Le temps de premier retour est constant sur chaque élément de la partition, et sa loi est donnée par $\mathbb{P}(\varphi_A = n) = p(n-1)$.

Le système induit correspondant à une suite d'excursions indépendantes et identiquement distribuées, la méthode des articles [22] et [23] fonctionne, ce qui évite les complications présentées dans cette thèse.

On peut étudier, avec les mêmes outils, des matrices de transitions un peu plus générales, pour lesquelles la suite $(P_{n,n-1})$ est à variation régulière d'indice $\alpha \in [1, 2]$. Il suffit d'utiliser le théorème de Karamata pour obtenir les résultats annoncés ci-dessus.

1.5.2 Applications de Pomeau-Manneville

La classe des systèmes hyperboliques avec un point fixe (ou périodique) indifférent a été particulièrement étudiée ces dernières années. Même si la plupart des recherches ont porté sur des applications ayant une unique mesure de probabilité invariante et absolument continue par rapport à la mesure de Lebesgue, certains résultats restent valables en mesure infinie.

Les *applications de Pomeau-Manneville*, aussi appelées applications de *Liverani-Saussol-Vaienti*, représentent un classe d'exemples combinatoirement très simples, et qui servent à étudier les phénomènes qui apparaissent dans les systèmes hyperboliques ayant un point fixe neutre. Ce sont des transformations de l'intervalle $(0, 1]$ avec un point fixe neutre en 0. Nous renvoyons le lecteur aux articles de M. Thaler [67] [68] et de C. Liverani, B. Saussol et S. Vaienti [53] pour leurs propriétés élémentaires.

Définition et propriétés élémentaires

Posons $\Omega := (0,1]$, que l'on munit de la topologie usuelle. Pour tout $\alpha \geq 0$ et pour tout $x \in \Omega$, on pose :

$$T_\alpha x = \begin{cases} x(1+(2x)^\alpha) & \forall x \in (0,1/2] \ ; \\ 2x-1 & \forall x \in (1/2,1]. \end{cases} \tag{1.5.1}$$

La transformation T_0 est l'application dyadique, qui préserve la mesure de Lebesgue. Quand on augmente le paramètre α, le système $(\Omega, T_\alpha, \mathrm{Leb})$ reste ergodique, mais la mesure de Lebesgue n'est plus préservée. Il existe cependant toujours, à constante près, une unique mesure invariante par T_α, absolument continue par rapport à la mesure de Lebesgue, positive et ergodique. Sous celle-ci, le système est conservatif. La densité de cette mesure a un pôle d'ordre α en 0 [67], et peut être supposée continue (et même localement lipschitzienne sur $(0,1]$). Cette propriété reflète l'idée selon laquelle, quand α augmente, la première branche de T_α est de plus en plus proche de la première bissectrice, donc le processus met de plus en plus de temps à s'éloigner de l'origine, et donc passe de plus en plus de temps près de l'origine.

Si $\alpha \in (0,1)$, cette mesure est finie, et on peut sans perte de généralité supposer que c'est une mesure de probabilité. Si $\alpha \geq 1$, en revanche, cette mesure est infinie ; on la normalisera en imposant que l'intervalle $(1/2,1]$ soit de mesure 1. On notera μ_α la mesure obtenue, et h_α la densité continue correspondante.

Pour tout $\alpha \geq 0$, le système induit sur $A := (1/2,1]$ peut être muni d'une structure Gibbs-Markov. Pour cela, on définit par récurrence une suite $(x_n)_{n\geq 0}$ de la façon suivante : si $T_{\alpha,1}$ est la branche de T_α sur l'intervalle $(0,1/2]$, alors :

$$\begin{cases} x_0 & = & 1 \ ; \\ x_1 & = & 1/2 \ ; \\ x_{n+1} & = & T_{\alpha,1}^{-1}(x_n) \ \forall n \geq 0. \end{cases}$$

On pose $A_n := \{x \in A \ : \ T_\alpha x \in (x_{n+1}, x_n]\}$ pour tout $n \geq 0$. C'est vis-à-vis de cette partition que l'application induite en A est Gibbs-Markov, et de plus $\{\varphi_A = n\} = A_{n-1}$ pour tout $n \geq 1$. D'après [53],

$$\mu_\alpha(\varphi_A \geq n) \sim \frac{h_\alpha(1/2)}{4(\alpha n)^{\frac{1}{\alpha}}}.$$

En d'autres termes, la condition (1.3.3) est satisfaite, et l'indice de variation de la fonction ψ est $1/\alpha$.

Remarquons au passage que les applications de Pomeau-Manneville sont aussi des exemples d'applications AFN, qui ont été évoquées en partie 1.4.

Applications de Holland

La définition des applications de Pomeau-Manneville peut se généraliser, donnant plus de liberté quant au comportement de la transformation au voisinage du point fixe neutre. On peut alors obtenir des *applications de Holland* [38]. Initialement, ces applications étaient destinées à étudier la décroissance des corrélations en mesure finie dans le cas critique $\alpha = 1$, mais leurs définitions et propriétés élémentaires restent valable en mesure infinie. La présentation que nous en donnons est proche de celle de [32].

Posons $\Omega := (0,1]$, que l'on munit de la topologie usuelle. Soit $\alpha > 0$, et soit h une fonction définie sur $(0,1/2]$ et à variation régulière de paramètre $1+\alpha$ en 0. Supposons que h est positive, de classe $\mathcal{C}^2$, que $h(1/2) = 1/2$, que $h''(x) = O(x^{-2}h(x))$ en 0, et que $h' > 0$. On pose alors pour tout $x \in \Omega$:

$$T_h x = \begin{cases} x + h(x) & \forall x \in (0,1/2] \ ; \\ 2x-1 & \forall x \in (1/2,1]. \end{cases}$$

Une telle application est dite de Holland. C'est un cas particulier d'application AFN, pour lequel le système induit sur $(1/2,1]$ peut être muni d'une structure Gibbs-Markov. Les applications des Pomeau-Manneville se retrouvent en prenant $h(x) = 2^\alpha x^{1+\alpha}$.

Ces applications ont toujours, à constante près, une unique mesure invariante, ergodique et absolument continue par rapport à la mesure de Lebesgue. Si $\alpha > 1$, cette mesure est infinie ; si $\alpha = 1$, cela dépend de la fonction h choisie. En effet, si $h(x) = x(x/\ell(1/x))^\alpha$, les queues du temps de premier retour en $(1/2, 1]$ vérifient :

$$\mathbb{P}(\varphi_A > n) \sim \frac{C}{n^{1/\alpha}\ell^{\#}(n^{1/\alpha})},$$

où $\ell^{\#}$ est la *conjuguée de de Bruijn* de ℓ [11, partie 1.5.7]. On en déduit que, si $\alpha = 1$, les mesures invariantes absolument continues sont infinies si et seulement si $\int_0^{1/2} x/h(x) \, \mathrm{d}x = +\infty$.

L'un des intérêts de ces applications est que l'on peut obtenir toutes les queues possibles pour le temps de premier retour en A, sauf celles à variation lente. Dans la sous-partie 2.7.2, nous ne traiterons que le cas des applications de Pomeau-Manneville, mais le cas plus général des applications de Holland est similaire.

Point de vue physique : le chaos intermittent

Nous donnons ici un bref point de vue historique sur les applications de Pomeau-Manneville. Celles-ci ne sont pas dues à Y. Pomeau et P. Manneville. Sous leur forme actuelle, elles ont été définies dans l'article de C. Liverani, B. Saussol et S. Vaienti en 1999 [53]. Une famille proche de transformations a été étudiée par P. Gaspard et X.J. Wang [28], qui ont établi certaines propriétés combinatoires et procédé à une étude numérique.

Ces applications devaient, initialement, modéliser en partie le phénomène de turbulences intermittentes. Ce phénomène se manifeste par exemple avec le modèle de Lorenz :

$$\left\{ \begin{array}{rcl} \dot{x} & = & \sigma(y-z) \ ; \\ \dot{y} & = & -xz + rx - y \ ; \\ \dot{z} & = & xy - bz. \end{array} \right.$$

Prenons $\sigma = 10$ et $b = 8/3$. Pour r proche de 166, le système est stable et présente des oscillations périodiques. Pour des valeurs de r assez grandes, le système est chaotique. Le passage du régime stable au régime turbulent se fait *via* un régime de chaos intermittent : il existe des intervalles de temps pendant lesquels le système oscille de façon régulière, et ces intervalles sont entrecoupés de périodes de turbulence [58]. Notons r_T le paramètre critique (numériquement, $r_T \simeq 166,06$).

En prenant une section de Poincaré du flot, on se ramène à une application de la forme $T_\lambda(x) = h(x) + \lambda$ sur $[0, 1]$, où le paramètre λ dépend de r et h est convexe croissante, nulle en 0 et propre. Pour de petites valeurs de λ, l'application T_λ a au moins deux points fixes, dont un point fixe attractif, ce qui correspond à un régime stable. Si λ est grand, le graphe de T_λ est strictement au-dessus de la première bissectrice, ce qui laisse la possibilité d'un régime turbulent. Entre les deux, il existe une valeur de λ (correspondant à r_T) pour laquelle la première bissectrice est tangente au graphe de T_λ ; la transformation T_λ a alors un point fixe neutre répulsif. Les applications de Pomeau-Manneville modélisent le comportement de cette application critique, à la bifurcation entre régime stable et régime turbulent (dans ce modèle précis, il s'agit d'une bifurcation de type selle-nœud, aussi appelée de type I). Les périodes de stabilité correspondent à des intervalles de temps où le système est proche du point fixe neutre.

Au passage, cela souligne l'importance de l'étude des applications de Pomeau-Manneville en mesure infinie. En effet, si h est de classe $\mathcal{C}^2$, alors à la bifurcation, on a $T_\lambda(x - x_0) = (x - x_0) + h''(x_0)(x - x_0)^2 + o((x - x_0)^2)$, où x_0 est le point fixe de T_λ. Ainsi, génériquement dans $\mathcal{C}^2$, le comportement de T_λ à la bifurcation est celui de l'application de Pomeau-Manneville pour $\alpha = 1$, pour laquelle la mesure invariante intéressante est infinie.

Autres exemples

Il existe, dans le bestiaire mathématique, d'autres exemples de transformations hyperboliques ayant un point fixe neutre, et préservant à constante multiplicative près une unique mesure absolument continue par rapport à la mesure de Lebesgue. Citons-en quelques unes.

La *transformation de Boole* est l'application :

$$T \ : \ \left\{ \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x - \frac{1}{x} \end{array} \right. . \tag{1.5.2}$$

Cette application préserve la mesure de Lebesgue, vis-à-vis de laquelle elle est ergodique et conservative. On peut y penser comme une application ayant un point fixe neutre en $+\infty$. Pour rendre cette intuition rigoureuse, on peut conjuguer cette transformation avec la transformation de Möbius $z \mapsto (z-i)/(z+i)$ pour obtenir une transformation du cercle unité, qui aura alors deux branches hyperboliques et un point fixe neutre en 1 [2, chapitre 6]. On peut induire cette transformation par exemple sur $[-1, 1]$ ou sur $[0, 1]$. La condition (1.3.3) est satisfaite, avec [2, exercice 1.3.4] :

$$\psi(n) = \sqrt{\frac{n}{2}}.$$

Un autre exemple intéressant, cette fois-ci sur $\mathbb{R}_+^*$, est l'application T qui à x associe $1/x$ si $x \in (0, 1]$, et $x - 1$ si $x > 1$. Cette application préserve à constante multiplicative près une unique mesure invariante et absolument continue par rapport à la mesure de Lebesgue, qui est de plus infinie, ergodique et conservative. L'application induite sur l'intervalle $(0, 1]$ est l'application de Gauss $x \mapsto 1/x - \lfloor 1/x \rfloor$, d'une grande importance en théorie des nombres. Celle-ci n'est pas Gibbs-Markov, car elle est de dérivée égale à -1 en 1. Cependant, ce point neutre n'est pas fixe. Si l'on induit le système sur l'intervalle $(0, 1/2]$, on obtient bien une application Gibbs-Markov. L'induction sur $(0, 1]$ permet cependant de vérifier facilement que la condition (1.3.3) est satisfaite, avec :

$$\psi(n) = \frac{1}{n \ln(2)}.$$

Terminons par la *transformation de Farey*. Elle aussi est d'un grand intérêt en théorie des nombres : on peut relier les propriétés spectrales de l'opérateur de transfert de cette transformation au lieu des zéros de la fonction Zeta de Riemann, et les préimages successives de 0 forment une suite de Farey modifiée. Elle est définie sur $[0, 1]$ par $Tx = x/(1-x)$ si $x \in [0, 1/2]$, et $Tx = (1-x)/x$ si $x \in [1/2, 1]$. Le point 0 est un point fixe neutre ; pour cette application, la mesure $\mathrm{d}x/x$ est à constante multiplicative près l'unique mesure invariante et absolument continue par rapport à la mesure de Lebesgue, qui est de plus infinie, ergodique et conservative. On obtient une application Gibbs-Markov en induisant par exemple sur l'intervalle $A := [1/2, 1]$. Cette fois-ci, la condition (1.3.3) est vérifiée avec :

$$\psi(n) = \frac{1}{n \ln(2)}.$$

Notons au passage que, si T^{φ_A} est l'application induite en A par la transformation de Farey, alors $T^{\varphi_A + 1}$ est l'application de Gauss. De plus, il n'est pas étonnant que certaines des propriétés de la transformation de Farey soient proches des propriétés de la transformation précédente. Si l'on pose :

$$T'x = \left\{ \begin{array}{ll} x - 1 & \forall x \in (1, +\infty) \ ; \\ \frac{1}{x} - 1 & \forall x \in (0, 1], \end{array} \right.$$

alors $h \circ T' = T \circ h$, où $h(x) = 1/(1+x)$.

1.5.3 Marches aléatoires et $\mathbb{Z}^d$-extensions

Une autre grande classe de systèmes dynamiques munis d'une mesure ergodique infinie est composée de marches aléatoires et, plus généralement, de $\mathbb{Z}^d$-extensions de systèmes dynamiques munis d'une mesure de probabilité.

Soit $d \geq 0$, et soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires indépendantes, identiquement distribuées, et à valeurs dans $\mathbb{R}^d$. Alors le processus $(S_n)_{n \geq 0} = (\sum_{k=0}^{n-1} X_k)_{n \geq 0}$ est une *marche aléatoire* à valeurs dans $\mathbb{R}^d$ partant de 0. On appelle la loi de X_0 le *noyau de transition* de la marche aléatoire associée. Si X_0 est de variance finie, la loi forte des grands nombres et le théorème central limite décrivent le comportement asymptotique global de la loi de S_n. Plus généralement, si X_0 est dans le bassin d'attraction d'une loi stable de Lévy de paramètre α, alors il existe des suites (A_n) et (B_n) à variation régulière telles que $(S_n - A_n)/B_n$ converge vers une loi stable de Lévy de paramètre α (voir la sous-partie 1.3.4) ; si $\alpha \in (1, 2]$, alors $A_n = n\mathbb{E}(X_0)$.

Dans cette thèse, on se limitera aux marches aléatoires dont les sauts (X_n) sont à valeurs dans un sous-réseau de $\mathbb{R}^d$. Sans perte de généralité, on peut supposer que ces sauts sont à valeurs dans $\mathbb{Z}^d$ (quitte à changer la valeur de d). Cette hypothèse simplifiera l'analyse, et la rapprochera de modèles plus évolués que nous présenterons par la suite, tels le flot géodésique sur des surfaces périodiques.

Supposons que $\alpha \in (1, 2]$, que $d = 1$, et que les pas X_n sont d'espérance nulle et dans le bassin d'attraction d'une loi stable de Lévy de paramètres $(\alpha, \beta, 0, c)$. Alors il existe une suite $(B_n)_{n\geq 0}$ à variation régulière d'indice $1/\alpha$ telle que S_n/B_n converge vers une loi stable de Lévy de paramètres $(\alpha, \beta, 0, 1)$. Soit $p_{\alpha,\beta,0,1}$ la densité d'une telle loi par rapport à la mesure de Lebesgue sur $\mathbb{R}$. Au premier abord, pour de grandes valeurs de n, la loi de S_n est donnée par $B_n^{-1} p_{\alpha,\beta,0,1}(x/B_n)\, \mathrm{d}x$, d'où :

$$\mathbb{P}(S_n = 0) \approx \frac{p_{\alpha,\beta,0,1}(0)}{B_n}. \tag{1.5.3}$$

Cette heuristique est fausse, comme en témoigne la marche aléatoire simple sur $\mathbb{Z}$ partant de 0 : pour tout entier n impair, $S_n \neq 0$. Pour éviter ce problème, on va imposer une condition de non dégénérescence. Dans le cadre des marches aléatoires, on parle aussi d'*apériodicité forte*.

Définition 1.5.1 ((Non) dégénérescence d'une marche aléatoire)**.**
Soit $d \geq 0$, et soit $(S_n)_{n\geq 0}$ une marche aléatoire sur $\mathbb{Z}^d$ de sauts $(X_n)_{n\geq 0}$. La marche aléatoire est dite dégénérée si X_0 prend ses valeurs dans un translaté d'un sous-réseau strict de $\mathbb{Z}^d$, et non dégénérée sinon.

Cette définition de la dégénérescence recouvre deux problèmes distincts : le cas où $(S_n)_{n\geq 0}$ prend ses valeurs dans un sous-réseau strict de $\mathbb{Z}^d$ (auquel cas la marche aléatoire sur $\mathbb{Z}^d$ n'est pas ergodique), et le cas où la marche aléatoire prend ses valeurs dans tout $\mathbb{Z}^d$, mais a une période, comme la marche aléatoire simple sur $\mathbb{Z}$. Cette notion de période est liée à la période d'une application Gibbs-Markov ergodique mais non mélangeante.

Sous l'hypothèse de non dégénérescence, l'équation (1.5.3) peut être rendue rigoureuse (voir par exemple [41, théorème 4.5.1]).

Théorème 1.5.2.
Soit $(S_n)_{n\geq 0}$ une marche aléatoire sur $\mathbb{Z}$ de sauts $(X_n)_{n\geq 0}$. Supposons que X_0 est dans le bassin d'attraction d'une loi stable de Lévy de paramètre $(\alpha, \beta, 0, 1)$; soit (B_n) la suite de renormalisation associée. Supposons que la marche aléatoire n'est pas dégénérée. Alors :

$$\mathbb{P}(S_n = 0) \sim \frac{p_{\alpha,\beta,0,1}(0)}{B_n}.$$

Remarquons que, si $\alpha \in (1, 2]$, ces hypothèses contraignent $\mathbb{E}(X_0)$ à être nul.

Si la marche aléatoire est dégénérée et non identiquement nulle, alors quitte à prendre un sous-réseau, on peut supposer que la marche aléatoire est à valeurs dans $\mathbb{Z}$. De plus, si la période de la marche aléatoire est M, alors :

$$\frac{1}{M} \sum_{k=n}^{n+M-1} \mathbb{P}(S_k = 0) \sim \frac{p_{\alpha,\beta,0,1}(0)}{B_n}.$$

Dans tous les cas, si la série de terme général $1/B_n$ diverge, alors :

$$\mathbb{E}\left(\sum_{k=0}^{n-1} 1_{\{0\}}(S_k) \right) = \sum_{k=0}^{n-1} \mathbb{P}(S_k = 0) \sim p_{\alpha,\beta,0,1}(0) \sum_{k=0}^{n-1} \frac{1}{B_k}.$$

Le théorème de Karamata permet de décrire précisément le comportement de la série de terme général $1/B_n$. Ce théorème admet des variantes en dimension supérieure, même si la suite (B_n) ne sera pas la même.

Si la série de terme général $1/B_n$ converge, alors le temps passé en 0 est fini presque sûrement, et donc, par invariance par translation de la marche aléatoire, le système est transient. Réciproquement, on montre que la marche aléatoire est récurrente si et seulement si la suite de terme général $1/B_n$ diverge.

En particulier, en dimension 1, cela contraint le paramètre α à être dans l'intervalle $[1, 2]$; de plus, si $\alpha > 1$, la marche aléatoire est automatiquement récurrente. Si $\alpha = 1$, la situation dépend des queues de X_0. En dimension 2, si la marche aléatoire est récurrente alors $\langle w, X_0 \rangle$ appartient au bassin d'attraction d'une loi normale pour tout $w \in \mathbb{T}^d$. Cette dernière condition n'est pas suffisante ; là encore, la situation dépend des queues de $\langle w, X_0 \rangle$. En dimension supérieure ou égale à trois, une marche aléatoire non dégénérée n'est jamais récurrente.

Il reste à reformuler les marches aléatoires dans le contexte de la théorie ergodique en mesure infinie. Une première méthode, proche de la méthode de la sous-partie 1.5.1, consiste à voir la marche aléatoire comme un décalage unilatère sur $(\mathbb{Z}^d)^{\mathbb{N}}$. Il sera cependant plus prolifique de voir une marche aléatoire sur $\mathbb{Z}^d$ comme une $\mathbb{Z}^d$ extension d'un système dynamique hyperbolique. Pour cela, donnons nous un alphabet dénombrable $\mathcal{A}$ et une fonction $F : \mathcal{A} \to \mathbb{Z}^d$ (par exemple, $\mathcal{A} = \mathbb{Z}^d$ et F est l'identité). Soit T_0 le décalage sur $A := \mathcal{A}^{\mathbb{N}}$, et μ_0 une mesure sur A correspondant à une suite de variables aléatoires indépendantes et identiquement distribuées. Alors le processus (A, μ_0, T_0) peut être muni d'une structure Gibbs-Markov. Posons $\Omega := A \times \mathbb{Z}^d$, étendons la mesure μ_0 en une mesure μ sur Ω par $\mu(B \times q) = \mu_0(B)$ pour tout borélien $B \subset A$ et tout $q \in \mathbb{Z}^d$, et posons :

$$T \ : \ \begin{cases} \Omega & \to \quad \Omega \\ (x, q) = ((x_n)_{n \geq 0}, q) & \mapsto \quad (T_0 x, q + F(x_0)) \end{cases} .$$

Le vocabulaire des marches aléatoires est alors la traduction du vocabulaire des systèmes dynamiques. Par exemple, considérons une marche aléatoire (S_n) sur $\mathbb{Z}^d$, et un point $q \in \mathbb{Z}^d$. Le temps local de la marche aléatoire en q est le processus stochastique défini, pour tout entier $n \geq 0$, par :

$$\xi_n(q) := \sum_{k=1}^{n} 1_q(S_k) = \mathrm{Card}\{0 < k \leq n \ : \ S_n = q\}.$$

Soit $(A \times \mathbb{Z}^d, \mu, T)$ le système dynamique correspondant à la marche aléatoire (S_n). Posons $A_q := A \times \{q\}$. Alors le temps local de la marche aléatoire en q (au sens probabiliste) est conjugué au temps local du système dynamique en A_q (au sens de la théorie ergodique). La même traduction fonctionne pour le temps de premier retour, les temps de retour successifs, etc.

La construction ci-dessus se généralise pour donner les $\mathbb{Z}^d$ extension de processus Gibbs-Markov, qui peuvent se voir comme des "marches aléatoires" dont les sauts ne sont pas nécessairement indépendants, mais sont générés par un processus Gibbs-Markov sous-jacent.

Définition 1.5.3 ($\mathbb{Z}^d$ extensions de processus Gibbs-Markov)**.**
Soit $(A, \pi, d_A, \mathcal{B}, \mu_A, T_0)$ une application Gibbs-Markov ergodique. Soit d un entier positif, et soit F une fonction définie sur A, à valeurs dans $\mathbb{Z}^d$ et $\sigma(\pi)$-mesurable (c'est-à-dire presque sûrement constante sur chaque élément de la partition π).

La $\mathbb{Z}^d$ extension de base $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$ et de fonction de saut F est le système dynamique (Ω, μ, T) défini de la façon suivante :
- $\Omega = A \times \mathbb{Z}^d$ *;*
- $\mu(B \times q) = \mu_0(B)$ *pour tout borélien* $B \subset A$ *et tout* $q \in \mathbb{Z}^d$ *;*
- $T(x, q) = (T_0 x, q + F(x))$ *pour tout* $(x, q) \in \Omega$.

Une telle extension est dite dégénérée si F peut être écrite comme la somme d'un cobord et d'une fonction à valeurs dans un translaté d'un sous-réseau propre de $\mathbb{Z}^d$.

Si une $\mathbb{Z}^d$ extension d'un processus Gibbs-Markov est ergodique et conservative, alors le temps de premier retour φ_A en $A \times \{0\}$ est fini presque sûrement. De plus, pour tout $n \geq 0$, l'évènement "le temps de premier retour en $A \times \{0\}$ est inférieur ou égal à n" ne dépend que des n premier sauts, c'est-à-dire de $F(x_0)$, ..., $F(x_{n-1})$. La fonction F étant supposée $\sigma(\pi)$-mesurable, cet évènement ne dépend que de x_0, ..., x_{n-1}. Le temps de premier retour en $A \times \{0\}$ est donc un temps d'arrêt strictement positif et presque sûrement fini pour la filtration naturelle sur A. De plus, l'application induite en $A \times \{0\}$ préserve la mesure μ_A. D'après le lemme 1.2.6, on peut donc munir le système dynamique induit (A, μ_A, T_A) d'une structure Gibbs-Markov. La $\mathbb{Z}^d$ extension est alors un facteur de la tour de Rokhlin de base (A, μ_A, T_A) et de hauteur φ_A. Les observables de la $\mathbb{Z}^d$ extension se relèvent naturellement en des observables de cette tour de Rokhlin, dont nous devrons vérifier l'intégrabilité et la régularité. Les mesures de probabilité absolument continues sur la $\mathbb{Z}^d$ extension se relèvent aussi en mesures de probabilité absolument continues sur cette tour de Rokhlin : il suffit de relever les densités.

Si le système est une marche aléatoire, alors, de même qu'en sous-partie 1.5.1, par la propriété de Markov forte, les excursions sont indépendantes et identiquement distribuées. Le système induit (A, μ_A, T_A) est donc mélangeant.

Le cas des vraies marches aléatoires est bien compris, et sera traité au cours du chapitre 2. Pour les $\mathbb{Z}^d$ extensions d'applications Gibbs-Markov en général, même si le comportement du temps local est bien connu grâce entre autres aux travaux de J. Aaronson et de M. Denker [3] [4], la question de la régularité et de l'intégrabilité des observables est plus délicate. En fait, l'étude des $\mathbb{Z}^d$ extensions d'applications Gibbs-Markov ne sera pas faite explicitement, pour éviter la redondance avec la théorie des $\mathbb{Z}^d$ extensions de semi-flots au-dessus d'applications Gibbs-Markov du chapitre 3. Cependant, la théorie en temps continu est plus complexe, et recouvre notamment la théorie en temps discret. Nous laissons au lecteur intéressé le soin de déduire les théorèmes en temps discret des théorèmes en temps continu.

1.5.4 Flot géodésique

Soit M une variété riemannienne, de fibré tangent TM. Une *géodésique* est une courbe $x : I \to M$, où I est un intervalle réel, qui est de vitesse constante. En d'autres termes, le vecteur vitesse d'une géodésique est invariant par transport parallèle le long de la même géodésique, ou, plus formellement,

$$\nabla_{\dot{x}} \dot{x} = 0.$$

Ceci est une équation différentielle ordinaire du second ordre portant sur x. En particulier, la valeur de x en tout temps est déterminée par sa valeur et la valeur de sa dérivée en un temps donné. Autrement dit, on dispose d'un système dynamique qui agit sur TM de la façon suivante : à un point $(x(0), \dot{x}(0)) \in TM$ on associe sa géodésique $(x(t), \dot{x}(t)) \in TM$ qui est bien définie sur un petit voisinage de 0 dans $\mathbb{R}$. Pour tout $t \in \mathbb{R}$, soit g_t l'application qui, à une position et une vitesse d'une particule $(x(0), \dot{x}(0)) \in TM$ à l'instant 0, associe la position et la vitesse $(x(t), \dot{x}(t)) \in TM$ de cette même particule à l'instant t.

À t fixé, l'application g_t n'est pas forcément définie sur tout TM. Elle l'est si le flot géodésique est *complet* ; c'est le cas si et seulement si la variété M est complète. Alors $(g_t)_{t \in \mathbb{R}}$ est un groupe de difféomorphismes lisses de TM pour la composition.

Cependant, le transport parallèle préserve la norme des vecteurs vitesse, et donc le flot géodésique la préserve aussi. Pour tout $r > 0$, le sous-ensemble $\{(x, \dot{x}) \in TM : \|\dot{x}\| = r\}$ est préservé par le flot géodésique ; on dispose donc (si l'on ignore les courbes de vitesse nulle) d'un feuilletage trivial invariant de l'espace des géodésiques. Pour étudier les propriétés d'ergodicité du flot géodésique, on peut sans perte de généralité se limiter aux courbes de vitesse unitaire. Soit $T^1 M$ le fibré tangent unitaire de M. Alors le flot $(g_t)_{t \in \mathbb{R}}$ préserve $T^1 M$, ainsi que la *mesure de Liouville* μ_0 sur $T^1 M$. En d'autres termes, si M est une variété riemannienne complète, alors le système dynamique $(T^1 M, \mu_0, (g_t)_{t \in \mathbb{R}})$ préserve la mesure.

En général, ce système dynamique n'est pas ergodique ; par exemple, le flot géodésique sur le fibré tangent unitaire d'une sphère a une infinité de composantes ergodiques supportées par les grands cercles. On s'intéressera dans cette thèse aux flots hyperboliques sur des variétés de courbure sectionnelle strictement négative. Rappelons quelques faits fondamentaux [6].

Proposition 1.5.4.
Soit M une variété riemannienne connexe, compacte, et de courbure sectionnelle strictement négative. Alors, il existe un paramètre $\kappa > 0$ ayant la propriété suivante. En tout point $x \in T^1 M$, l'espace tangent en x est somme directe de trois sous-espaces $E_-(x)$, $E_+(x)$ et $E_0(x)$, tels que :

- *Ces trois sous-espaces sont préservés par les difféomorphismes g_t pour tout $t \in \mathbb{R}$;*

- *La direction $E_-(x)$ est contractée à vitesse exponentielle : pour tout $t \geq 0$, l'application linéaire $D_{g_t} : E_-(x) \to E_-(g_t(x))$ contracte d'un facteur au moins $e^{-\kappa t}$;*
- *La direction $E_+(x)$ est dilatée à vitesse exponentielle : pour tout $t \leq 0$, l'application linéaire $D_{g_t} : E_+(x) \to E_+(g_t(x))$ contracte d'un facteur au moins $e^{-\kappa|t|}$;*
- *L'espace $E_0(x)$ est de dimension 1, et engendré par l'évaluation en x du champ de vecteurs engendrant le flot géodésique sur T^1M.*

En d'autres termes, un tel flot géodésique est un flot d'Anosov. Il est de plus ergodique et mélangeant pour la mesure de Liouville.

Si la variété M est de volume fini, mais n'est pas compacte et est de courbure variable, alors la situation est plus compliquée. On se restreindra alors au flot géodésique sur des surfaces de courbure strictement négative et constante : dans ce cas, une telle surface peut être vue comme un quotient de $\mathrm{SL}_2(\mathbb{R})$ par un sous-groupe Fuchsien. On pourra aussi traiter le cas du flot géodésique sur des orbifolds, tels que la surface modulaire $\mathrm{SL}_2(\mathbb{R})/\mathrm{SL}_2(\mathbb{Z})$; de tels flots ne sont pas définis partout, mais presque partout pour la mesure de Liouville (il suffit d'éliminer les trajectoires passant par un point singulier).

Les flots que nous avons définis jusqu'à présent sont ergodiques, mais la mesure associée est finie. Pour obtenir des flots en mesure infinie, nous allons procéder comme dans la sous-partie précédente : nous allons considérer le flot géodésique sur une variété hyperbolique compacte comme un générateur de hasard, et en prendre une $\mathbb{Z}^d$ extension. On se restreindra aux extensions qui ont un “sens géométrique” ; cela revient à étudier le flot géodésique sur des $\mathbb{Z}^d$ *revêtements de variétés hyperboliques*. De tels objets ont été étudiés entre autres par M. Babillot, F.Ledrappier pour le flot géodésique, et par J. Aaronson, F. Ledrappier et O Sarig pour le flot horocyclique.

Définition 1.5.5 ($\mathbb{Z}^d$ revêtement d'une variété riemannienne)**.**
Soit $p : N \to M$ un revêtement d'une variété riemannienne, et soit $d \geq 0$. On dit que p est un $\mathbb{Z}^d$ revêtement de M s'il existe une action de $\mathbb{Z}^d$ par isométries sur N qui préservent la projection p, et si cette action est fidèle et transitive sur la fibre.

Soit M une variété riemannienne compacte de dimension 2 et de courbure strictement négative. Alors on peut générer tout $\mathbb{Z}^d$ revêtement de la façon suivante. Soit g le genre de M. Tout d'abord, on décompose M en $3g-3$ pantalons $(M_k)_{0 \leq k < 3g-3}$, en découpant le long de $2g-2$ géodésiques fermées $(c_n)_{0 \leq n < 2g-2}$. Chacune de ces géodésiques définit un bord sur deux composantes (potentiellement identiques). Notons c_n^+ et c_n^- ces deux bords. Ensuite, on se donne $\mathbb{Z}^d$ copies de chacune des composantes M_k, que l'on notera $(M_{k,q})_{q \in \mathbb{Z}^d}$; cela engendre $\mathbb{Z}^d$ copies de chacun des $2g-2$ bords $(c_{n,q}^\pm)_{q \in \mathbb{Z}^d}$. Pour tout $0 \leq n < 2g-2$, on se donne un élément $F(n) \in \mathbb{Z}^d$. Enfin, pour tout $0 \leq n < 2g-2$, on recolle les bords $c_{n,q}^+$ et $c_{n,q+F(n)}^-$ canoniquement. Tous les $\mathbb{Z}^d$ revêtements peuvent s'obtenir ainsi.

Le flot géodésique sur le fibré tangent unitaire d'un $\mathbb{Z}^d$ revêtement connexe d'une variété riemannienne compacte de courbure sectionnelle strictement négative a de nombreux points communs avec une marche aléatoire sur $\mathbb{Z}^d$ de sauts bornés. Par exemple, la mesure de Liouville est ergodique et conservative si et seulement si $d \leq 2$. Les propriétés plus fines seront étudiées en partie 3.4.

Chapitre 2

Étude en temps discret

Sommaire

Ce second chapitre est consacré aux résultats en temps discret. Étant donnés un système dynamique muni d'une mesure infinie (Ω, μ, T) induisant un système Gibbs-Markov avec un temps de premier retour dont les queues sont à variation régulière, et une fonction f régulière, d'intégrale nulle et suffisamment intégrable, nous allons trouver le comportement asymptotique distributionnel des sommes $\sum_{k=0}^{n-1} f \circ T^k$.

Le principe de la démonstration est le suivant. Soit f une fonction d'intégrale nulle sur (Ω, μ). Alors :

$$\sum_{k=0}^{n-1} f \circ T^k = \sum_{k=0}^{\tau_{\xi_n}-1} f \circ T^k + \sum_{k=\tau_{\xi_n}}^{n-1} f \circ T^k.$$

Si l'on impose de bonnes conditions à f, alors le dernier terme est négligeable, et alors, en regroupant les termes appartenant à la même excursion :

$$\sum_{k=0}^{n-1} f \circ T^k \approx \sum_{k=0}^{\tau_{\xi_n}-1} f \circ T^k = \sum_{k=0}^{\xi_n-1} X_f \circ T_A^k \approx \sigma(f)\sqrt{\xi_n}\mathcal{N} \approx \sigma(f)\sqrt{a_n Y_\beta}\mathcal{N},$$

pour peu que X_f soit suffisamment régulière et intégrable. Ce raisonnement souffre de nombreuses failles ; par exemple, même si un processus stochastique (par exemple un mouvement brownien renormalisé) converge en loi vers une gaussienne, le même processus évalué à des instants aléatoires n'a aucune raison a priori de converger vers la même loi.

Après avoir exposé les deux théorèmes principaux de ce chapitre, nous nous inspirerons des travaux de E. Csáki et A. Földes pour démontrer un résultat d'indépendance asymptotique entre la somme de Birkhoff pour le système induit ($\sum_{k=0}^{n-1} X_f \circ T_A^k$) et la suite des temps de passage (τ_n). C'est cette indépendance qui va sauver l'heuristique ci-dessus. Il restera un peu de travail en partie 2.3 pour utiliser l'indépendance asymptotique dans le but de montrer la convergence forte en loi des sommes de Birkhoff.

Nous discuterons dans les parties 2.4 à 2.6 un certain nombre de généralisations de nos deux théorèmes principaux. Dans l'ordre, nous affaiblirons la condition de mélange sur le système induit, nous affaiblirons la condition de régularité sur les observables, et nous étudierons le cas d'observables à valeurs dans des espaces de Hilbert.

Enfin, nous appliquerons dans la partie 2.7 nos théorèmes limites à deux exemples : les applications de Pomeau-Manneville, et les marches aléatoires.

Les parties 2.1 à 2.3 sont issues de l'article [69]. Les parties 2.4 à 2.6 sont issues de l'article [70]. Les exemples sont en partie tirés de l'article [69]. Nous avons cependant rajouté des travaux propres à cette thèse, comme une discussion des conditions de régularité dans le cadre des applications de Pomeau-Manneville, ainsi qu'une discussion de la condition d'intégrabilité dans le cadre des marches aléatoires.

2.1 Résultats principaux

Cette partie a pour but de présenter les deux théorèmes centraux de cette thèse. Le premier est un résultat d'indépendance asymptotique ; c'est un intermédiaire important dans la preuve du second théorème, qui résout partiellement la problématique de la thèse. La plupart des résultats dans la suite de la thèse sont des raffinements de ces deux théorèmes.

Le résultat d'indépendance asymptotique est formulé en termes de *couplages*. Rappelons quelques définitions et propriétés fondamentales de ceux-ci. Soient X et Y deux variables aléatoires à valeurs dans des espaces polonais $\mathcal{X}$ et $\mathcal{Y}$ respectivement. Un couplage entre X et Y est une variable aléatoire à valeurs dans $\mathcal{X} \times \mathcal{Y}$, dont la première marginale (la projection sur $\mathcal{X}$) a la même loi que X et la seconde marginale (la projection sur $\mathcal{Y}$) la même loi que Y.

On peut concaténer les couplages au sens suivant. Soient X, Y et Z trois variables aléatoires à valeurs dans des espaces polonais $\mathcal{X}$, $\mathcal{Y}$ et $\mathcal{Z}$ respectivement. Supposons que l'on dispose d'un couplage (X_1, Y_1) entre X et Y et d'un couplage (Y_2, Z_2) entre Y et Z. Alors on peut coupler simultanément X, Y et Z sur le même espace de façon compatible avec les deux couplages précédents, c'est-à-dire qu'il existe une variable aléatoire (X_3, Y_3, Z_3) telle que $(X_3, Y_3) = (X_1, Y_1)$ et $(Y_3, Z_3) = (Y_2, Z_2)$ en loi. Il suffit de poser, pour toutes fonctions mesurables bornées f et g définies sur $\mathcal{X}$ et $\mathcal{Z}$ respectivement et à valeurs réelles,

$$\mathbb{E}(f(X_3)g(Z_3)|Y_3) := \mathbb{E}(f(X_3)|Y_3)\mathbb{E}(g(Z_3)|Y_3),$$

où le membre de droite est bien défini par les égalités en loi souhaitées. De telles concaténations de couplages ne sont en général pas uniques ; nous nous satisferons de savoir qu'elles existent.

Dans ce qui suit, (A, π, d, μ_A, T_A) sera une application Gibbs-Markov ergodique ; on supposera même pour commencer que cette application est aussi mélangeante. Les fonctions X et φ définies sur A seront des variables aléatoires à valeurs dans $\mathbb{R}$ et $\mathbb{R}_+$ respectivement ; X sera une variable aléatoire vérifiant de bonnes conditions d'intégrabilité, et φ une variable aléatoire à queue lourde et à variation régulière. On pourra penser à φ comme étant la hauteur d'une tour de base (A, π, d, μ_A, T_A), ou de façon équivalente comme le temps de premier retour en A pour un système dynamique (Ω, μ, T) de mesure infinie. Plus précisément, on supposera que $X \in \mathbb{L}^p(A, \mu_A)$ pour un certain $p > 2$, et que φ vérifie la condition (1.3.2) ou la condition (1.3.3). On supposera aussi que X et φ sont des fonctions régulières, par exemple que $\mathbb{E}(D(X))$ et $\mathbb{E}(D(\varphi))$ sont tous deux finis ; cette dernière hypothèse sera par la suite affaiblie (voir la partie 2.5).

Théorème 2.1.1.
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soient X et φ des fonctions mesurables sur A à valeurs dans $\mathbb{R}$ et $\mathbb{R}_+$ respectivement.

On pose $(X_i, \varphi_i) := (X \circ T_A^i, \varphi \circ T_A^i)$. Soient $(\tilde{X}_i)_{i \in \mathbb{N}}$ et $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ des copies des processus $(X_i)_{i \in \mathbb{N}}$ et $(\varphi_i)_{i \in \mathbb{N}}$ respectivement, telles que $(\tilde{X}_i)_{i \in \mathbb{N}}$ et $(\tilde{\varphi}_i)_{i \in \mathbb{N}}$ soient mutuellement indépendantes.

Supposons que :
- *$\mathbb{E}(D(X))$ et $\mathbb{E}(D(\varphi))$ sont tous deux finis ;*
- *Il existe $p > 2$ tel que $X \in \mathbb{L}^p(A, \mu_A)$;*
- *φ vérifie la condition (1.3.2) pour un certain paramètre $\beta \in [0, 1)$ et une certaine fonction auxiliaire ψ.*

Alors il existe $r \in [0, 1)$ et un couplage entre $(X_i, \varphi_i)_{i \in \mathbb{N}}$ et $(\tilde{X}_i, \tilde{\varphi}_i)_{i \in \mathbb{N}}$ tel que presque sûrement, pour tout entier N suffisamment grand,

$$\left| \sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i \right| \leq N^{\frac{r}{2}},$$
$$\left| \sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i \right| \leq \psi^*(N^r).$$

Nous insistons sur le fait que ce premier théorème n'est valable que pour un paramètre $\beta \in [0, 1)$.

Supposons que φ vérifie l'hypothèse plus restrictive (1.3.3), et que X est d'espérance nulle et n'est pas un cobord. Alors, typiquement, la somme $\sum_{i=0}^{N-1} X_i$ est de l'ordre $\sqrt{N}$ (d'après la proposition 1.1.7, on dispose d'un théorème central limite pour la suite des sommes partielles sous ces hypothèses du théorème), alors que la somme $\sum_{i=0}^{N-1} \varphi_i$ est de l'ordre de $\psi^*(N)$. Par conséquent, ce théorème affirme que, pour un certain couplage, la distance entre le processus originel (X_i, φ_i) et le processus $(\tilde{X}_i, \tilde{\varphi}_i)$ est petite en comparaison de la valeur typique de ces deux processus.

Remarque 2.1.2 (Choix du paramètre r).
La situation la plus favorable pour le choix du paramètre r dans le théorème 2.1.1 est celle où X appartient à $\mathbb{L}^p$ pour tout p fini. Cette condition est garantie par exemple si X est induite par une fonction höldérienne à support compact pour l'application de Pomeau-Manneville ou l'application de Boole, ou par une fonction à support fini dans $\mathbb{Z}$ ou $\mathbb{Z}^2$ pour les marches aléatoires (voir les discussions dans les sous-parties 2.7.1 et 3.3.2). Alors nos techniques montrent que l'on peut choisir le paramètre r dans l'intervalle $((3 - \beta)/(7 - 5\beta), 1)$.

La meilleure borne que l'on peut espérer obtenir avec notre méthode est de $3/7$ pour $\beta = 0$. On peut aussi choisir de dissocier les paramètres contrôlant les différences $\left|\sum_{i=0}^{N-1} X_i - \sum_{i=0}^{N-1} \tilde{X}_i\right|$ et $\left|\sum_{i=0}^{N-1} \varphi_i - \sum_{i=0}^{N-1} \tilde{\varphi}_i\right|$, ce qui permet d'améliorer (dans une certaine mesure) le contrôle sur l'une de ces sommes en sacrifiant le contrôle sur l'autre somme.

Au passage, si $\beta = 1$, l'intervalle dans lequel nous pouvons choisir r est vide, ce qui montre que l'argument utilisé ne fonctionne plus.

Le théorème 2.1.1, bien qu'intéressant par lui-même, est une étape intermédiaire dans la dérivation de limites distributionnelles explicites pour les sommes de Birkhoff de fonctions d'intégrale nulle. On rappelle que, si (Ω, μ, T) est un système dynamique ergodique, conservatif et préservant la mesure, si la mesure μ est infinie, et si A est un borélien de mesure 1, alors on note ψ l'inverse des queues du temps de premier retour en A. Autrement dit, pour tout $t \geq 0$,

$$\psi(t) := \frac{1}{\mu_{|A}(\varphi_A > t)}.$$

Une telle fonction ψ est alors nécessairement croissante, càdlàg, bornée inférieurement par 1 et tendant vers l'infini. Nous pouvons à présent formuler notre second théorème principal :

Théorème 2.1.3.

Soit (Ω, μ, T) un système dynamique ergodique, conservatif et préservant la mesure. Supposons que μ est σ-finie et infinie, et qu'il existe un sous-ensemble borélien A de mesure 1 sur lequel (Ω, μ, T) induit un système Gibbs-Markov (A, π, d, μ_A, T_A). Supposons que la fonction ψ associée au temps de premier retour φ_A est à variation régulière d'indice $\beta \in [0,1)$.

Soit f une fonction mesurable définie sur Ω et à valeurs réelles. Supposons que :
- *Il existe $p > 2$ tel que $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$;*
- *$\int_\Omega f \, \mathrm{d}\mu = 0$;*
- *$\mathbb{E}(D(X_f))$ est fini.*

Alors :

$$\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \to \sigma(f)\sqrt{Y_\beta}\mathcal{N} \text{ fortement en loi,} \tag{2.1.1}$$

où Y_β suit une loi de Mittag-Leffler standard de paramètre β, où $\mathcal{N}$ suit une loi normale standard, où Y_β et $\mathcal{N}$ sont indépendants et où :

$$\sigma(f)^2 = \int_A X_f^2 \, \mathrm{d}\mu + 2 \sum_{i=1}^{+\infty} \int_A X_f \cdot X_f \circ T_A^i \, \mathrm{d}\mu. \tag{2.1.2}$$

De plus, $\sigma(f) = 0$ si et seulement si f est un cobord.

Remarquons que, contrairement au théorème 2.1.1, la régularité du temps de premier retour n'apparaît pas dans cet énoncé. Cette hypothèse est en fait implicite : dans la définition d'un système dynamique induisant un système Gibbs-Markov, il est supposé que le temps de premier retour est constant sur chaque élément de la partition. La condition de régularité des temps de premier retour est donc trivialement satisfaite.

Comme on le verra dans la remarque 2.3.4, le jeu d'hypothèses :
- Il existe $p > 2$ tel que $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$;
- $\int_\Omega f \, \mathrm{d}\mu = 0$;
- $\mathbb{E}(D(X_f))$ est fini,

peut être remplacé par le jeu d'hypothèses suivant :
- Il existe $p > 2$ tel que $\sup_{N \leq \varphi_A} \left| \sum_{k=0}^{N-1} f \circ T^k \right| \in \mathbb{L}^p(A, \mu_A)$;
- $\int_A X_f \, \mathrm{d}\mu_A = 0$;
- $\mathbb{E}(D(X_f))$ est fini.

Cette deuxième liste d'hypothèses est plus lourde, mais plus générale. Elle a l'avantage de couvrir un certain nombre d'exemples de fonctions f qui ne sont pas intégrables, et d'être stable si l'on ajoute un cobord borné à f.

Nous donnons aussi une version du théorème 2.1.3 dans le cas $\beta = 1$. Son énoncé est un peu différent, pour deux raisons. Premièrement, une variable aléatoire suivant une loi de Mittag-Leffler standard de paramètre 1 vaut 1 presque sûrement ; Y_1 disparaît donc de l'énoncé. Deuxièmement, la renormalisation n'est pas la même, car dans la proposition 1.4.8 on a dû séparer le cas $\beta \in [0,1)$ et le cas $\beta = 1$. Finalement, on obtient un théorème central limite non standard.

De plus, la démonstration du théorème 2.1.4 ne fait pas appel au théorème d'indépendance asymptotique 2.1.1, qui ne fonctionne que pour $\beta \in [0,1)$. Elle est donc de nature assez différente. La conséquence la plus intéressante est que l'on n'utilise pas explicitement la structure Gibbs-Markov du système induit, mais seulement des propriétés spectrales de l'opérateur de transfert du système induit. En particulier, le théorème 2.1.4 reste valable dans un contexte plus général, incluant entre autres les applications AFN.

Théorème 2.1.4.

Supposons que les hypothèses du théorème 2.1.3 sont vérifiées, excepté que le paramètre β vaut 1. Alors :

$$\sqrt{\frac{1}{N} \sum_{i=0}^{N-1} \frac{1}{\psi(i)}} \sum_{k=0}^{N-1} f \circ T^k \to \sigma(f)\mathcal{N} \text{ fortement en probabilité,} \tag{2.1.3}$$

où $\mathcal{N}$ est une variable aléatoire suivant une loi normale standard et où la variance $\sigma(f)^2$ est donnée par l'équation (2.1.2)*. De plus, $\sigma(f) = 0$ si et seulement si f est un cobord.*

Le théorème 2.1.1 sera démontré dans la partie 2.2. Les théorèmes 2.1.3 et 2.1.4 seront démontrés dans la partie 2.3.

2.1.1 Description de la loi limite

Avant de passer à la démonstration des théorèmes limites que nous venons d'énoncer, faisons quelques remarques sur les lois limites impliquées. Si le paramètre β vaut 1, la loi limite est une gaussienne. Il est plus difficile d'obtenir une densité explicite pour ces lois si $\beta < 1$, ne serait-ce que parce que les lois de Mittag-Leffler n'ont pas de densité qui s'expriment avec des fonctions usuelles. Quelques remarques sont cependant possibles. Dans cette sous-partie, pour tout $\beta \in [0, 1]$, on note X_β une variable aléatoire réelle, dont la loi est celle de $\sqrt{Y_\beta}\mathcal{N}$, où $\mathcal{N}$ une variable aléatoire de loi normale standard, Y_β une variable aléatoire de loi de Mittag-Leffler standard de paramètre β, et où $\mathcal{N}$ et Y_β sont indépendantes.

Proposition 2.1.5 (Densité de la loi limite)**.**
Soit $\beta \in [0, 1)$. Soit $p_\beta \in \mathbb{L}^1(\mathbb{R}_+)$ la densité d'une loi de Mittag-Leffler standard de paramètre β. Alors la densité de la loi de X_β est donnée par :

$$x \mapsto \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_+} e^{-\frac{x^2}{2y^2}} p_\beta(y^2)\, \mathrm{d}y. \tag{2.1.4}$$

Démonstration.
Soit t un réel. On va calculer $\mathbb{P}(X_\beta \geq t)$. La loi de X_β étant symétrique, on peut supposer sans perte de généralité que $t \geq 0$. Alors :

$$\begin{aligned}
\mathbb{P}(\sqrt{Y_\beta}\mathcal{N} \geq t) &= \int_{\mathbb{R}_+} \mathbb{P}(\sqrt{y}\mathcal{N} \geq t) p_\beta(y)\, \mathrm{d}y \\
&= \int_{\mathbb{R}_+} \int_{t/\sqrt{y}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, \mathrm{d}y \cdot p_\beta(y)\, \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}_+} \int_t^{+\infty} \frac{1}{\sqrt{y}} e^{-\frac{x^2}{y}}\, \mathrm{d}x \cdot p_\beta(y)\, \mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}} \int_t^{+\infty} \int_{\mathbb{R}_+} \frac{1}{\sqrt{y}} e^{-\frac{x^2}{2y}} p_\beta(y)\, \mathrm{d}y\, \mathrm{d}x.
\end{aligned}$$

La densité de la loi de X_β est donc donnée par :

$$\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}_+} \frac{1}{\sqrt{y}} e^{-\frac{x^2}{2y}} p_\beta(y)\, \mathrm{d}y = \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_+} e^{-\frac{x^2}{2y^2}} p_\beta(y^2)\, \mathrm{d}y. \qquad \square$$

Dans les cas où la loi de Y_β peut être explicitée, nous pouvons aller plus loin. Pour $\beta = 1/2$, nous trouvons comme densité de $X_{1/2}$:

$$x \mapsto \frac{1}{\pi}\sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_+} e^{-\left(\frac{x^2}{2y^2} + \frac{y^4}{\pi}\right)}\, \mathrm{d}y.$$

À constante près, cette formule avait déjà été remarquée par Dobrushin.

En utilisant conjointement les propositions 1.4.5 et 2.1.5, on peut montrer que, pour tout $\beta \in (0, 1)$, la densité de la loi de X_β a une version continue et bornée sur $\mathbb{R}$, infiniment différentiable sur $\mathbb{R}^*$, et dont toutes les dérivées sont à décroissance rapide en l'infini. Pour tout $\beta \in (0, 1)$, on peut de plus trouver des constantes strictement positives C, C' telles que, pour tout $x \geq 0$,

$$\mathbb{P}(|X_\beta| \geq t) \leq Ce^{-C'x^{\frac{2}{2-\beta}}}.$$

Le cas $\beta = 1$ est trivial (X_1 suit une loi normale standard), et le cas $\beta = 0$ peut se traiter explicitement :

Corollaire 2.1.6.
La densité de la loi de X_0 est donnée par :

$$x \mapsto \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|}. \tag{2.1.5}$$

En d'autres termes, X_0 suit une loi de Laplace centrée de paramètre $1/\sqrt{2}$.

Démonstration.
Soit x un réel. D'après l'équation (2.1.4), la densité que l'on cherche à calculer est :

$$\sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_+} e^{-\frac{x^2}{2y^2}-y^2}\, \mathrm{d}y = \sqrt{\frac{2}{\pi}} e^{-\sqrt{2}|x|} \int_{\mathbb{R}_+} e^{-\left(y-\frac{|x|}{\sqrt{2}y}\right)^2}\, \mathrm{d}y.$$

Mais, pour tout réel x, la transformation $y \mapsto y - |x|/(\sqrt{2}y)$ préserve la mesure de Lebesgue [13, équation (3), p. 780]. Par conséquent,

$$\sqrt{\frac{2}{\pi}} \int_{\mathbb{R}_+} e^{-\frac{x^2}{2y^2}-y^2}\, \mathrm{d}y = \sqrt{\frac{2}{\pi}} e^{-\sqrt{2}|x|} \int_{\mathbb{R}_+} e^{-y^2}\, \mathrm{d}y = \frac{1}{\sqrt{2}} e^{-\sqrt{2}|x|}. \qquad \square$$

Les moments de X_β sont plus faciles à calculer. En effet, pour tout entier positif n, par indépendance, on a $\mathbb{E}(X_\beta^n) = \mathbb{E}(Y_\beta^{n/2})\mathbb{E}(\mathcal{N}^n)$. Si n est impair, par parité de la gaussienne, on a $\mathbb{E}(\mathcal{N}^n) = 0$. On obtient donc :

Proposition 2.1.7 (Moments de la loi limite)**.**
Soit $\beta \in [0,1]$. Alors, pour tout entier positif n,

$$\begin{cases} \mathbb{E}(X_\beta^{2n+1}) & = & 0\ ; \\ \mathbb{E}(X_\beta^{2n}) & = & \frac{\Gamma(1+\beta)^n (2n)!}{2^n \Gamma(1+\beta n)}. \end{cases}$$

En particulier, pour tout entier positif n,

$$\begin{cases} \mathbb{E}(X_0^{2n+1}) & = & 0\ ; \\ \mathbb{E}(X_0^{2n}) & = & \frac{(2n)!}{2^n}. \end{cases}$$

Nous retrouvons bien le fait que X_0 suit une loi de Laplace centrée de paramètre $1/\sqrt{2}$. Finalement, nous pouvons aussi calculer la fonction caractéristique de la loi limite.

Proposition 2.1.8 (Fonction caractéristique de la loi limite)**.**
Soit $\beta \in (0,1]$. Alors, pour tout réel ξ,

$$\mathbb{E}(e^{i\xi X_\beta}) = \sum_{n=0}^{+\infty} \frac{1}{\Gamma(1+\beta n)} \left(-\frac{\Gamma(1+\beta)\xi^2}{2}\right)^n. \tag{2.1.6}$$

De plus, pour tout réel ξ,

$$\mathbb{E}(e^{i\xi X_0}) = \frac{1}{1+\frac{\xi^2}{2}}. \tag{2.1.7}$$

Démonstration.
Nous savons déjà que X_0 suit une loi de Laplace de paramètre $1/\sqrt{2}$, d'où l'équation (2.1.7). Soit $\beta \in (0,1]$. Soient Y_β et $\mathcal{N}$ deux variables aléatoires indépendantes, qui suivent une loi de Mittag-Leffler standard de paramètre β et une loi normale standard respectivement. Alors, pour tout réel ξ :

$$\mathbb{E}(e^{i\xi X_\beta}) = \mathbb{E}(e^{i\xi\sqrt{Y_\beta}\mathcal{N}}) = \mathbb{E}(\mathbb{E}(e^{i\xi\sqrt{Y_\beta}\mathcal{N}}|Y_\beta)) = \mathbb{E}(e^{-\frac{\xi^2}{2}Y_\beta}).$$

Or les lois de Mittag-Leffler sont définies par leur fonction caractéristique, qui pour $\beta \in (0,1]$ converge sur $\mathbb{C}$, d'où le résultat. $\square$

Si l'on dispose d'une formule close pour la fonction de Mittag-Leffler restreinte au paramètre β, alors on dispose d'une formule close pour la fonction caractéristique de X_β. Par exemple, la fonction caractéristique de $X_{1/2}$ a une expression relativement succincte. Soit erf la fonction d'erreur, définie pour tout réel x par :

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^{+\infty} e^{-y^2} \, \mathrm{d}y.$$

Alors :

$$\mathbb{E}(e^{i\xi X_{1/2}}) = e^{\left(\frac{\sqrt{\pi}\xi^2}{4}\right)^2} \left(1 - \mathrm{erf}\left(\frac{\sqrt{\pi}\xi^2}{4}\right)\right). \tag{2.1.8}$$

Par conséquent, $\mathbb{E}(e^{i\xi X_{1/2}}) \sim 4/(\pi\xi^2)$ en l'infini. De plus, par la formule (2.1.4), la densité de $X_{1/2}$ a une version $\mathcal{C}^\infty$ sur $\mathbb{R}^*$ et dont toutes les dérivées sont à décroissance rapide en l'infini. Tout comme la densité de la loi de Laplace, la densité de $X_{1/2}$ a un défaut de régularité en 0 : elle ne peut pas avoir de version de classe $\mathcal{C}^2$. En fait, pour tout $\beta \in [0,1)$, la fonction caractéristique de X_β décroît en ξ^{-2} en l'infini. Ce défaut de régularité est donc présent pour tout $\beta \in [0,1)$.

2.2 Couplage de Csáki - Földes

La première étape vers la preuve du théorème 2.1.3 est le théorème d'indépendance asymptotique 2.1.1. La preuve que nous en offrons est adaptée de deux articles de E. Csáki et A. Földes [22] [23]. Le couplage fourni est explicite. Nous commençons par en expliquer la construction dans le cadre plus simple de suites de variables aléatoires indépendantes et identiquement distribuées, avant d'aborder le cadre des applications Gibbs-Markov. Nous espérons que ce détour aidera le lecteur à saisir l'heuristique sous-jacente sans être distrait par des ornements techniques.

2.2.1 Méthode dans le cas indépendant et adaptation

Soit $(X_i, \varphi_i)_{i\geq 0}$ une suite de variables aléatoire indépendantes et identiquement distribuées, telles que $X_0 \in \mathbb{L}^p$ pour un certain $p > 2$ et la fonction $t \mapsto \mathbb{P}(\varphi_0 > t)$ est à variation régulière d'indice $-\beta$ pour un certain $\beta \in [0,1)$. De plus, on suppose que φ_0 est presque sûrement positive, et que X_0 est d'espérance nulle.

Soit $(\tilde{X}_i, \tilde{\varphi}_i)_{i\geq 0}$ un processus stochastique tel que $(\tilde{X}_i) = (X_i)$ en loi, $(\tilde{\varphi}_i) = (\varphi_i)$ en loi, et $(\tilde{X}_i)$ et $(\tilde{\varphi}_i)$ sont indépendants. On veut tout d'abord construire un couplage entre $(X_i, \varphi_i)_{i\geq 0}$ et $(\tilde{X}_i, \tilde{\varphi}_i)_{i\geq 0}$ vérifiant les conclusions du théorème 2.1.1.

Le couplage

Présentons l'heuristique sur laquelle repose ce théorème. Le processus $(\sum_{i=0}^{n-1} X_i)$ est une somme de variables aléatoires de moyenne nulle et vérifiant de fortes conditions d'intégrabilité. La valeur du processus est majoritairement due à l'accumulation d'un très grand nombre de petits pas, qui ne se compensent pas entièrement ; un tel processus converge en un certain sens vers un mouvement brownien. À l'inverse, la variable φ_0 a une queue lourde, et donc la valeur du processus $(\sum_{i=0}^{n-1} \varphi_i)$ dépend d'un petit nombre de très grands sauts. Ainsi, les valeurs de ces deux processus sont presque indépendantes ; par exemple, on peut changer les grandes valeurs de la suite (φ_i) de telle sorte à modifier beaucoup $(\sum_{i=0}^{n-1} \varphi_i)$, mais en changeant peu la valeur de $(\sum_{i=0}^{n-1} X_i)$.

Il reste à rendre ce raisonnement rigoureux. Nous offrons une version simplifiée du couplage construit dans les articles [22] et [23]. On se donne deux copies indépendantes de (X_i, φ_i), que l'on notera $(X_i^{(1)}, \varphi_i^{(1)})$ et $(X_i^{(2)}, \varphi_i^{(2)})$. On cherche à reconstruire un processus de même loi que (X_n, φ_n) en choisissant certains termes dans le processus numéroté (1), et d'autres termes dans le processus numéroté (2).

Pour cela, on fixe un paramètre $r \in (0,1)$, et on découpe l'ensemble des entiers naturels en blocs $(I_n)_{n\geq 0}$ consécutifs tels que $|I_n| = 2^n$. Soient $n \geq 0$ et $j \in \{1,2\}$. Pour tout $i \in I_n$, on dit que le couple $(X_i^{(j)}, \varphi_i^{(j)})$ est *grand* si $\varphi_i^{(j)} \geq \psi^*(2^{rn}) = \psi^*(|I_n|^r)$, et *petit* sinon. Ensuite, pour tout $n \geq 0$ et pour tout $i \in I_n$:

- Si $(X_i^{(1)}, \varphi_i^{(1)})$ ou $(X_i^{(2)}, \varphi_i^{(2)})$ est grand, on pose $(X_i', \varphi_i') := (X_i^{(1)}, \varphi_i^{(1)})$;
- Si $(X_i^{(1)}, \varphi_i^{(1)})$ et $(X_i^{(2)}, \varphi_i^{(2)})$ sont petits, on pose $(X_i', \varphi_i') := (X_i^{(2)}, \varphi_i^{(2)})$.

Le processus ainsi créé a la même loi que le processus initial :

Lemme 2.2.1. *Le processus (X_i', φ_i') a la même loi que (X_i, φ_i).*

Démonstration.

Soit i un entier naturel. La valeur de (X_i', φ_i') ne dépend que des valeurs de $(X_i^{(1)}, \varphi_i^{(1)})$ et de $(X_i^{(2)}, \varphi_i^{(2)})$. Le processus $(X_i', \varphi_i')_{i \in \mathbb{N}}$ est donc une suite de variables aléatoires indépendantes. Il reste à montrer que, pour tout entier i, la loi de (X_i', φ_i') est la même que la loi de (X_i, φ_i).

Soit i un entier naturel. Soit A_i (respectivement A_i', $A_i^{(1)}$, $A_i^{(2)}$) l'évènement "le couple (X_i, φ_i) est grand" (respectivement (X_i', φ_i'), $(X_i^{(1)}, \varphi_i^{(1)})$, $(X_i^{(2)}, \varphi_i^{(2)})$). Le couple (X_i', φ_i') est grand si et seulement si $(X_i^{(1)}, \varphi_i^{(1)})$ est grand. On a donc $\mathbb{P}(A_i') = \mathbb{P}(A_i^{(1)}) = \mathbb{P}(A_i)$, et la loi de (X_i', φ_i') conditionnée par A_i' est la même que la loi de (X_i, φ_i) conditionnée par A_i.

Si (X_i', φ_i') est petit, alors il y a deux possibilités. Ou bien $(X_i^{(2)}, \varphi_i^{(2)})$ est grand, auquel cas $(X_i', \varphi_i') = (X_i^{(1)}, \varphi_i^{(1)})$, ou bien $(X_i^{(2)}, \varphi_i^{(2)})$ est petit, auquel cas $(X_i', \varphi_i') = (X_i^{(2)}, \varphi_i^{(2)})$. La variable aléatoire (X_i', φ_i') conditionnée par $(A_i')^c$ vaut donc $(X_i^{(1)}, \varphi_i^{(1)})$ conditionnée par $A_i^{(1),c}$ avec probabilité $\mathbb{P}(A_i)$, et vaut $(X_i^{(2)}, \varphi_i^{(2)})$ conditionnée par $A_i^{(2),c}$ avec probabilité $1 - \mathbb{P}(A_i)$. Mais la loi de $(X_i^{(1)}, \varphi_i^{(1)})$ conditionnée par $A_i^{(1),c}$ et la loi de $(X_i^{(2)}, \varphi_i^{(2)})$ conditionnée par $A_i^{(2),c}$ sont identiques. La loi de (X_i', φ_i') conditionnée par $(A_i')^c$ est donc la même que la loi de (X_i, φ_i) conditionnée par A_i^c. Finalement, $(X_i', \varphi_i') = (X_i, \varphi_i)$ en loi. $\square$

Pour tout $i \in I_n$, la probabilité qu'un couple (X_i, φ_i) soit grand est de l'ordre de 2^{-rn}, et tend donc vers 0 quand n tend vers l'infini à vitesse exponentielle en n (mais polynomiale en i). La densité des grands couples tend donc vers 0 en probabilité. La plupart du temps, on aura donc $(X_i', \varphi_i') = (X_i^{(2)}, \varphi_i^{(2)})$. En particulier, si r est bien choisi alors $(\sum_{i=0}^{n-1} X_i') \approx (\sum_{i=0}^{n-1} X_i^{(2)})$. De plus, si r est bien choisi, la majeure partie du processus $(\sum_{i=0}^{n-1} \varphi_i')$ est due aux grandes valeurs de (X_i', φ_i') - mais alors, $(\sum_{i=0}^{n-1} \varphi_i') \approx (\sum_{i=0}^{n-1} \varphi_i^{(1)})$.

Posons $(\tilde{X}_i, \tilde{\varphi}_i) := (X_i^{(2)}, \varphi_i^{(1)})$. D'après l'argument ci-dessus, si r est bien choisi, alors le couplage entre (X_i', φ_i') et $(\tilde{X}_i, \tilde{\varphi}_i)$ vérifie les hypothèses du théorème.

Cet argument repose de façon cruciale sur l'indépendance du processus (X_i, φ_i). Sans cette indépendance, le processus (X_i', φ_i') que l'on a reconstruit n'a aucun raison d'avoir la même loi que (X_i, φ_i), et donc on n'obtient pas le couplage souhaité. Pour remédier à ce problème, on va utiliser les propriétés de décorrélation des applications Gibbs-Markov pour forcer les processus à être indépendants par blocs. Ensuite, on travaillera avec un couplage où l'on échangera non pas des couples de variables aléatoires, mais des blocs entiers. La notion que nous utiliserons est celle de *processus i.i.d. par morceaux.*

Adaptation du couplage

Soient q et ε dans $(0,1)$. Pour tous entier n, k strictement positifs et tels que $k < 2^{(1-q)n}$, posons $I_{n,k} := \{i \in \mathbb{N} \ : \ 2^n + k2^{qn} \leq i < 2^n + (k+1)2^{qn} - 2^{q\varepsilon n}\}$ et $J_{n,k} := \{i \in \mathbb{N} \ : \ 2^n + (k+1)2^{qn} - 2^{q\varepsilon n} \leq i < 2^n + (k+1)2^{qn}\}$. Si $k \geq 2^{(1-q)n}$, on pose $I_{n,k} = J_{n,k} = \varnothing$ par convention. Enfin, posons $I := \bigcup_{(n,k) \in \mathbb{N}^2} I_{n,k}$ et $J := \bigcup_{(n,k) \in \mathbb{N}^2} J_{n,k}$.

Définition 2.2.2 (Processus i.i.d. par morceaux)**.**

Soit $\mathbb{B}$ un espace de Banach. Une suite de variables aléatoires $(Y_i)_{i \in \mathbb{N}}$ à valeurs dans $\mathbb{B}$ est un processus i.i.d. par morceaux de paramètres q et ε si :

- *Les variables aléatoires $(Y_i)_{i \in I}$ sont identiquement distribuées ;*
- *$Y_i = 0$ pour tout $i \in J$;*
- *Pour tout $n > 0$, la suite de variables aléatoires $((Y_i)_{i \in I_{n,k}})_{0 \leq k < 2^{(1-q)n}}$ à valeurs dans $\mathbb{B}^{|I_{n,k}|}$ est identiquement distribuée ;*
- *Les variables aléatoires $((Y_i)_{i \in I_{n,k}})_{(k,n) \in \mathbb{N}^2}$ sont indépendantes.*

Soit (A, π, d, μ_A, T_A) un système dynamique Gibbs-Markov. Soient X et φ des observables sur A vérifiant les hypothèses habituelles. Dans les trois prochaines sous-parties, nous couplerons le processus $(X \circ T_A^i, \varphi \circ T_A^i)$ avec divers processus, y compris des processus i.i.d. par morceaux, et vérifierons que ces couplages ont de bonnes propriétés (par exemple, les sommes partielles d'intérêt doivent être asymptotiquement proches les unes des autres). Pour que chacun de ces couplages ait de bonnes propriétés, il faudra que certaines relations entre les paramètres β, p, r, q et ε soient satisfaites. À la fin de la démonstration du théorème 2.1.1, nous vérifierons que l'on peut choisir les paramètres r, q et ε dans ce but.

2.2.2 Décorrélation pour les systèmes Gibbs-Markov

A l'inverse du reste de cette thèse, on ne distingue pas dans cette sous-partie les variables aléatoires X et φ. En effet, les conditions d'intégrabilité sur les observables sont sans importance dans le lemme qui suit ; seules les conditions de régularité importent.

Lemme 2.2.3.
Soit (A, μ_A, T_A) une application Gibbs-Markov mélangeante. Soit $(\mathbb{B}, \|\cdot\|_{\mathbb{B}})$ un espace de Banach, et soit Y une fonction définie sur A à valeurs dans $\mathbb{B}$. Soient q et ε dans $(0,1)$.

Posons $Y_i := Y \circ T_A^i$. Soit (Y_i^*) un processus i.i.d. par morceaux de paramètres q et ε tel que, pour tous entier n et k, les suites de variables aléatoires $(Y_i)_{i \in I_{n,k}}$ et $(Y_i^*)_{i \in I_{n,k}}$ aient la même loi.

Si $\mathbb{E}(D(Y)) < +\infty$, alors il existe un couplage entre $(Y_i)_{i \in \mathbb{N}}$ et $(Y_i^*)_{i \in \mathbb{N}}$ tel que, presque sûrement,

$$\sum_{i \in I} \|Y_i - Y_i^*\|_{\mathbb{B}} < +\infty.$$

Démonstration.
On procède en quatre étapes. Dans la première, on contrôle la semi-norme lipschitzienne des itérés de Y. Dans un second temps, on introduit les mesures conditionnelles, puis on démontre un résultat de décorrélation rapide (à vitesse exponentielle) pour des observables qui sont des fonction indicatrices de cylindres. Enfin, on utilisera ces deux ingrédients dans les étapes trois et quatre pour construire le couplage désiré.

Première étape : Contrôle de la semi-norme lipschitzienne

Soit $\overline{a} := [\overline{a}_0, \cdots, \overline{a}_{|I_{n,k}|+2^{q\varepsilon n-1}-1}]$ un cylindre de longueur $2^{qn} - 2^{q\varepsilon n} + 2^{q\varepsilon n-1} = |I_{n,k}| + 2^{q\varepsilon n-1}$, et soit $x_{\overline{a}}$ un élément de $\overline{a}$. Alors, pour tout x dans $\overline{a}$ et pour tout entier $i < |I_{n,k}| + 2^{q\varepsilon n-1}$:

$$\begin{aligned}\left\|Y \circ T_A^i x - Y \circ T_A^i x_{\overline{a}}\right\|_{\mathbb{B}} &\leq |Y|_{\mathrm{Lip}(\overline{a}_i)} \, d(T_A^i x, T_A^i x_{\overline{a}}) \\ &\leq C \, |Y|_{\mathrm{Lip}(\overline{a}_i)} \, \lambda^{-(|I_{n,k}|+2^{q\varepsilon n-1}-i)}.\end{aligned}$$

Par conséquent :

$$\sum_{i=0}^{|I_{n,k}|-1} \left\|Y \circ T_A^i x - Y \circ T_A^i x_{\overline{a}}\right\|_{\mathbb{B}} \leq C \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\mathrm{Lip}(\overline{a}_i)} \, \lambda^{-(|I_{n,k}|-i)} \lambda^{-2^{q\varepsilon n-1}}. \tag{2.2.1}$$

Deuxième étape : Décorrélation pour des fonctions indicatrices

Soit y dans A. On définit une mesure de probabilité sur A par :

$$\tilde{\mu}_y^{(2^{qn})} := \sum_{T_A^{2^{qn}} x = y} g^{(2^{qn})}(x) \delta_x.$$

Pour toute fonction f mesurable et μ_A-intégrable, pour μ_A-presque tout y, l'intégrale de f par rapport à la mesure $\tilde{\mu}_y^{(2^{qn})}$ est l'espérance conditionnelle de $f(x)$ sachant que $T_A^{2^{qn}} x = y$. Pour tout cylindre $\overline{a}$, on a $\tilde{\mu}_y^{(2^{qn})}(\overline{a}) = \mathcal{L}^{2^{qn}} 1_{\overline{a}}(y)$.

Soit $\overline{a}$ un cylindre non vide de longueur $|I_{n,k}|+2^{q\varepsilon n-1}$. Soit x dans $T_A^{|I_{n,k}|+2^{q\varepsilon n-1}}\overline{a}$. Alors :

$$\mathcal{L}^{|I_{n,k}|+2^{q\varepsilon n-1}}1_{\overline{a}}(x) = g^{(|I_{n,k}|+2^{q\varepsilon n-1})}(\overline{a}x) \leq C\mu_A(\overline{a}),$$

où $\overline{a}x$ est l'unique point de $\overline{a}$ dont l'image par $T_A^{|I_{n,k}|+2^{q\varepsilon n-1}}$ est x. On utilise la propriété de distorsion bornée des applications Gibbs-Markov (lemme 1.1.2). Pour tous x et y dans $T_A^{|I_{n,k}|+2^{q\varepsilon n-1}}\overline{a}$,

$$\begin{aligned}\left|\mathcal{L}^{|I_{n,k}|+2^{q\varepsilon n-1}}1_{\overline{a}}(x) - \mathcal{L}^{|I_{n,k}|+2^{q\varepsilon n-1}}1_{\overline{a}}(y)\right| &= \left|g^{(|I_{n,k}|+2^{q\varepsilon n-1})}(\overline{a}x) - g^{(|I_{n,k}|+2^{q\varepsilon n-1})}(\overline{a}y)\right| \\ &\leq Cd(x,y)g^{(|I_{n,k}|+2^{q\varepsilon n-1})}(\overline{a}x) \\ &\leq Cd(x,y)\mu_A(\overline{a}).\end{aligned}$$

Par conséquent, la fonction $\mathcal{L}^{|I_{n,k}|+2^{q\varepsilon n-1}}1_{\overline{a}}$ appartient à Lip^∞, et à constante multiplicative près sa norme est bornée par $\mu_A(\overline{a})$. Étant donné que $\mathcal{L}^{2^{qn}}1_{\overline{a}} = \mathcal{L}^{2^{q\varepsilon n-1}}\mathcal{L}^{|I_{n,k}|+2^{q\varepsilon n-1}}1_{\overline{a}}$, d'après l'équation (1.1.5),

$$\left|\tilde{\mu}_y^{(2^{qn})}(\overline{a}) - \mu_A(\overline{a})\right| = \left|\mathcal{L}^{2^{qn}}1_{\overline{a}}(y) - \mu_A(\overline{a})\right| \leq C\mu_A(\overline{a})\rho^{2^{q\varepsilon n-1}}. \tag{2.2.2}$$

Troisième étape : Construction du couplage

Soient n et k des entiers naturels. On rappelle que, pour tout $i \in I_{n,k}$, on a $Y_i = Y \circ T_A^i$. Soit y dans A, et soit $(Y_{i,y})_{i\in I_{n,k}}$ le processus stochastique $(Y \circ T_A^i x)_{i\in I_{n,k}}$ quand la distribution de $T_A^{2^n+k2^{qn}}x$ est $\tilde{\mu}_y^{(2^{qn})}$. Informellement, il s'agit du processus $(Y \circ T_A^i x)_{i\in I_{n,k}}$ sachant que $T_A^{2^n+(k+1)2^{qn}}x = y$. Pour tout $y \in A$, on se donne une copie $(Y_{i,y}^*)$ de (Y_i).

On démontrera dans la quatrième étape qu'il existe une constante C, indépendante de n, telle que pour μ_A-presque tout y, il existe un couplage entre $(Y_{i,y})$ et $(Y_{i,y}^*)$ vérifiant l'inégalité suivante :

$$\mathbb{P}\left(\sum_{i\in I_{n,k}}\left\|Y_{i,y} - Y_{i,y}^*\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}}\right) \leq C\max\left\{\rho, \lambda^{-1}\right\}^{2^{q\varepsilon n-2}}. \tag{2.2.3}$$

Tout d'abord, supposons que cette propriété soit satisfaite. Posons $(Y_i^*) := \int_A (Y_{i,y}^*)\ \mathrm{d}\mu_A(y)$. Ce processus stochastique est naturellement couplé avec $(Y_i) = \int_A (Y_{i,y})\ \mathrm{d}\mu_A(y)$; de plus, la loi de $(Y_{i,y}^*)$ ne dépend pas de $y = T_A^{2^n+(k+1)2^{qn}}x$, et donc la variable aléatoire $(Y_i^*)_{i\in I_{n,k}}$ est indépendante du futur à partir du temps $2^n+(k+1)2^{qn}$, ou, plus formellement, est indépendante de la tribu $T_A^{-2^n-(k+1)2^{qn}}\mathcal{B}$.

D'après l'inéquation (2.2.3), et grâce à un théorème de sélection mesurable, on peut construire un couplage entre $(Y_i^*)_{i\in I_{n,k}}$ et $(Y_i)_{i\in I_{n,k}}$ tel que :

$$\mathbb{P}\left(\sum_{i\in I_{n,k}}\left\|Y_i - Y_i^*\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}}\right) \leq C\max\left\{\rho, \lambda^{-1}\right\}^{2^{q\varepsilon n-2}}. \tag{2.2.4}$$

On itère cette opération. Soit $((Y_i^*)_{i\in I_{m,k}})_{m\leq n,\ 0\leq k<2^{(1-q)m}}$ une suite de variables aléatoires indépendantes telle que les variables $((Y_i^*)_{i\in I_{m,k}})_{0\leq k<2^{(1-q)m}}$ soient de même loi que $(Y_i)_{i\in I_{m,0}}$. On construit un couplage entre la variable aléatoire $((Y_i^*)_{i\in I_{m,k}})_{m\leq n,\ 0\leq k<2^{(1-q)m}}$ et la variable aléatoire $((Y_i)_{i\in I_{m,k}})_{m\leq n,\ 0\leq k<2^{(1-q)m}}$ tel que, pour tous $m \leq n$:

$$\mathbb{P}\left(\sum_{k=0}^{2^{(1-q)m}-1}\sum_{i\in I_{m,k}}\left\|Y_i - Y_i^*\right\|_{\mathbb{B}} > C2^{-m}\right) \leq C2^{-m}. \tag{2.2.5}$$

Soit $(Y_i^*)_{i\in I}$ un processus tel que les variables $((Y_i^*)_{i\in I_{n,k}})_{(n,k)\in\mathbb{N}^2}$ soient indépendantes et, pour tout couple d'entiers (n,k), les variables aléatoires $(Y_i^*)_{i\in I_{n,k}}$ à valeurs dans $\mathbb{B}^{|I_{n,k}|}$ aient la même loi que $(Y_i)_{i\in I_{n,k}}$. Alors, par le théorème d'extension de Kolmogorov, il existe un couplage entre

$(Y \circ T_A^i)_{i \in I}$ et $(Y_i^*)_{i \in I}$ tel que, pour tout entier m, la borne (2.2.5) soit vérifiée. D'après le lemme de Borel-Cantelli, la somme $\sum_{i \in I} \|Y_i - Y_i^*\|_{\mathbb{B}}$ est finie μ_A-presque sûrement. Ceci conclut la preuve du lemme 2.2.3.

Quatrième étape : Un couplage proche de la diagonale

On va maintenant construire un couplage vérifiant l'inégalité (2.2.3). Soit n un entier, et soit $k < 2^{(1-q)n}$. On veut coupler les variables aléatoires $(Y_{i,y})_{i \in I_{n,k}}$ et $(Y_{i,y}^*)_{i \in I_{n,k}}$ de telle sorte qu'elle soient proches l'une de l'autre avec grande probabilité. Soit $\mathcal{X}$ une variable aléatoire à valeurs dans A et de loi μ_A ; soit $\mathcal{X}_y$ une variable aléatoire à valeurs dans A et de loi $\tilde{\mu}_y^{(2^{qn})}$. Alors on peut identifier $(Y \circ T_A^i(\mathcal{X}_y))_{i < |I_{n,k}|-1}$ avec $(Y_{i,y})_{i \in I_{n,k}}$ et $(Y \circ T_A^i(\mathcal{X}))_{i < |I_{n,k}|-1}$ avec $(Y_{i,y}^*)_{i \in I_{n,k}}$. En d'autres termes, $(Y_{i,y})_{i \in I_{n,k}}$ et $(Y_{i,y}^*)_{i \in I_{n,k}}$ sont la même fonction sur A, mais les mesures sur A correspondantes sont telles que la loi de $T_A^{2^n + k2^{qn}} x$ soit respectivement $\tilde{\mu}_y^{(2^{qn})}$ et μ_A.

D'après la première étape de cette démonstration, on dispose d'une bonne majoration de $\sum_{i=0}^{|I_{n,k}|-1} \left\|Y \circ T_A^i x - Y \circ T_A^i z\right\|_{\mathbb{B}}$ dès que x et z appartiennent au même cylindre de longueur $|I_{n,k}| + 2^{q\varepsilon n-1} = 2^{qn} - 2^{q\varepsilon n-1}$. Maintenant, notons que μ_A et $\tilde{\mu}_y^{(2^{qn})}$ donnent approximativement la même masse à chacun de ces cylindres.

D'après l'équation (2.2.2), il existe des constantes $C > 0$ et $\rho \in (0,1)$ telles que, pour tout cylindre $\overline{a}$ de longueur $|I_{n,k}| + 2^{q\varepsilon n-1}$,

$$\left|\tilde{\mu}_y^{(2^{qn})}(\overline{a}) - \mu_A(\overline{a})\right| \leq C\mu_A(\overline{a})\rho^{2^{q\varepsilon n-1}}.$$

Par conséquent, il existe un couplage entre $\mathcal{X}$ et $\mathcal{X}_y$ tel que, avec probabilité au moins $1 - C\rho^{2^{q\varepsilon n-1}}$, ces deux variables aléatoires prennent leur valeur dans le même cylindre de longueur $|I_{n,k}| + 2^{q\varepsilon n-1}$. Ce couplage entre $\mathcal{X}$ et $\mathcal{X}_y$ induit un couplage entre $(Y_{i,y})_{i \in I_{n,k}}$ et $(Y_{i,y}^*)_{i \in I_{n,k}}$. C'est ce couplage qui vérifie la majoration (2.2.3) ; démontrons-le.

D'après l'inégalité (2.2.1), il existe une constante $C_{(1)}$ telle que, dès que $\mathcal{X}$ et $\mathcal{X}_y$ prennent leur valeur dans le même cylindre $\overline{a}$ de longueur $|I_{n,k}| + 2^{q\varepsilon n-1}$,

$$\sum_{i \in I_{n,k}} \left\|Y_i - Y_{i,y}^*\right\|_{\mathbb{B}} \leq C_{(1)} \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\mathrm{Lip}(\overline{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2^{q\varepsilon n-1}}.$$

Afin d'éviter de manipuler des indices trop longs, pour tout évènement E, on désigne la fonction indicatrice de E par $1(E)$ (à la place de 1_E). Soit $\tau \in (\lambda^{-1}, 1)$. Soit γ une variable aléatoire à valeurs dans π et de loi $\mathbb{P}(\gamma = a) = \mu_A(a)$, de telle sorte que :

$$\mathbb{E}\left(|Y|_{\mathrm{Lip}(\gamma)}\right) = \mathbb{E}(D(Y)).$$

Alors, pour tout $\delta > 0$:

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i \in I_{n,k}} \left\|Y_{i,y} - Y_{i,y}^*\right\|_{\mathbb{B}} > \delta\right) \\
&\quad \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{\overline{a} \in \pi^{|I_{n,k}|}} \mu_A(\overline{a}) 1\left(C_{(1)} \sum_{i=0}^{|I_{n,k}|-1} |Y|_{\mathrm{Lip}(\overline{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2^{q\varepsilon n-1}} > \delta\right) \\
&\quad \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{\overline{a} \in \pi^{|I_{n,k}|}} \mu_A(\overline{a}) \sum_{i=0}^{|I_{n,k}|-1} 1\left(C_{(1)} |Y|_{\mathrm{Lip}(\overline{a}_i)} \lambda^{-(|I_{n,k}|-i)} \lambda^{-2^{q\varepsilon n-1}} > \frac{\tau^{|I_{n,k}|-i}}{1-\tau}\delta\right) \\
&\quad = C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \sum_{\overline{a} \in \pi^{|I_{n,k}|}} \mu_A(\overline{a}) 1\left(|Y|_{\mathrm{Lip}(\overline{a}_i)} > C_{(2)} (\lambda\tau)^{|I_{n,k}|-i} \lambda^{2^{q\varepsilon n-1}} \delta\right) \\
&\quad = C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}\left(|Y|_{\mathrm{Lip}(\gamma)} > C_{(2)} (\lambda\tau)^{|I_{n,k}|-i} \lambda^{2^{q\varepsilon n-1}} \delta\right). \qquad (2.2.6)
\end{aligned}$$

On applique l'inégalité de Markov :

$$
\begin{aligned}
\mathbb{P}&\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \delta\right) \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}\left(|Y|_{\mathrm{Lip}(\gamma)} > C_{(2)}\lambda^{2^{q\varepsilon n-1}}(\lambda\tau)^{|I_{n,k}|-i}\delta\right) \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \frac{\mathbb{E}(D(Y))}{C_{(2)}\lambda^{2^{q\varepsilon n-1}}(\lambda\tau)^{|I_{n,k}|-i}\delta} \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \frac{\mathbb{E}(D(Y))}{C_{(2)}\lambda^{2^{q\varepsilon n-1}}\delta} \sum_{i=0}^{|I_{n,k}|-1} (\lambda\tau)^{-(|I_{n,k}|-i)} \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \frac{\lambda\tau}{1-\lambda\tau}\frac{\mathbb{E}(D(Y))}{C_{(2)}\lambda^{2^{q\varepsilon n-1}}\delta}. \qquad (2.2.7)
\end{aligned}
$$

On choisit $\delta = \lambda^{-2^{q\varepsilon n-2}}$. Alors :

$$
\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}}\right) \leq C\max\left\{\rho, \lambda^{-1}\right\}^{2^{q\varepsilon n-2}}. \qquad (2.2.8)
$$

□

On posera par la suite $Y := (X, \varphi)$. Si X est à valeurs réelles, alors Y est à valeurs dans $\mathbb{B} := \mathbb{R}^2$, que l'on munira de la norme du supremum. On a alors $D(Y) = \max\{D(X), D(\varphi)\}$. En particulier, si $\mathbb{E}(D(X))$ et $\mathbb{E}(D(\varphi))$ sont tous deux finis, alors $\mathbb{E}(D(Y))$ l'est aussi. Cette construction s'adapte aisément si X est à valeurs dans un espace de Banach, ce dont on se servira en partie 2.6.

2.2.3 Contrôle des "trous"

Nous vérifions maintenant que ne pas tenir compte des variables aléatoires dont l'indice est dans J, c'est-à-dire dans les "trous" entre les blocs, ne change pas le comportement asymptotique des sommes partielles des suites $(\varphi_i) = (\varphi \circ T_A^i)$ et $(X_i) = (X \circ T_A^i)$.

La plupart de la croissance de la suite des sommes partielles de $(\varphi_i)_{i\in\mathbb{N}}$ est due à une petite proportion de grands sauts, qui ont peu de chances d'arriver dans J. Retirer les $(\varphi_i)_{i\in J}$ ne changera donc pas la vitesse de croissance des sommes partielles. Le contrôle de la suite des sommes partielles de $(X_i)_{i\in J}$ est un peu plus compliquée. Elle repose sur le fait que la densité de J est nulle, mais pour exploiter cela, on aura besoin d'introduire un peu d'indépendance entre les blocs de J.

On commence par étudier la suite $(1_{i\in J}\varphi_i)$.

Lemme 2.2.4.

Supposons que φ vérifie la condition (1.3.2). *Soit $r \in (1-(1-\varepsilon)q, 1)$. Presque sûrement, pour tout entier N suffisamment grand,*

$$
\sum_{\substack{i\leq N\\ i\in J}} \varphi_i \leq \psi^*(N^r). \qquad (2.2.9)
$$

Démonstration.

Soit $r' \in (1-(1-\varepsilon)q, r)$ et n un entier strictement positif. Dans le calcul suivant, on borne les variables aléatoires φ_i au seuil $\psi^*(2^{r'n})$, puis on applique l'inégalité de Markov :

$$\begin{aligned}
\mathbb{P}\left(\sum_{\substack{2^n \le i < 2^{n+1} \\ i \in J}} \varphi_i > \psi^*(2^{r'n})\right) &\le \mathbb{P}\left(\exists 2^n \le i < 2^{n+1},\ i \in J,\ \varphi_i > \psi^*(2^{r'n})\right) \\
&\quad + \mathbb{P}\left(\sum_{\substack{2^n \le i < 2^{n+1} \\ i \in J}} \min\{\varphi_i, \psi^*(2^{r'n})\} > \psi^*(2^{r'n})\right) \\
&\le 2^{(1-(1-\varepsilon)q)n}\mathbb{P}\left(\varphi > \psi^*(2^{r'n})\right) + 2^{(1-(1-\varepsilon)q)n}\frac{\mathbb{E}\left(\min\{\varphi, \psi^*(2^{r'n})\}\right)}{\psi^*(2^{r'n})} \\
&\le \frac{2^{(1-(1-\varepsilon)q)n}}{\psi(\psi^*(2^{r'n}))} + 2^{(1-(1-\varepsilon)q)n}\frac{\mathbb{E}\left(\min\{\varphi, \psi^*(2^{r'n})\}\right)}{\psi^*(2^{r'n})} \\
&\le 2^{(1-(1-\varepsilon)q-r')n} + 2^{(1-(1-\varepsilon)q)n}\frac{\mathbb{E}\left(\min\{\varphi, \psi^*(2^{r'n})\}\right)}{\psi^*(2^{r'n})}.
\end{aligned}$$

De plus, d'après le théorème de Karamata (théorème 1.3.4) :

$$\begin{aligned}
\mathbb{E}\left(\min\{\varphi, \psi^*(2^{r'n})\}\right) &\le \int_0^{\psi^*(2^{r'n})} \frac{1}{\psi(t)}\,\mathrm{d}t + \frac{\psi^*(2^{r'n})}{\psi(\psi^*(2^{r'n}))} \\
&\le C'\frac{\psi^*(2^{r'n})}{\psi(\psi^*(2^{r'n}))} \\
&\le C' 2^{-r'n}\psi^*(2^{r'n}).
\end{aligned}$$

En mettant ensemble ces deux dernières inégalités, on a montré que, pour tout n :

$$\mathbb{P}\left(\sum_{\substack{2^n \le i < 2^{n+1} \\ i \in J}} \varphi_i > \psi^*(2^{r'n})\right) \le C' 2^{-[r'-(1-(1-\varepsilon)q)n]}.$$

Le membre de droite est sommable. D'après le lemme de Borel-Cantelli, presque sûrement, pour tout entier n suffisamment grand,

$$\sum_{\substack{2^n \le i < 2^{n+1} \\ i \in J}} \varphi_i \le \psi^*(2^{r'n}).$$

On additionne maintenant ces inégalités. Presque sûrement, il existe une constante C telle que, pour tout entier N suffisamment grand,

$$\sum_{\substack{i \le N \\ i \in J}} \varphi_i \le \sum_{\substack{i < 2^{\lfloor \log_2 N \rfloor + 1} \\ i \in J}} \varphi_i \le C \sum_{k=0}^{\lfloor \log_2 N \rfloor} \psi^*(2^{r'k}) \le C(\lfloor \log_2 N \rfloor + 1)\psi^*(2^{r'\lfloor \log_2 N \rfloor}).$$

D'après le lemme 1.3.9, il existe presque sûrement une constante C telle que, pour tout entier N suffisamment grand,

$$\sum_{\substack{i \le N \\ i \in J}} \varphi_i \le C(\lfloor \log_2 N \rfloor + 1)\psi^*(2^{r'\lfloor \log_2 N \rfloor}) \le \psi^*(2^{r\lfloor \log_2 N \rfloor}) \le \psi^*(N^r). \qquad \square$$

Nous sommes parvenus à contrôler les sommes partielles de la suite $(1_{i\in J}\varphi \circ T_A^i)$. Nous allons maintenant démontrer le même type de majoration pour les sommes partielles de la suite $(1_{i\in J} X \circ T_A^i)$.

Lemme 2.2.5.

Soit $p > 2$, et soit $X \in \mathbb{L}^p(A, \mu_A)$. Soit $r \in (1-(1-\varepsilon)q, 1)$. Presque sûrement, pour tout entier N suffisamment grand,

$$\left| \sum_{\substack{i\leq N \\ i\in J}} X_i \right| \leq N^{\frac{r}{2}}. \tag{2.2.10}$$

Démonstration.

Soit n un entier naturel. Soit $((X_i^*)_{i\in J_{n,k}})_{0\leq k<2^{(1-q)n}}$ une suite de variables aléatoires telle que les $(X_i^*)_{i\in J_{n,k}}$ soient indépendants et aient la même loi que $(X_i)_{i\in J_{n,0}}$. La preuve du lemme 2.2.3 peut être adaptée pour montrer qu'il existe une constante $C > 0$ et un couplage entre (X_i) and (X_i^*) tels que :

$$\mathbb{P}\left(\sum_{k=0}^{2^{(1-q)n}-1} \sum_{i\in J_{n,k}} |X_i - X_i^*| > C2^{-n} \right) \leq C2^{-n}.$$

Par le lemme de Borel-Cantelli, la somme des différences $\sum_{i\in J} |X_i - X_i^*|$ est presque sûrement bornée, donc il suffit de démontrer le lemme 2.2.5 en remplaçant (X_i) par (X_i^*).

Soient $i < j$ des entiers ; pour commencer, on veut une majoration de l'espérance de $\left|\sum_{\substack{\ell\in J\\ i\leq\ell<j}} X_\ell^*\right|^p$. Remarquons que $\{\ell \in J :\ i \leq \ell < j\}$ est une suite de blocs B tels que les variables aléatoires $(X_\ell^*)_{\ell\in B}$ sont indépendantes : ces blocs B sont les ensembles $\{\ell \in J_{n,k} :\ i \leq \ell < j\}$ pour n et k dans $\mathbb{N}$. Soit $\mathbb{B}$ l'ensemble de tous les blocs non vides dans $\{\ell \in J :\ i \leq \ell < j\}$. D'après l'inégalité de Rosenthal pour les suites de variables aléatoires indépendantes [60, théorème 3], il existe une constante C ne dépendant que de p et telle que :

$$\mathbb{E}\left(\left| \sum_{\substack{\ell\in J \\ i\leq\ell<j}} X_\ell^* \right|^p \right) \leq C \left[\left(\sum_{B\in\mathbb{B}} \mathbb{E}\left(\left| \sum_{\ell\in B} X_\ell^* \right|^2 \right) \right)^{\frac{p}{2}} + \sum_{B\in\mathbb{B}} \mathbb{E}\left(\left| \sum_{\ell\in B} X_\ell^* \right|^p \right) \right]. \tag{2.2.11}$$

On utilise maintenant une autre version de l'inégalité de Rosenthal, la proposition 1.1.8. Il existe une constante C telle que, pour tous les blocs B, quelque soit leur longueur,

$$\mathbb{E}\left(\left| \sum_{\ell\in B} X_\ell^* \right|^2 \right) \leq C|B|,$$

et :

$$\mathbb{E}\left(\left| \sum_{\ell\in B} X_\ell^* \right|^p \right) \leq C|B|^{\frac{p}{2}}.$$

Par conséquent, l'inégalité (2.2.11) devient :

$$\begin{aligned} \mathbb{E}\left(\left| \sum_{\substack{\ell\in J \\ i\leq\ell<j}} X_\ell^* \right|^p \right) &\leq C \left[\left(\sum_{B\in\mathbb{B}} |B| \right)^{\frac{p}{2}} + \sum_{B\in\mathbb{B}} |B|^{\frac{p}{2}} \right] \\ &\leq C \left(\sum_{B\in\mathbb{B}} |B| \right)^{\frac{p}{2}} = C\,\mathrm{Card}\{\ell \in J :\ i \leq \ell < j\}^{\frac{p}{2}}. \end{aligned}$$

Un résultat de R.J. Serfling [62, corollaire B1] permet de contrôler les moments du maximum de la norme des sommes partielles si l'on dispose seulement d'un contrôle sur les moments de la norme des sommes partielles (il s'agit d'une généralisation de l'inégalité de Kolmogorov). Il existe donc une constante C telle que :

$$\mathbb{E}\left(\sup_{i\leq N}\left|\sum_{\substack{\ell\in J\\ 0\leq \ell<i}} X^*_\ell\right|^p\right) \leq C\,\mathrm{Card}\{\ell\in J:\ \ell\leq N\}^{\frac{p}{2}} \leq CN^{\frac{(1-(1-\varepsilon)q)p}{2}}.$$

Soit $r'\in(1-(1-\varepsilon)q, r)$. D'après l'inégalité de Markov, pour tout entier n :

$$\mathbb{P}\left(\sup_{\substack{i\in J\\ i\leq 2^n}}\left|\sum_{\substack{\ell\in J\\ 0\leq \ell<i}} X^*_\ell\right| > 2^{\frac{r'n}{2}}\right) \leq C\frac{2^{\frac{(1-(1-\varepsilon)q)pn}{2}}}{2^{\frac{pr'n}{2}}}.$$

Le membre de droite est sommable en tant que fonction de n. Par le lemme de Borel-Cantelli, presque sûrement, pour tout entier N suffisamment grand,

$$\left|\sum_{\substack{i\leq N\\ i\in J}} X^*_i\right| \leq \sup_{\substack{i\in J\\ i\leq 2^{\lfloor \log_2 N\rfloor+1}}}\left|\sum_{\substack{\ell\in J\\ 0\leq \ell<i}} X^*_\ell\right| \leq C2^{\frac{r'\lfloor \log_2 N\rfloor}{2}} \leq CN^{\frac{r'}{2}} \leq N^{\frac{r}{2}}. \qquad \square$$

2.2.4 Le couplage

Nous allons à présent construire un couplage inspiré du couplage de la sous-partie 2.2.1, mais adapté aux processus i.i.d. par morceaux. Le but est de démontrer la proposition suivante :

Proposition 2.2.6.
Soit $(X^*_i,\varphi^*_i)_{i\in\mathbb{N}}$ un processus i.i.d. par morceaux de paramètres q et ε. Soient $(\overline{X}_i)_{i\in\mathbb{N}}$ et $(\overline{\varphi}_i)_{i\in\mathbb{N}}$ deux processus indépendants, tels que $(\overline{X}_i)_{i\in\mathbb{N}}$ soit une copie de $(X^*_i)_{i\in\mathbb{N}}$ et que $(\overline{\varphi}_i)_{i\in\mathbb{N}}$ soit une copie de $(\varphi^*_i)_{i\in\mathbb{N}}$.

Supposons qu'il existe un réel $p>2$ et une constante C tels que, pour tout entier n,

$$\mathbb{E}\left(\sup_{i<2^{qn}}\left|\sum_{\ell=2^n}^{2^n+i} X^*_\ell\right|^p\right) \leq C2^{\frac{pqn}{2}}. \tag{2.2.12}$$

Supposons de plus que φ vérifie la condition (1.3.2). *Alors il existe un réel $r\in(0,1)$ et un couplage entre $(X^*_i,\varphi^*_i)_{i\in\mathbb{N}}$ et $(\overline{X}_i,\overline{\varphi}_i)_{i\in\mathbb{N}}$ tels que, presque sûrement, pour tout entier N suffisamment grand,*

$$\left|\sum_{i=0}^{N-1} X^*_i - \sum_{i=0}^{N-1}\overline{X}_i\right| \leq N^{\frac{r}{2}}, \tag{2.2.13}$$
$$\left|\sum_{i=0}^{N-1} \varphi^*_i - \sum_{i=0}^{N-1}\overline{\varphi}_i\right| \leq \psi^*(N^r).$$

Nous présentons maintenant le couplage annoncé. Soit $r\in(0,1)$. Soient $(X^{(1)}_i,\varphi^{(1)}_i)$ et $(X^{(2)}_i,\varphi^{(2)}_i)$ deux copies indépendantes du processus (X^*_i,φ^*_i). On dit qu'un bloc $(X^{(j)}_i,\varphi^{(j)}_i)_{i\in I_{n,k}}$ est *grand* s'il existe un entier $i\in I_{n,k}$ tel que $\varphi^{(j)}_i\geq\psi^*(2^{rn})$. Dans le cas contraire, un bloc est dit *petit*. On définit un processus (X'_i,φ'_i) de la façon suivante. Pour tout couple d'entier positifs (n,k),

- Si le bloc $(X^{(1)}_i,\varphi^{(1)}_i)_{i\in I_{n,k}}$ ou le bloc $(X^{(2)}_i,\varphi^{(2)}_i)_{i\in I_{n,k}}$ est grand, on pose $(X'_i,\varphi'_i)_{i\in I_{n,k}} := (X^{(1)}_i,\varphi^{(1)}_i)_{i\in I_{n,k}}$;
- Si les deux blocs $(X^{(1)}_i,\varphi^{(1)}_i)_{i\in I_{n,k}}$ et $(X^{(2)}_i,\varphi^{(2)}_i)_{i\in I_{n,k}}$ sont petits, on pose $(X'_i,\varphi'_i)_{i\in I_{n,k}} := (X^{(2)}_i,\varphi^{(2)}_i)_{i\in I_{n,k}}$.

Le processus (X_i', φ_i') est alors i.i.d. par morceaux de paramètres q et ε. Un bloc $(X_i', \varphi_i')_{i\in I_{n,k}}$ est grand si et seulement si le bloc $(X_i^{(1)}, \varphi_i^{(1)})_{i\in I_{n,k}}$ est grand lui aussi ; on vérifie donc aisément que le processus (X_i', φ_i') a la même loi que le processus initial (X_i^*, φ_i^*). De plus, le processus $(X_i^{(2)}, \varphi_i^{(1)})$ a la même loi que le processus $(\overline{X}_i, \overline{\varphi}_i)$.

Nous allons démontrer que les processus (X_i', φ_i') et $(X_i^{(2)}, \varphi_i^{(1)})$ vérifient presque sûrement les inégalités (2.2.13). Ainsi, le couplage naturel entre (X_i', φ_i') et $(X_i^{(2)}, \varphi_i^{(1)})$ engendre le couplage désiré entre (X_i^*, φ_i^*) et $(\overline{X}_i, \overline{\varphi}_i)$. C'est l'objet des deux prochains lemmes. Comme dans la sous-partie précédente, on commence par étudier les processus (φ_i') et $(\varphi_i^{(1)})$.

Lemme 2.2.7.

Supposons que φ vérifie la condition (1.3.2). *Soit $r^* \in (\beta+(1-\beta)r, 1)$. Presque sûrement, pour tout entier N suffisamment grand,*

$$\left|\sum_{i=0}^{N-1} \varphi_i' - \sum_{i=0}^{N-1} \varphi_i^{(1)}\right| \leq \psi^*(N^{r^*}). \tag{2.2.14}$$

Démonstration.

Les variables aléatoires φ_i' et $\varphi_i^{(1)}$ ne diffèrent que dans les blocs où elles sont toutes deux petites. Par conséquent :

$$\left|\sum_{i=0}^{N-1} \varphi_i' - \sum_{i=0}^{N-1} \varphi_i^{(1)}\right| \leq \sum_{n=0}^{\lfloor \log_2 N\rfloor+1} \sum_{k=0}^{2^{(1-q)n}-1} \left(1_{\left\{(X_i',\varphi_i')_{i\in I_{n,k}} \text{ petit}\right\}} \sum_{i\in I_{n,k}} \varphi_i' + 1_{\left\{(X_i^{(1)},\varphi_i^{(1)})_{i\in I_{n,k}} \text{ petit}\right\}} \sum_{i\in I_{n,k}} \varphi_i^{(1)} \right).$$

On va majorer le terme contenant les sommes $\sum_{i\in I_{n,k}} \varphi_i'$; le terme portant sur $\varphi_i^{(1)}$ peut être majoré de la même manière. Soit $r' \in (\beta+(1-\beta)r, r^*)$. Soit n un entier. D'après l'inégalité de Markov et le théorème de Karamata (théorème 1.3.4),

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=0}^{2^{(1-q)n}-1} 1_{\left\{(X_i',\varphi_i')_{i\in I_{n,k}} \text{ petit}\right\}} \sum_{i\in I_{n,k}} \varphi_i' > \psi^*(2^{r'n}) \right) &\leq \frac{2^{(1-q)n}\mathbb{E}\left(1_{\left\{(X_i',\varphi_i')_{i\in I_{n,0}} \text{ petit}\right\}} \sum_{i\in I_{n,0}} \varphi_i'\right)}{\psi^*(2^{r'n})} \\
&\leq \frac{2^n \left(\sum_{\ell=0}^{\lfloor \psi^*(2^{rn})\rfloor} \frac{1}{\psi(\ell)}\right)}{\psi^*(2^{r'n})} \\
&\leq C 2^{(1-r)n} \frac{\psi^*(2^{rn})}{\psi^*(2^{r'n})}.
\end{aligned}$$

Comme $r' > \beta+(1-\beta)r$, on peut choisir un paramètre $\kappa > 1-r$ tel que $r' > r+\kappa\beta$. Alors, en appliquant le lemme 1.3.9, on obtient :

$$\mathbb{P}\left(\sum_{k=0}^{2^{(1-q)n}-1} 1_{\left\{(X_i',\varphi_i')_{i\in I_{n,k}} \text{ petit}\right\}} \sum_{i\in I_{n,k}} \varphi_i' > \psi^*(2^{r'n}) \right) \leq C 2^{(1-r)n} \frac{\psi^*(2^{rn})}{\psi^*(2^{r'n})} \leq C 2^{-(\kappa-(1-r))n}.$$

Le membre de droite est sommable. D'après le lemme de Borel-Cantelli, presque sûrement, pour tout entier n suffisamment grand,

$$\sum_{k=0}^{2^{(1-q)n}-1} \left(1_{\left\{(X_i',\varphi_i')_{i\in I_{n,k}} \text{ petit}\right\}} \sum_{i\in I_{n,k}} \varphi_i' \right) \leq \psi^*(2^{r'n}).$$

On additionne cette inégalité pour n entre 0 et $\lfloor \log_2 N \rfloor + 1$. La fin de cette démonstration est similaire à la fin de la démonstration du lemme 2.2.4. Presque sûrement, pour tout entier N suffisamment grand,

$$\left| \sum_{i=0}^{N-1} \varphi'_i - \sum_{i=0}^{N-1} \varphi_i^{(1)} \right| \leq C \lfloor \log_2 N \rfloor \psi^*(N^{r'}) \leq \psi^*(N^{r^*}). \qquad \square$$

Passons maintenant à l'étude des processus (X'_i) et $(X_i^{(2)})$.

Lemme 2.2.8.
Supposons qu'il existe un réel $p > 2$ et une constante C tels que, pour tout entier n,

$$\mathbb{E}\left(\sup_{i<2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^* \right|^p \right) \leq C 2^{\frac{pqn}{2}}.$$

Soit $z > 0$ tel que $1 + (p-2)q/2 < p(z + r - 1)$. Presque sûrement, pour tout entier N suffisamment grand,

$$\left| \sum_{i=0}^{N-1} X'_i - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq N^z. \tag{2.2.15}$$

Démonstration.
Soit n un entier. On note $\nu_n^{(j)}$ le nombre de grands blocs du processus $(X_i^{(j)}, \varphi_i^{(j)})_{2^n \leq i < 2^{n+1}}$; cette variable aléatoire suit une loi binomiale $B(M, P)$ de paramètres $M = 2^{(1-q)n}$ et P, où la probabilité P est majorée par $2^{qn}\mathbb{P}(\varphi \geq \psi^*(2^{rn})) \leq 2^{-(r-q)n}$. En utilisant l'inégalité de Markov, on obtient une borne exponentielle sur le nombre de grands blocs :

$$\begin{aligned}
\mathbb{P}(\nu_n^{(j)} > 2^{(1-r)n+1}) &\leq e^{-2^{(1-r)n+1}} \mathbb{E}(e^{\nu_n^{(j)}}) \\
&\leq e^{-2^{(1-r)n+1}} (1 + (e-1)2^{(q-r)n})^{2^{(1-q)n}} \\
&\leq e^{-2^{(1-r)n+1}} e^{(e-1)2^{(q-r)n+(1-q)n}} \\
&= e^{-(3-e)2^{(1-r)n}}.
\end{aligned}$$

Posons $M_n^{(j)} := \sup_{k<2^{(1-q)n}} \sup_{i \in I_{n,k}} \left| \sum_{\ell=2^n+k2^{qn}}^{i} X_\ell^{(j)} \right|$. Soit N un entier, et soit $n \leq \lfloor \log_2 N \rfloor + 1$. Les processus X' et $X^{(2)}$ ne diffèrent que dans les blocs où l'un des $(X^{(1)}, \varphi^{(1)})$ et $(X^{(2)}, \varphi^{(2)})$ est grand, et si c'est le cas, les sommes $\sum_{\ell=2^n+k2^{qn}}^{i} X'_\ell$ et $\sum_{\ell=2^n+k2^{qn}}^{i} X_\ell^{(2)}$ diffèrent au plus de $M_n^{(1)} + M_n^{(2)}$. Par conséquent :

$$\left| \sum_{i=0}^{N-1} X'_i - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq \sum_{n=0}^{\lfloor \log_2 N \rfloor + 1} (\nu_n^{(1)} + \nu_n^{(2)})(M_n^{(1)} + M_n^{(2)}).$$

Soit $z' \in (0, z)$ tel que l'inégalité $1 + (p-2)q/2 < p(z' + r - 1)$ reste vraie. Alors :

$$\begin{aligned}
&\mathbb{P}\left((\nu_n^{(1)} + \nu_n^{(2)})\ (M_n^{(1)} + M_n^{(2)}) > 2^{z'n} \right) \\
&\leq 2\mathbb{P}(\nu_n^{(j)} > 2^{(1-r)n+1}) + 2\mathbb{P}\left(M_n^{(j)} > \frac{2^{z'n}2^{-(1-r)n}}{8} \right) \\
&\leq 2e^{-(3-e)2^{(1-r)n}} + 2 \cdot 2^{(1-q)n} \mathbb{P}\left(\sup_{i<2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^{(j)} \right| > \frac{2^{(z'+r-1)n}}{8} \right) \\
&\leq 2e^{-(3-e)2^{(1-r)n}} + C2^{(1-q)n} \frac{\mathbb{E}\left(\sup_{i<2^{qn}} \left| \sum_{\ell=2^n}^{2^n+i} X_\ell^{(j)} \right|^p \right)}{2^{p(z'+r-1)n}} \\
&\leq 2e^{-(3-e)2^{(1-r)n}} + C2^{(1-q)n} 2^{\frac{pqn}{2}} 2^{-p(z'+r-1)n}.
\end{aligned}$$

Le membre de droite est une fonction de n sommable. Par le lemme de Borel-Cantelli, presque sûrement, pour tout entier N suffisamment grand,

$$\left| \sum_{i=0}^{N-1} X_i' - \sum_{i=0}^{N-1} X_i^{(2)} \right| \leq CN^{z'}.$$

Quitte à prendre des valeurs de N encore plus grande, on a $CN^{z'} \leq N^z$ et la majoration (2.2.15) est démontrée. □

Pour terminer la démonstration de la proposition 2.2.6, il nous faut choisir les paramètres r (utilisé dans la définition du couplage) et z de façon à ce qu'ils vérifient l'inégalité suivante :

$$1 + \frac{p-2}{2} q < p(z + r - 1).$$

Rappelons que $r \in (0,1)$ et que $z \in (0,1/2)$ (sinon, le lemme 2.2.8 serait trivial). Sachant que $p > 2$ par hypothèse, si on choisissait de prendre $r = 1$ et $z = 1/2$, l'inégalité serait vérifiée pour tout $q \in (0,1)$. Par continuité, il suffit de choisir r suffisamment proche de 1 et z suffisamment proche de $1/2$.

2.2.5 Fin de la démonstration

Nous pouvons maintenant terminer la démonstration du théorème 2.1.1. Soient X et φ des fonctions satisfaisant les hypothèses du théorème. Pour tout i, on note $(X_i, \varphi_i) = (X \circ T_A^i, \varphi \circ T_A^i)$. Soit $(\tilde{X}_i, \tilde{\varphi}_i)$ un processus tel que $(\tilde{X}_i)$ et $(\tilde{\varphi}_i)$ soient des copies de (X_i) et (φ_i) respectivement, et tel que $(\tilde{X}_i)$ et $(\tilde{\varphi}_i)$ soient indépendants. On veut coupler les processus (X_i, φ_i) et $(\tilde{X}_i, \tilde{\varphi}_i)$.

Soient q et ε dans $(0,1)$. Voici les étapes successives de la construction du couplage, et les résultats que l'on utilise à chaque fois :

- On couple le processus (X_i, φ_i) avec un processus (X_i^*, φ_i^*) qui est i.i.d. par morceaux de paramètres q et ε, en utilisant le lemme 2.2.3. De par les lemmes 2.2.4 et 2.2.5, les sommes partielles du processus (X_i^*, φ_i^*) sont asymptotiquement proches des sommes partielles du processus initial.
- On utilise encore le lemme 2.2.3 pour coupler le processus $(\tilde{X}_i)$ avec un processus $(\overline{X}_i)$ qui est i.i.d. par morceaux de paramètres q et ε. De par le lemme 2.2.5, les sommes partielles de ces deux processus sont asymptotiquement proches.
- On utilise une dernière fois le lemme 2.2.3 pour coupler le processus $(\tilde{\varphi}_i)$ avec un processus $(\overline{\varphi}_i)$ qui est i.i.d. par morceaux de paramètres q et ε. De par le lemme 2.2.4, les sommes partielles de ces deux processus sont asymptotiquement proches.
- On couple les processus $(\tilde{X}_i, \overline{X}_i)$ et $(\tilde{\varphi}_i, \overline{\varphi}_i)$ de façon triviale (c'est-à-dire indépendante). On obtient alors un couplage entre $(\tilde{X}_i, \tilde{\varphi}_i)$ et $(\overline{X}_i, \overline{\varphi}_i)$. De plus, $(\tilde{X}_i)$ et $(\tilde{\varphi}_i)$ sont indépendants, de même que $(\overline{X}_i)$ et $(\overline{\varphi}_i)$. Les processus $(\tilde{X}_i)$ et (X_i^*) ont la même loi, de même que les processus $(\tilde{\varphi}_i)$ et (φ_i^*).
- On couple les processus (X_i^*, φ_i^*) et $(\overline{X}_i, \overline{\varphi}_i)$ en utilisant la proposition 2.2.6.

2.3 Du couplage au théorème limite

Nous allons maintenant exploiter le théorème 2.1.1, qui formalise l'indépendance asymptotique des sommes de Birkhoff $(\sum_{k=0}^{N-1} X \circ T_A^k)$ et $(\sum_{k=0}^{N-1} \varphi \circ T_A^k)$, pour déterminer la limite en loi des sommes de Birkhoff de fonctions d'intégrale nulle. En premier lieu, nous allons démontrer le théorème 2.1.3, qui porte sur le cas où l'indice β de variation régulière des queues du temps de premier retour est dans $[0,1)$. Le théorème 2.1.4, qui porte sur le cas $\beta = 1$, sera traité à part. Enfin, nous démontrerons une majoration presque sûre des sommes de Birkhoff.

Dans l'ensemble de cette partie, (Ω, μ, T) est un système dynamique muni d'une mesure infinie et σ-finie μ qui est T-invariante. On suppose de plus que ce système dynamique est conservatif et ergodique. On notera par A un sous-ensemble mesurable de Ω tel que $\mu(A) = 1$, par μ_A la mesure restreinte à A, et par T_A l'application induite en A. On suppose que le système dynamique

(A, μ_A, T_A) peut être muni d'une structure Gibbs-Markov. Enfin, on pose $\psi(x) := \mu_A(\varphi_A \geq x)^{-1}$ pour tout $x \geq 0$, et on supposera que la fonction ψ est à variation régulière d'indice $\beta \in [0,1]$ en l'infini.

2.3.1 Démonstration du théorème principal

Nous rappelons que, pour démontrer le théorème 2.1.3, nous avons adopté la méthode suivante. Tout d'abord, on démontre un résultat de convergence en loi si l'on suppose les processus $(\sum_{k=0}^{N-1} X_f \circ T_A^k)$ et $(\sum_{k=0}^{N-1} \varphi_A \circ T_A^k)$ indépendants ; c'est le lemme 2.3.1. Ensuite, on utilise le théorème 2.1.1 : ces processus sont asymptotiquement indépendants. Le problème est qu'il y a malgré tout une marge d'erreur dans le couplage que l'on a construit ; en particulier, le temps local en A pour ces deux processus va être différent. Les lemmes 2.3.2 et 2.3.3 ont pour objet de contrôler cette erreur.

Nous commençons par démontrer le lemme 2.3.1.

Lemme 2.3.1.
Soient $(A_N)_{N\in\mathbb{N}}$ et $(B_N)_{N\in\mathbb{N}}$ deux processus stochastiques indépendants, où les A_N sont des entiers positifs et les B_N prennent leurs valeurs dans un espace polonais E. Supposons que :

- *Il existe une suite (a_N) telle que $\lim_{N\to+\infty} a_N = +\infty$ et le processus (A_N/a_N) converge en loi vers une variable aléatoire positive $\mathcal{A}$;*
- *$\mathbb{P}(\mathcal{A} > 0) = 1$;*
- *Le processus (B_N) converge en loi vers une variable aléatoire $\mathcal{B}$ à valeurs dans E.*

Alors $(A_N/a_N, B_{A_N})$ converge en loi vers une variable aléatoire $(\mathcal{A}^*, \mathcal{B}^*)$, où $\mathcal{A}^*$ est une copie de $\mathcal{A}$, où $\mathcal{B}^*$ est une copie de $\mathcal{B}$, et où $\mathcal{A}^*$ et $\mathcal{B}^*$ sont indépendants.

Démonstration.
La suite de variables aléatoires $(A_N/a_N, B_{A_N})$ est tendue, et donc a des points limite. Il suffit de caractériser ces points limites.

Soient $0 < m < M$; on pose $I := (m, M)$. Soit J un sous-ensemble mesurable de E. Supposons que $\mathbb{P}(\mathcal{A} \in \partial I) = \mathbb{P}(\mathcal{B} \in \partial J) = 0$. Soit $\varepsilon > 0$. Étant donné que (B_N) converge vers $\mathcal{B}$ et a_N tend vers $+\infty$, pour tout entier N suffisamment grand, pour tout $k \geq m a_N$:

$$|\mathbb{P}(B_k \in J) - \mathbb{P}(\mathcal{B} \in J)| \leq \varepsilon.$$

Quitte à prendre des valeurs de N encore plus grandes, on peut aussi supposer que :

$$|\mathbb{P}(A_N/a_N \in I) - \mathbb{P}(\mathcal{A} \in I)| \leq \varepsilon.$$

Ainsi, pour tout entier N suffisamment grand :

$$\begin{aligned}
|\mathbb{P}(A_N \in a_N I,\ B_{A_N} \in J) & - \mathbb{P}(\mathcal{A} \in I)\mathbb{P}(\mathcal{B} \in J)| \\
& \leq \left| \sum_{k \in a_N I} \mathbb{P}(A_N = k,\ B_k \in J) - \mathbb{P}(A_N \in a_N I)\mathbb{P}(\mathcal{B} \in J) \right| \\
& \qquad + |\mathbb{P}(A_N \in a_N I)\mathbb{P}(\mathcal{B} \in J) - \mathbb{P}(\mathcal{A} \in I)\mathbb{P}(\mathcal{B} \in J)| \\
& \leq \sum_{k \in a_N I} \mathbb{P}(A_N = k) \, |\mathbb{P}(B_k \in J) - \mathbb{P}(\mathcal{B} \in J)| \\
& \qquad + |\mathbb{P}(A_N \in a_N I) - \mathbb{P}(\mathcal{A} \in I)| \, \mathbb{P}(\mathcal{B} \in J) \\
& \leq \varepsilon \sum_{k \in a_N I} \mathbb{P}(A_N = k) + \varepsilon \mathbb{P}(\mathcal{B} \in J) \leq 2\varepsilon.
\end{aligned}$$

Ainsi, toute mesure ν vers laquelle converge la loi d'une sous-suite de $(A_N/a_N, B_{A_N})$ vérifie $\nu((m, M) \times J) = \mathbb{P}(\mathcal{A} \in (m, M))\mathbb{P}(\mathcal{B} \in J)$ pour tous les ensembles de continuité (m, M) et J. Comme la plupart des boules sont des ensembles de continuité, de telles boîtes $(m, M) \times J$ engendrent la tribu borélienne sur $\mathbb{R}_+ \times E$, et ν est la mesure produit. □

Le lemme suivant nous permettra de contrôler l'accroissement du temps local de façon localement uniforme.

Lemme 2.3.2.
Soit (Ω, μ, T) un système dynamique qui induit une application Gibbs-Markov sur un ensemble borélien A. Supposons que la variable aléatoire φ_A sur l'espace (A, μ_A) satisfasse la condition (1.3.3).

Soient $r^* > r > 0$. Alors, μ_A-presque sûrement, pour tout entier N suffisamment grand,

$$\sup_{k \leq N} \left(\xi_{k+\psi^*(\psi(N)^r)} - \xi_k \right) \leq \psi(N)^{r^*}. \tag{2.3.1}$$

Démonstration.
Cette démonstration est inspirée de la preuve du lemme 3.5 dans [23]. Toutes les probabilités sont prises par rapport à la mesure μ_A. Pour tout entier naturel n, posons :

$$\mathbb{P}(n) := \mathbb{P}\left(\max_{k \leq \psi^*(n)} \left(\xi_{k+\psi^*(n^r)} - \xi_k \right) \geq (n-1)^{r^*} \right). \tag{2.3.2}$$

Les fonctions $\psi(k)$ et ξ_k sont croissantes en k. Par conséquent, si $\sup_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq (n-1)^{r^*}$ pour tout entier N tel que $\psi^*(n-1) \leq N < \psi^*(n)$,

$$\begin{aligned} \sup_{k \leq N} \left(\xi_{k+\psi^*(\psi(N)^r)} - \xi_k \right) &\leq \max_{k \leq \psi^*(n)} \left(\xi_{k+\psi^*(n^r)} - \xi_k \right) \\ &\leq (n-1)^{r^*} \leq \psi(N)^{r^*}. \end{aligned}$$

Il suffit donc de montrer que $\max_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq (n-1)^{r^*}$ pour tout entier n suffisamment grand. Par le lemme de Borel-Cantelli, il suffit de montrer que la suite $\mathbb{P}(n)$ est sommable. De plus, le maximum $\max_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k)$ est atteint juste avant que ξ_k ne "saute", c'est-à-dire quand $k = t_\ell - 1$ pour un certain entier ℓ. En effet, tant que le temps local ξ_k est constant (c'est-à-dire si $t_{\ell-1} \leq k \leq t_\ell - 1$ pour un certain entier ℓ), la fonction $\xi_{k+\psi^*(n^r)} - \xi_k$ est croissante. En faisant un pas de plus, ξ_k augmente de 1, et on obtient :

$$\max_{k \leq \psi^*(n)} (\xi_{k+\psi^*(n^r)} - \xi_k) \leq \max_{k:\ t_k \leq \psi^*(n)} \left(\xi_{t_k+\psi^*(n^r)} - \xi_{t_k} \right) + 1.$$

Quitte à prendre une valeur de r^* un peu plus petite et des valeurs de n plus grandes, on peut ignorer le terme constant. Alors :

$$\begin{aligned} \mathbb{P}(n) &\leq \mathbb{P}\left(\max_{k:\ t_k \leq \psi^*(n)} \left(\xi_{t_k+\psi^*(n^r)} - \xi_{t_k} \right) \geq (n-1)^{r^*} \right) \\ &\leq \mathbb{P}\left(\max_{k:\ t_k \leq \psi^*(n)} \left(\xi_{t_k+\psi^*(n^r)} - \xi_{t_k} \right) \geq (n-1)^{r^*},\ \xi_{\psi^*(n)} \leq n^2 \right) \\ &\quad + \mathbb{P}\left(\xi_{\psi^*(n)} \geq n^2 \right). \end{aligned}$$

La mesure μ_A est invariante par T_A, de telle sorte que :

$$\mathbb{P}(n) \leq n^2 \mathbb{P}\left(\xi_{\psi^*(n^r)} \geq (n-1)^{r^*} \right) + \mathbb{P}\left(\xi_{\psi^*(n)} \geq n^2 \right).$$

Remarquons que :

$$\mathbb{P}\left(\xi_{\psi^*(n^r)} \geq (n-1)^{r^*} \right) = \mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k \leq \psi^*(n^r) \right).$$

Soit $q \in (0,1)$ tel que $(1-q)r^* > r$. Afin de regagner de l'indépendance, on ne prend en compte que les valeurs $\varphi_A \circ T_A^{k(n-1)^{qr^*}}$ pour k compris entre 0 et $(n-1)^{(1-q)r^*} - 1$:

$$\mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k \leq \psi^*(n^r) \right) \leq \mathbb{P}\left(\sup_{0 \leq k < (n-1)^{(1-q)r^*}-1} \varphi_A \circ T_A^{k(n-1)^{qr^*}} \leq \psi^*(n^r) \right).$$

Soit $(\varphi'_k)_{k\geq 0}$ une famille de variables aléatoires i.i.d. de même loi que φ_A. Le temps de premier retour φ_A est constant sur chaque élément de la partition de A liée à l'application Gibbs-Markov. On peut donc appliquer les même techniques que dans la preuve du lemme 2.2.3. On couple les familles de variables aléatoires $(\varphi_A \circ T_A^{k(n-1)^{qr^*}})_{0\leq k<(n-1)^{(1-q)r^*}-1}$ et $(\varphi'_k)_{k\geq 0}$, de telle sorte que leur différence décroisse à vitesse exponentielle. Soit $C > 1$. Il existe des constantes $\rho \in (0,1)$ et $C', C'' > 0$ telles que, pour tout entier n suffisamment grand,

$$\mathbb{P}\left(\max_{0\leq k<(n-1)^{(1-q)r^*}-1} |\varphi_A \circ T_A^{k(n-1)^{qr^*}} - \varphi'_k| \geq C''\right) \leq C'(n-1)^{(1-q)r^*}\rho^{(n-1)^{qr^*}},$$

ce qui nous donne :

$$\begin{aligned}\mathbb{P}&\left(\max_{0\leq k<(n-1)^{(1-q)r^*}-1} \varphi_A \circ T_A^{k(n-1)^{qr^*}} \leq \psi^*(n^r)\right)\\ &\leq \mathbb{P}\left(\max_{0\leq k<(n-1)^{(1-q)r^*}-1} \varphi'_k \leq \psi^*(n^r) + C''\right) + C'(n-1)^{(1-q)r^*}\rho^{(n-1)^{qr^*}}.\end{aligned}$$

Quitte à prendre une valeur de r un peu plus grande, on peut oublier la constante C''. Ensuite :

$$\begin{aligned}\mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k \leq \psi^*(n^r)\right) &\leq \mathbb{P}\left(\varphi_A \leq \psi^*(n^r)\right)^{(n-1)^{(1-q)r^*}} + C'(n-1)^{(1-q)r^*}\rho^{(n-1)^{qr^*}}\\ &= \left(1 - \frac{1}{\psi(\psi^*(n^r))}\right)^{(n-1)^{(1-q)r^*}} + C'(n-1)^{(1-q)r^*}\rho^{(n-1)^{qr^*}} \quad ;\end{aligned}$$

Soit $C > 1$. Si n est suffisamment grand, alors $\psi(\psi^*(n^r)) \leq Cn^r$, d'où :

$$\mathbb{P}\left(\sum_{k=0}^{(n-1)^{r^*}-1} \varphi_A \circ T_A^k \leq \psi^*(n^r)\right) \leq e^{-\frac{(n-1)^{(1-q)r^*}}{Cn^r}} + C'(n-1)^{(1-q)r^*}\rho^{(n-1)^{qr^*}}.$$

Cette dernière quantité est sommable. En prenant $r = 1$ et $r^* = 2$, on s'aperçoit que la suite $n^2\mathbb{P}\left(\xi_{\psi^*(n)} \geq n^2\right)$ est elle aussi sommable. Par conséquent, $\mathbb{P}(n)$ est sommable. □

Nous avons aussi besoin du même type de contrôle sur les sommes partielles de la suite $(X_f \circ T_A^i)$, où f est une observable d'intégrale nulle sur Ω.

Lemme 2.3.3.

Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soit X une fonction à valeurs réelles sur A telle que $X \in \mathbb{L}^p(A, \mu_A)$ pour un certain $p > 2$, que $\mathbb{E}(D(X)) < +\infty$, et que $\mathbb{E}(X) = 0$.

Pour tout $r \in [0,1]$, pour tout $r^* > 2/p + (1-2/p)r$, presque sûrement, pour tout entier N suffisamment grand,

$$\sup_{k\leq N}\sup_{i\leq N^r}\left|\sum_{j=k}^{k+i-1} X \circ T_A^j\right| \leq N^{\frac{r^*}{2}}. \tag{2.3.3}$$

Démonstration.

Commençons par découper l'ensemble des entier naturels en morceaux de longueur 2^n pour $n \in \mathbb{N}$. Ensuite, découpons chacun de ces morceaux en $2^{(1-r)n}$ segments de longueur 2^{rn}. On peut contrôler la variation des sommes partielles de $(X \circ T_A^i)$ sur chacun de ces segments : pour tout

$r' > r$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{2^n \leq k < 2^{n+1}} \sup_{i \leq 2^{rn}} \left|\sum_{j=k}^{k+i-1} X \circ T_A^j\right| \geq 2^{\frac{r'(n-1)}{2}}\right) \\
&\qquad \leq \mathbb{P}\left(\sup_{k < 2^{(1-r)n}} \sup_{i \leq 2^{rn}} \left|\sum_{j=2^n+k2^{rn}}^{2^n+k2^{rn}+i-1} X \circ T_A^j\right| \geq \frac{2^{\frac{r'(n-1)}{2}}}{3}\right) \\
&\qquad \leq 2^{(1-r)n}\mathbb{P}\left(\sup_{i \leq 2^{rn}} \left|\sum_{j=0}^{i-1} X \circ T_A^j\right| \geq \frac{2^{\frac{r'(n-1)}{2}}}{3}\right) \\
&\qquad \leq C2^{(1-r)n}2^{-\frac{r'pn}{2}}\mathbb{E}\left(\sup_{i \leq 2^{rn}} \left|\sum_{j=0}^{i-1} X \circ T_A^j\right|^p\right).
\end{aligned}$$

D'après l'inégalité de Burkholder-Rosenthal 1.1.8, il existe une constante C telle que, pour tout entier N,

$$\mathbb{E}\left(\left|\sum_{i=0}^{N-1} X \circ T_A^i\right|^p\right) \leq CN^{\frac{p}{2}}.$$

Le processus $(X \circ T_A^i)$ étant stationnaire, on peut appliquer une fois encore le corollaire B1 de [62] :

$$\mathbb{E}\left(\sup_{i \leq 2^{rn}} \left|\sum_{j=0}^{i-1} X \circ T_A^j\right|^p\right) \leq C2^{\frac{rpn}{2}}.$$

Par conséquent, il existe une constante C telle que, pour tout n :

$$\mathbb{P}\left(\sup_{2^n \leq k < 2^{n+1}} \sup_{i \leq 2^{rn}} \left|\sum_{j=k}^{k+i-1} X \circ T_A^j\right| \geq 2^{\frac{r'n}{2}}\right) \leq C2^{(1-r)n}2^{-\frac{(r'-r)pn}{2}}.$$

On suppose maintenant que $r' \in (2/p + (1-2/p)r, r^*)$. Le membre de droite de cette inégalité est sommable. Par le lemme de Borel-Cantelli, presque sûrement, il existe une constante C telle que, pour tout N,

$$\sup_{k \leq N} \sup_{i \leq N^r} \left|\sum_{j=k}^{k+i-1} X \circ T_A^j\right| \leq CN^{\frac{r'}{2}}.$$

Étant donné que $r' < r^*$, on a démontré le lemme. □

Nous passons maintenant à la démonstration du théorème 2.1.3.

Démonstration du théorème 2.1.3.
Afin d'alléger les notations, on écrira :

$$K := \frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)}}.$$

La démonstration est divisée en trois étapes. Dans la première étape, on suppose que le point de départ x est choisi dans A suivant la mesure μ_A, et on obtient un théorème limite pour l'application induite T_A. Lors de la deuxième étape, on suppose toujours que x est choisi dans A suivant la mesure μ_A, mais on démontre cette fois-ci un théorème limite pour l'application initiale T. Enfin, lors de la troisième partie, on montre la convergence forte en loi ; en d'autre termes, on ne suppose plus que le point de départ est choisi selon μ_A, mais qu'il est choisi selon n'importe quelle mesure de probabilité absolument continue par rapport à μ.

Première étape : Sommes de Birkhoff pour l'application induite

On choisit le point de départ dans A suivant la mesure de probabilité μ_A. Posons $(X_i, \varphi_i) := (X_f \circ T_A^i, \varphi_A \circ T_A^i)$. Soit $(\tilde{X}_i, \tilde{\varphi}_i)$ un processus tel que $(\tilde{X}_i)$ soit une copie de (X_i), que $(\tilde{\varphi}_i)$ soit une copie de (φ_i), et que $(\tilde{X}_i)$ et $(\tilde{\varphi}_i)$ soient indépendants.

On pose $\tau_N := \sum_{i=0}^{N-1} \varphi_i$ ainsi que $\tilde{\tau}_N := \sum_{i=0}^{N-1} \tilde{\varphi}_i$. A x fixé, on note $\tilde{\xi}_N = (\tilde{\tau})^*(N)$ l'inverse généralisé de la fonction $\tilde{\tau}_N$. Le processus $(\tilde{\xi}_N)$ est une copie du processus (ξ_N), et est indépendant de $(\tilde{X}_i)$.

On couple les processus (X_i, φ_i) et $(\tilde{X}_i, \tilde{\varphi}_i)$ de façon à satisfaire les conclusions du théorème 2.1.1. D'après les propositions 1.1.7 et 1.4.8 et le lemme 2.3.1,

$$K \left(\frac{\tilde{\xi}_N}{\psi(N)} \right)^{\frac{1}{2}} \frac{1}{\sqrt{\tilde{\xi}_N}} \sum_{\ell=0}^{\tilde{\xi}_N - 1} \tilde{X}_\ell \to \sigma(f) \sqrt{Y_\beta} \mathcal{N},$$

où la convergence est en loi, où Y_β et $\mathcal{N}$ sont indépendants, Y_β suit une loi de Mittag-Leffler standard de paramètre β, où $\mathcal{N}$ suit une loi normale standard, et où la variance $\sigma(f)^2$ est donnée par la formule (2.1.2). De plus, $\sigma(f) = 0$ si et seulement si X_f est un cobord, donc si et seulement si f est un cobord (voir la discussion en sous-partie 1.2.1).

De plus, par le théorème 2.1.1, il existe $r < 1$ tel que, presque sûrement, pour tout entier N suffisamment grand,

$$\frac{K}{\sqrt{\psi(N)}} \left| \sum_{\ell=0}^{\tilde{\xi}_N - 1} \tilde{X}_\ell - X_f \circ T_A^\ell \right| \leq \frac{K}{\psi(N)^{\frac{1-r}{2}}} \left(\frac{\tilde{\xi}_N}{\psi(N)} \right)^{\frac{r}{2}},$$

et le membre de droite tend vers 0 en probabilité par la proposition 1.4.8. Par conséquent,

$$\frac{K}{\sqrt{\psi(N)}} \sum_{\ell=0}^{\tilde{\xi}_N - 1} X_f \circ T_A^\ell \to \sigma(f) \sqrt{Y_\beta} \mathcal{N}.$$

Remarquons que, pour tout N, on a $\xi_{\tau_N} = N$ et $\tau_{\xi_N} \leq N$, de telle sorte que :

$$\tilde{\xi}_N - \xi_N = \xi_{\tau_{\tilde{\xi}_N}} - \xi_N \leq \xi_{\tau_{\tilde{\xi}_N}} - \xi_{\tilde{\tau}_{\tilde{\xi}_N}}.$$

Soit $\kappa > 1$. Par le lemme 2.3.2 (on prend $r = 1$), presque sûrement, pour tout entier N suffisamment grand, la fonction ξ_N (respectivement $\tilde{\xi}_N$) est majorée par $\psi(N)^\kappa$. Comme le couplage entre τ_N et $\tilde{\tau}_N$ vérifie les conclusions du théorème 2.1.1, presque sûrement, pour tout entier N suffisamment grand,

$$\left| \tau_{\tilde{\xi}_N} - \tilde{\tau}_{\tilde{\xi}_N} \right| \leq \psi^*(\psi(N)^{\kappa r}).$$

Ceci est vrai pour tout $\kappa > 1$. Donc, pour tout $r' \in (r, 1)$, la différence $\left| \tau_{\tilde{\xi}_N} - \tilde{\tau}_{\tilde{\xi}_N} \right|$ est majorée par $\psi^*(\psi(N)^{r'})$ pour tout entier N suffisamment grand. Par le lemme 2.3.2, pour tout $r^* \in (r, 1)$ et pour tout N suffisamment grand,

$$\tilde{\xi}_N - \xi_N \leq \psi(N)^{r^*}.$$

On majore $\xi_N - \tilde{\xi}_N$ de la même façon. Ainsi, $\left| \xi_N - \tilde{\xi}_N \right| \leq \psi(N)^{r^*}$ pour tout $r^* \in (r, 1)$ et pour tout entier N suffisamment grand.

Soit $\kappa > 1$, et soit $r' \in (2/p + (1 - 2/p) r^*, 1)$. Par le lemme 2.3.3, pour tout entier N suffisamment grand,

$$\begin{aligned} \frac{K}{\sqrt{\psi(N)}} \left| \sum_{\ell=0}^{\tilde{\xi}_N - 1} X_f \circ T_A^\ell - \sum_{\ell=0}^{\xi_N - 1} X_f \circ T_A^\ell \right| &\leq \frac{K}{\sqrt{\psi(N)}} \sup_{k \leq \psi(N)^\kappa} \sup_{i \leq \psi(N)^{r^*}} \left| \sum_{\ell=k}^{k+i-1} X_f \circ T_A^\ell \right| \\ &\leq K \psi(N)^{\frac{r'\kappa - 1}{2}}. \end{aligned}$$

Si κ est suffisamment proche de 1, ceci converge en probabilité vers 0 quand N tend vers l'infini. Ainsi,

$$\frac{K}{\sqrt{\psi(N)}} \sum_{\ell=0}^{\xi_N - 1} X_f \circ T_A^\ell \to \sigma(f)\sqrt{Y_\beta}\mathcal{N}. \tag{2.3.4}$$

Deuxième étape : Théorème limite pour la loi initiale μ_A

On choisit encore le point de départ dans A suivant la mesure μ_A. On cherche à faire disparaître le temps local dans l'expression (2.3.4).

Pour tous $\ell \geq 0$ et $\varepsilon > 0$,

$$\mathbb{P}(X_{|f|} \circ T_A^\ell > \ell^{\frac{1+\varepsilon}{p}}) \leq C\ell^{-(1+\varepsilon)}.$$

Donc, par le lemme de Borel-Cantelli, presque sûrement, il existe une constante C telle que $X_{|f|} \circ T_A^\ell \leq C\ell^{\frac{1+\varepsilon}{p}}$. Si on remplace ℓ par $\xi_N - 1$, on voit que presque sûrement, il existe une constante C telle que $X_{|f|} \circ T_A^{\xi_N - 1} \leq C\xi_N^{\frac{1+\varepsilon}{p}}$. Par le lemme 2.3.2, pour tout $\kappa > 1$, presque sûrement, pour tout entier n suffisamment grand, $X_{|f|} \circ T_A^{\xi_N - 1} \leq C\psi(N)^{\frac{(1+\varepsilon)\kappa}{p}}$. Ainsi, pour tout entier N suffisamment grand,

$$\begin{aligned} \frac{1}{\sqrt{\psi(N)}} \left| \sum_{i=\tau_{\xi_N}}^{N-1} f \circ T^i \right| &\leq \frac{1}{\sqrt{\psi(N)}} X_{|f|} \circ T_A^{\xi_N - 1} \\ &\leq C\psi(N)^{\frac{(1+\varepsilon)\kappa}{p} - \frac{1}{2}}. \end{aligned}$$

Si on choisit ε suffisamment proche de 0 et κ suffisamment proche de 1, l'expression ci-dessus converge presque sûrement vers 0 quand N tend vers l'infini. Par conséquent :

$$\frac{K}{\sqrt{\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i = \frac{K}{\sqrt{\psi(N)}} \left(\sum_{\ell=0}^{\xi_N - 1} X_f \circ T_A^\ell + \sum_{i=\tau_{\xi_N}}^{N-1} f \circ T^i \right) \to \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \tag{2.3.5}$$

où la convergence est en loi quand le membre de gauche est vu comme une fonction à valeurs réelles définie sur l'espace probabilisé (A, μ_A).

Troisième étape : Convergence forte en distribution

Cette dernière étape est une application directe de [75, théorème 1] ; il ne reste qu'une hypothèse à vérifier. Soit N un entier. Pour tout $\varepsilon > 0$:

$$\mu\left(\frac{K}{\sqrt{\psi(N)}} \left| \sum_{k=0}^{N-1} (f \circ T - f) \circ T^k \right| > \varepsilon \right) \leq 2\mu\left(|f| > C\sqrt{\psi(N)}\varepsilon \right),$$

et, comme $f \in \mathbb{L}^1(\Omega, \mu)$, le membre de droite converge vers 0 quand N tend vers l'infini. Par conséquent, la convergence dans l'équation (2.3.5) est vrai non seulement sur (A, μ_A), mais sur (Ω, ν) pour toute mesure de probabilité $\nu \ll \mu$. □

Remarque 2.3.4.

En regardant cette démonstration d'un peu plus près, on peut voir que les hypothèses $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$ et $\int_\Omega f \, \mathrm{d}\mu = 0$ du théorème 2.1.3 peuvent être remplacées par l'ensemble d'hypothèses suivant, qui est plus faible :

- $\sup_{N \leq \varphi_A} \left| \sum_{k=0}^{N-1} f \circ T^k \right| \in \mathbb{L}^p(A, \mu_A)$;
- $\int_A X_f \, \mathrm{d}\mu_A = 0$;
- *Il existe $M > 0$ tel que $\mu(|f| > M) < +\infty$.*

En fait, on verra même en partie 3.1 que la troisième de ces hypothèses n'est pas nécessaire. Il suffit donc d'avoir $\sup_{N \leq \varphi_A} \left| \sum_{k=0}^{N-1} f \circ T^k \right| \in \mathbb{L}^p(A, \mu_A)$ *pour un certain* $p > 2$ *et* $\int_A X_f \, \mathrm{d}\mu_A = 0$.

Cette amélioration est suffisante pour étendre ce théorème à certaines "fonctions non intégrables d'intégrale nulle", telles que $f \notin \mathbb{L}^1(\Omega, \mu)$ *mais* $X_f \in \mathbb{L}^p(A, \mu_A)$.

L'avantage de ce critère plus fin est qu'il reste vérifié si l'on ajoute à f *un cobord borné, ce qui n'est pas vrai du critère initial* $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$. *Ce sera utile en partie 3.4.*

2.3.2 Étude du cas $\beta = 1$

Nous démontrons maintenant le théorème 2.1.4, qui est une version du théorème 2.1.3 pour un paramètre $\beta = 1$. Notre résultat d'indépendance asymptotique central, le théorème 2.1.1, ne peut pas s'appliquer dans cette situation. Cependant, une variable aléatoire ayant une loi de Mittag-Leffler standard de paramètre 1 vaut 1 presque sûrement, et une variable aléatoire constante est indépendante de n'importe quelle variable aléatoire. On dispose donc malgré tout d'une forme d'indépendance asymptotique ! On applique alors une astuce, utilisée par exemple dans [21, Section 2.3], pour contrôler les fluctuations stochastiques du temps local autour de sa limite.

Démonstration du théorème 2.1.4.
Pour cette démonstration, on pose $\tilde{\psi}(N) := N / \sum_{k=0}^{N-1} \frac{1}{\psi(k)}$. De plus, pour la durée de cette démonstration, les sommes sont orientées : pour tout $a > b$, on pose $\sum_{n=a}^{b-1} u_n = - \sum_{n=b}^{a-1} u_n$.

De même que dans la démonstration du théorème 2.1.3, on procède en trois étapes. Tout d'abord, on obtient un théorème limite pour la transformation induite et pour la mesure initiale μ_A. Ensuite, on étend ce théorème à la transformation initiale, mais toujours en travaillant avec la mesure μ_A. Enfin, on étend le résultat à n'importe quelle mesure initiale absolument continue par rapport à μ. Comme les deux dernières étapes sont exactement les mêmes que dans la démonstration du théorème 2.1.3, nous n'écrivons ici que la première étape.

On veut donc démontrer, sur l'espace probabilisé (A, μ_A), la convergence en loi suivante :

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\xi_N - 1} X_f \circ T_A^\ell \to \sigma(f) \mathcal{N}, \tag{2.3.6}$$

où $\sigma(f)^2$ est donné par la formule (2.1.2) et où $\mathcal{N}$ est une variable aléatoire suivant une loi normale standard. On divise la somme en deux morceaux :

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\xi_N - 1} X_f \circ T_A^\ell = \frac{1}{\sqrt{\tilde{\psi}(N)}} \left(\sum_{\ell=0}^{\tilde{\psi}(N)-1} X_f \circ T_A^\ell + \sum_{\ell=\tilde{\psi}(N)}^{\xi_N - 1} X_f \circ T_A^\ell \right).$$

D'après le théorème central limite 1.1.7 appliqué à X_f,

$$\frac{1}{\sqrt{\tilde{\psi}(N)}} \sum_{\ell=0}^{\tilde{\psi}(N)-1} X_f \circ T_A^\ell \to \sigma(f) \mathcal{N},$$

où la convergence est en loi. Il nous suffit donc de majorer le terme d'erreur. Par la proposition 1.4.8, la suite de variables aléatoires $(\xi_N / \tilde{\psi}(N))$ converge en loi, et donc en probabilité, vers 1. Soit $\varepsilon > 0$. Pour tout entier N suffisamment grand,

$$\mathbb{P}(|\xi_N - \tilde{\psi}(N)| > \varepsilon \tilde{\psi}(N)) \leq \varepsilon.$$

D'après l'inégalité de Burkholder-Rosenthal 1.1.8, il existe une constante C, ne dépendant que de f et de p, telle que, pour tout n,

$$\mathbb{E} \left(\sup_{0 < k \leq n} \left| \sum_{\ell=0}^{k-1} X_f \circ T_A^\ell \right|^p \right) \leq C n^{\frac{p}{2}},$$

ce qui implique, par l'inégalité de Markov, que :

$$\begin{aligned}\mathbb{P}\left(\frac{1}{\sqrt{\tilde{\psi}(N)}}\left|\sum_{\ell=\tilde{\psi}(N)}^{\xi_N-1} X_f \circ T_A^\ell\right| > \varepsilon^{\frac{1}{3}}\right) &\leq \varepsilon + \mathbb{P}\left(\sup_{0<k\leq\varepsilon\tilde{\psi}(N)}\left|\sum_{\ell=0}^{k-1} X_f \circ T_A^\ell\right| > \varepsilon^{\frac{1}{3}}\sqrt{\tilde{\psi}(N)}\right) \\ &\leq \varepsilon + \frac{C(\varepsilon\tilde{\psi}(N))^{\frac{p}{2}}}{\varepsilon^{\frac{p}{3}}\tilde{\psi}(N)^{\frac{p}{2}}} \\ &= \varepsilon + C\varepsilon^{\frac{p}{6}}.\end{aligned}$$

Le terme d'erreur converge donc en probabilité vers 0, ce que l'on voulait démontrer. □

2.3.3 Borne presque sûre

Pour clore cette partie, nous allons démontrer un dernier théorème, qui donne une majoration presque sûre de la croissance des sommes de Birkhoff de fonctions d'intégrale nulle. Ce résultat n'est pas aussi fin qu'une loi du logarithme itéré, mais se démontre facilement avec les outils que nous avons développés.

Théorème 2.3.5.
Soit (Ω, μ, T) un système dynamique préservant la mesure qui induit une application Gibbs-Markov mélangeante sur un ensemble borélien A. Supposons que l'inverse des queues de la variable aléatoire φ_A sur (A, μ_A) est à variation régulière de paramètre $\beta \in [0,1]$. Soit $f \in \mathbb{L}^1(\Omega, \mu)$ telle que $\int_\Omega f \, \mathrm{d}\mu = 0$, que $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$ pour un certain $p > 2$ et que $\mathbb{E}(D(X_f)) < +\infty$.

Alors μ-presque partout, pour tout $\varepsilon > 0$,

$$\lim_{N\to+\infty} \frac{1}{N^{\frac{\beta}{2}+\epsilon}} \sum_{k=0}^{N-1} f \circ T^k = 0. \tag{2.3.7}$$

La démonstration de ce théorème utilise les mêmes outils que la démonstration du théorème 2.1.3, mais d'une façon nettement moins subtile.

Démonstration.
Pour commencer, on choisit le point de départ x dans A suivant la loi μ_A. Supposons que $\beta \in [0,1)$. Presque sûrement, ξ_N tend vers l'infini. Par le lemme 2.3.3 (on prend $r = 1$), pour tout $\varepsilon > 0$, presque sûrement, pour tout entier N suffisamment grand,

$$\left|\sum_{k=0}^{\xi_N-1} X_f \circ T_A^k\right| \leq \xi_N^{\frac{1}{2}+\varepsilon}. \tag{2.3.8}$$

D'après le lemme 2.3.2 (on prend $r = 1$), pour tout $\varepsilon > 0$, presque sûrement, pour tout entier N suffisamment grand,

$$\xi_N \leq \psi(N)^{1+\varepsilon}. \tag{2.3.9}$$

On utilise conjointement les inégalités (2.3.8) et (2.3.9). Pour tout $\varepsilon > 0$, presque sûrement, pour tout entier N suffisamment grand,

$$\left|\sum_{k=0}^{\xi_N-1} X_f \circ T_A^k\right| \leq \psi(N)^{\frac{1}{2}+\varepsilon}.$$

Si $\beta = 1$, il suffit d'utiliser le lemme 2.3.3 en choisissant $r = 1$.

On réutilise la deuxième étape de la démonstration du théorème 2.1.3. Pour tout $\varepsilon > 0$, presque sûrement, pour tout entier N suffisamment grand,

$$\left|\sum_{k=0}^{N-1} f \circ T^k\right| \leq \psi(N)^{\frac{1}{2}+\varepsilon}.$$

Maintenant, on veut une propriété valable μ-presque partout dans Ω. Mais, pour μ-presque tout x, il existe un entier k tel que $T^k x$ appartienne au sous-ensemble de A de mesure pleine des points qui satisfont la propriété désirée. De plus, la somme des $f \circ T^i x$ pour $i < k$ est aussi finie pour μ-presque tout x dans Ω.

Finalement, il reste à intervertir les quantificateurs "pour tout $\varepsilon > 0$" et "μ-presque partout". Le procédé est classique. Pour tout entier $M > 0$, l'ensemble Ω_M des points x tels que $\sum_{k=0}^{n-1} f \circ T^k x = o(N^{\beta/2+1/M})$ est de mesure pleine. L'ensemble $\bigcap_{M \in \mathbb{N}^*} \Omega_M$ est donc lui aussi de mesure pleine. □

Remarque 2.3.6.

Une fois de plus, on peut démontrer ce théorème sous des hypothèses plus faibles, qui ne requièrent pas que f soit intégrable. Les hypothèses $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$ et $\int_\Omega f \, \mathrm{d}\mu = 0$ dans le théorème 2.3.5 peuvent être remplacées par :

- $\sup_{N \leq \varphi_A} \left| \sum_{k=0}^{N-1} f \circ T^k \right| \in \mathbb{L}^p(A, \mu_{|A})$;
- $\int_A X_f \, \mathrm{d}\mu_A = 0$.

La vitesse de convergence du théorème 2.3.5 pourrait être améliorée jusqu'à ce que l'on obtienne une loi du logarithme itéré, comme l'ont fait M.B. Marcus et J. Rosen pour le temps local de marches aléatoires sur $\mathbb{Z}$ et $\mathbb{Z}^2$ [54, théorème 1.1].

2.4 Bases non mélangeantes

Nous allons maintenant affaiblir l'hypothèse de mélange présente dans les théorèmes 2.1.1 et 2.1.3. Nous allons la remplacer par une hypothèse d'ergodicité, au prix d'une expression plus compliquée de la variance. Autrement dit, nous allons démontrer dans cette sous-partie la proposition suivante :

Proposition 2.4.1.

Le terme "mélangeante" peut être remplacé par le terme "ergodique" dans les énoncés des théorèmes 2.1.1, 2.1.3, 2.1.4 et 2.3.5, pourvu que la formule de la variance donnée par l'équation (2.1.2) *soit remplacée par :*

$$
\begin{aligned}
\sigma(f)^2 &= \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i \right)^2 \mathrm{d}\mu_A \\
&= \frac{1}{\sqrt{M}} \left[\int_A \left(\sum_{i=0}^{M-1} X_f \circ T_A^i \right)^2 \mathrm{d}\mu_A \right. \\
&+ \left. 2 \sum_{n=1}^{+\infty} \int_A \left(\sum_{i=0}^{M-1} X_f \circ T_A^i \right) \cdot \left(\sum_{i=0}^{M-1} X_f \circ T_A^i \right) \circ T_A^{Mn} \, \mathrm{d}\mu_A \right],
\end{aligned}
$$

où M est la période de l'application Gibbs-Markov.

Pour démontrer cette proposition, nous allons utiliser le fait, évoqué dans la sous-partie 1.1.2, que, si une application Gibbs-Markov est ergodique, alors l'un de ses itérés aura une décomposition en composantes mélangeantes. Rappelons que pour une application Gibbs-Markov ergodique $(A, \pi, \lambda, \mu_A, T_A)$, il existe un unique entier $M \geq 1$ et une partition $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$ de A en M sous-ensembles tels que :

- La partition $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$ est moins fine que π ;
- Les sous-ensembles A_k sont invariants par T_A^M à un ensemble de mesure nulle près ;
- $T_A(A_k) = A_{k+1}$ pour tout $k \in \mathbb{Z}/M\mathbb{Z}$;
- Les M systèmes dynamiques (A_k, d, T_A^M) sont topologiquement mélangeants.

On utilise la technique de saut de Schweiger décrite en sous-partie 1.2.5 avec le temps d'arrêt constant et égal à M. L'application T_A^M préserve la mesure μ_A. D'après le lemme 1.2.6, on obtient une application Gibbs-Markov $(A, \pi_M, \lambda, T_A^M, \mu_A)$. Ce système dynamique n'est pas ergodique, mais a exactement M composantes ergodiques $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}}$. De plus, ces composantes ont la même mesure, et le système dynamique sur chacune de ces composantes est mélangeant. On note $\mu_k := \mu_A(\cdot|A_k) = M\mu_{A|A_k}$, de telle sorte que $T_{A*}^\ell \mu_k = \mu_{k+\ell}$ pour tout entier positif ℓ et pour tout $k \in \mathbb{Z}/M\mathbb{Z}$. On note aussi $D = D_{\pi,d}$ et $D_M = D_{\pi_M, d_M}$.

Si l'on applique directement le théorème 2.1.1 à chacune de ces composantes, les conclusions du théorème sont valables pour les processus $(X_{Mi+\ell}, \varphi_{Mi+\ell})_{i \geq 0}$, où $0 \leq \ell < M$. Pour avoir un contrôle sur le processus initial $(X_i, \varphi_i)_{i \geq 0}$, nous travaillerons avec le processus à valeurs vectorielles $((X_{Mi+\ell}, \varphi_{Mi+\ell})_{0 \leq \ell < M})_{i \geq 0}$ défini sur chaque A_k, auxquels nous appliquerons une version du théorème 2.1.1.

Pour commencer, nous allons décrire la façon dont l'intégrabilité et la régularité des observables se comportent lorsque que l'on itère la transformation M fois. Soit f une fonction définie sur A et à valeurs réelles. Soient $k \in \mathbb{Z}/M\mathbb{Z}$ et $0 \leq \ell < M$. Si f appartient à $\mathbb{L}^p(A, \mu_A)$ pour un certain $p \geq 1$, alors :

$$\begin{aligned}
\left\| f \circ T_A^\ell \right\|_{\mathbb{L}^p(A_k, \mu_k)} &= \left(\int_{A_k} |f \circ T_A^\ell|^p M \, \mathrm{d}\mu_{A|A_k} \right)^{\frac{1}{p}} \\
&= M^{\frac{1}{p}} \left(\int_{A_{k+\ell}} |f|^p \, \mathrm{d}\mu_{A|A_{k+\ell}} \right)^{\frac{1}{p}} \\
&\leq M^{\frac{1}{p}} \left\| f \right\|_{\mathbb{L}^p(A, \mu_A)}.
\end{aligned}$$

Ainsi, les fonctions $f \circ T_A^\ell$ appartiennent à $\mathbb{L}^p(A_k, \mu_k)$ pour tout $\ell \geq 0$. De même, si f est positive et satisfait la condition (1.3.2) avec une fonction auxiliaire ψ, alors, pour tout $x \geq 0$:

$$\mathbb{P}_{\mu_k}(f \circ T_A^\ell > x) = \mathbb{P}_{\mu_{k+\ell}}(f > x) \leq M \mathbb{P}_{\mu_A}(f > x) \leq \frac{M}{\psi(x)}.$$

Ainsi, les fonctions $f \circ T_A^\ell$ satisfont toutes la condition (1.3.2) avec la fonction auxiliaire ψ/M.

Supposons maintenant que f est lipschitzienne sur chaque élément de la partition π. Soit $a_M \in \pi_M$, et soit $k \in \mathbb{Z}/M\mathbb{Z}$ tel que $a_M \subset A_k$. Soient x et y dans a_M. Soit $0 \leq \ell < M$, et soit a l'élément de la partition π tel que $T_A^\ell a_k \subset a$. Alors :

$$\begin{aligned}
|f(T_A^\ell x) - f(T_A^\ell y)| &\leq |f|_{\mathrm{Lip}_d(a)} d(T_A^\ell x, T_A^\ell y) \\
&= \lambda^{\ell - M} |f|_{\mathrm{Lip}_d(a)} d(T_A^M x, T_A^M y) \\
&\leq \lambda^{\ell - M + 1} |f|_{\mathrm{Lip}_d(a)} d_M(x, y),
\end{aligned}$$

et donc $|f \circ T_A^\ell|_{\mathrm{Lip}_{d_M}(a_M)} \leq \lambda^{\ell - M + 1} |f|_{\mathrm{Lip}(a)}$. En d'autres termes, $D_M(f \circ T_A^\ell) \leq \lambda^{\ell - M + 1} D(f) \circ T_A^\ell$. En intégrant la fonction $D_M(f \circ T_A^\ell)$ sur A_k, on obtient :

$$\mathbb{E}_{\mu_k} \left(D_M(f \circ T_A^\ell) \right) \leq \lambda^{\ell - M + 1} \mathbb{E}_{\mu_{k+\ell}} \left(D(f) \right) \leq M\lambda \mathbb{E}_{\mu_A} \left(D(f) \right).$$

Par conséquent, si $\mathbb{E}_{\mu_A}(D(f))$ est fini, alors il en est de même pour $\mathbb{E}_{\mu_k} \left(D_M(f \circ T_A^\ell) \right)$ pour tout $0 \leq \ell < M$ et pour tout $k \in \mathbb{Z}/M\mathbb{Z}$. Cette propriété reste vraie pour toute condition faible de régularité, comme celles que l'on manipulera en partie 2.5 (voir la condition (2.5.1)).

Maintenant que nous avons exposé les propriétés fondamentales de cette induction, nous allons pouvoir démontrer la proposition 2.4.1. Le lecteur pourra vérifier de lui-même que le théorème 2.1.1 peut être aisément modifié pour prendre en compte de multiples occurrences d'observables dans $\mathbb{L}^p$ et d'observables à queues épaisses (il suffit de modifier légèrement l'énoncé et la démonstration de la proposition 2.2.6) :

Proposition 2.4.2.

Soit $(A, \pi, \lambda, \mu_A, T_A)$ une application Gibbs-Markov mélangeante. Soit M un entier strictement positif. Soient $X^{(0)}, \cdots, X^{(M-1)}$ et $\varphi^{(0)}, \cdots, \varphi^{(M-1)}$ des fonctions mesurables définies sur A et à valeurs dans $\mathbb{R}$ et dans $\mathbb{R}_+$ respectivement, telles que $\mathbb{E}_{\mu_A}(D(X^{(i)}))$ et $\mathbb{E}_{\mu_A}(D(\varphi^{(i)}))$ soient finis *pour tout* $0 \leq i < M$. *On pose* $(X_i^{(\ell)}, \varphi_i^{(\ell)}) := (X^{(\ell)} \circ T_A^i, \varphi^{(\ell)} \circ T_A^i)$. *Soient* $((\tilde{X}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *et* $((\tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *des copies des processus* $((X_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *et* $((\varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *respectivement, telles que* $((\tilde{X}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *et* $((\tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *soient mutuellement indépendants.*

Supposons qu'il existe un réel $p > 2$ *tel que* $X^{(\ell)} \in \mathbb{L}^p(A, \mu_A)$. *Supposons qu'il existe* $\beta \in [0, 1)$ *tel que chaque* $\varphi^{(\ell)}$ *satisfait la condition* (1.3.2) *avec la même fonction auxiliaire* ψ. *Alors il existe* $r \in (0, 1)$ *et un couplage entre* $((X_i^{(\ell)}, \varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *et* $((\tilde{X}_i^{(\ell)}, \tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ *tels que, presque sûrement, pour tout entier* N *suffisamment grand,*

$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} X_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{X}_i^{(\ell)} \right| \leq N^{\frac{r}{2}},$$
$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} \varphi_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{\varphi}_i^{(\ell)} \right| \leq \psi^*(N^r).$$

Nous pouvons alors démontrer la proposition 2.4.1.

Démonstration de la proposition 2.4.1.

Démonstration du théorème 2.1.1 sans mélange

Soit $(A, \pi, \lambda, \mu_A, T_A)$ une application Gibbs-Markov ergodique. Soient X et φ des fonctions mesurables définies sur A et à valeurs dans $\mathbb{R}$ et $\mathbb{R}_+$ respectivement, telles que $\mathbb{E}_{\mu_A}(D(X))$ et $\mathbb{E}_{\mu_A}(D(\varphi))$ soient finis, telles que $X \in \mathbb{L}^p(A, \mu_A)$ pour un certain réel $p > 2$ et que φ satisfasse la condition (1.3.2) pour un certain $\beta \in [0, 1)$ et une certaine fonction auxiliaire ψ. Soient $M \geq 1$ et $(A_k)_{k \in \mathbb{Z}/M\mathbb{Z}} \subset \pi$ comme dans le début de cette sous-partie. Soit $k \in \mathbb{Z}/M\mathbb{Z}$.

Pour tout $0 \leq \ell < M$, on pose $X^{(\ell)} := X \circ T_A^\ell$ et $\varphi^{(\ell)} := \varphi \circ T_A^\ell$. Ces observables sont bien définies μ_k-presque partout sur A_k. De plus, on sait grâce à la discussion précédente que chaque fonction $X^{(\ell)}$ appartient à $\mathbb{L}^p(A_k, \mu_k)$, que chaque fonction $\varphi^{(\ell)}$ satisfait la condition (1.3.2) (avec le même paramètre β et la fonction auxiliaire ψ/M), et que les espérances $\mathbb{E}_{\mu_k}(D_M(X^{(\ell)}))$ et $\mathbb{E}_{\mu_k}(D_M(\varphi^{(\ell)}))$ sont finies. On peut donc appliquer la proposition 2.4.2 au système Gibbs-Markov $(A_k, \pi_M, \lambda, \mu_k, T_A^M)$ et aux familles d'observables $(X^{(\ell)})_{0 \leq \ell < M}$ et $(\varphi^{(\ell)})_{0 \leq \ell < M}$. Remarquons au passage que, pour tout x positif, $(\psi/M)^*(x) = \psi^*(Mx)$.

Soit $((\tilde{X}_i^{(\ell)}, \tilde{\varphi}_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ un processus stochastique comme défini dans la proposition 2.4.2 et couplé avec $((X_i^{(\ell)}, \varphi_i^{(\ell)})_{0 \leq \ell < M})_{i \in \mathbb{N}}$ de telle sorte que, μ_k-presque sûrement, pour tout entier N suffisamment grand,

$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} X_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{X}_i^{(\ell)} \right| \leq N^{\frac{r}{2}},$$
$$\sum_{\ell=0}^{M-1} \left| \sum_{i=0}^{N-1} \varphi_i^{(\ell)} - \sum_{i=0}^{N-1} \tilde{\varphi}_i^{(\ell)} \right| \leq \psi^*(MN^r).$$

Pour tous entiers $i \geq 0$ et $0 \leq \ell < M$, on pose $\overline{X}_{Mi+\ell} := \tilde{X}_i^{(\ell)}$. On définit de même un processus stochastique $(\overline{\varphi}_i)$. Alors les processus $(\overline{X}_i)$ et $(\overline{\varphi}_i)$ sont indépendants, et ont respectivement la loi de $(X \circ T_A^i x)$ et de $(\varphi \circ T_A^i x)$ quand x est tiré au hasard avec la loi μ_k. De plus, μ_k-presque sûrement, pour tout entier N suffisamment grand, si l'on écrit $N = N_0 M + N_1$ avec $0 \leq N_1 < M$,

$$\left| \sum_{i=0}^{N-1} \left(X \circ T_A^i - \overline{X}_i\right) \right| \leq \sum_{\ell=0}^{N_1-1} \left| \sum_{i=0}^{N_0+1} \left(X_i^{(\ell)} - \tilde{X}_i^{(\ell)}\right) \right| + \sum_{\ell=N_1}^{M-1} \left| \sum_{i=0}^{N_0} \left(X_i^{(\ell)} - \tilde{X}_i^{(\ell)}\right) \right| \leq 2 \left(\frac{N-1}{M} + 1 \right)^{\frac{r}{2}},$$

et l'on dispose d'une majoration similaire pour la différence entre les sommes partielles $\sum_{i=0}^{N-1} \varphi \circ T_A^i$ et $\sum_{i=0}^{N-1} \overline{\varphi}_i$. Quitte à choisir un paramètre r un peu plus grand, si N est suffisamment grand, le couplage que l'on a construit entre $(X \circ T_A^i, \varphi \circ T_A^i)$ et $(\overline{X}_i, \overline{\varphi}_i)$ satisfait les conclusions du théorème 2.1.1 quand le point de départ est choisit avec la mesure μ_k.

Finalement, remarquons que μ est la moyenne des mesures μ_k pour $k \in \mathbb{Z}/M\mathbb{Z}$:

$$\mu = \frac{1}{M} \sum_{k=0}^{M-1} \mu_k.$$

Pour chaque k, on dispose d'un couplage entre $(X \circ T_A^i, \varphi \circ T_A^i)$ sous la loi μ_k et le processus constitué d'une copie de $(X \circ T_A^i)$ et d'une copie de $(\varphi \circ T_A^i)$ indépendantes. En prenant la moyenne de ces couplages, on obtient un nouveau couplage entre $(X \circ T_A^i, \varphi \circ T_A^i)$ sous la loi μ et un processus $(\hat{X}_i, \hat{\varphi}_i)$ qui satisfait les conclusions du théorème 2.1.1.

Démonstration du théorème 2.1.3 sans mélange

Dans la sous-partie 2.3.1, on déduit le théorème 2.1.3 du théorème 2.1.1 tout d'abord en démontrant trois lemmes (les lemmes 2.3.1, 2.3.2 et 2.3.3), puis en les utilisant dans la preuve du théorème 2.1.1 proprement dite. Nous devons vérifier que chacun de ces résultats peut être transposé au cas où le système Gibbs-Markov induit est seulement ergodique.

Le lemme 2.3.1 peut être utilisé tel quel.

L'adaptation du lemme 2.3.2 est le point le moins facile de cette démonstration. Rappelons que ce lemme signifie que le processus $(\xi_N)_{n \geq 0}$ ne croît pas trop vite, ou de façon équivalente que son inverse $(\sum_{i=0}^{N-1} \varphi_A \circ T_A^i)$ croît suffisamment vite. La preuve repose sur le fait que, pour tous $0 < r < r^*$, la suite $\mathbb{P}(n)$ définie par l'équation (2.3.2) est sommable, ce qui demande une forme d'indépendance pour le processus $(\varphi_A \circ T_A^i)$. Dans le cas mélangeant, on se donnait un paramètre q bien choisit puis on travaillait avec la sous-suite $((\varphi_A \circ T_A^{n^{qr^*} i})_{0 \leq i < n^{(1-q)r^*}})_{n \geq 0}$. Dans le cas où la transformation Gibbs-Markov est ergodique mais pas mélangeante, on induit la transformation sur l'une des parties A_k. Ceci nous amène à travailler avec la sous-suite $((\sum_{\ell=0}^{M-1} \varphi_A \circ T_A^\ell) \circ T_A^{M n^{qr^*} i})$. La plupart des calculs dans la démonstration du lemme 2.3.2 empirent d'un terme constant dépendant de M, sans influence sur la sommabilité de la suite $\mathbb{P}(n)$. On peut donc remplacer l'hypothèse de mélange par une hypothèse d'ergodicité dans l'énoncé du lemme 2.3.2.

Dans la preuve du lemme 2.3.3, on utilise l'inégalité de Burkholder-Rosenthal ; cependant, on dispose déjà de cette inégalité le long de la sous-suite $(X_f \circ T_A^{Mi+\ell})_{i \geq 0}$ pour tout $0 \leq \ell < M$, ce qui permet d'en déduire l'inégalité de Burkholder-Rosenthal pour le processus $(X_f \circ T_A^i)_{i \geq 0}$.

Pour finir, le point auquel on doit faire attention dans la démonstration du théorème 2.1.3 est le théorème central limite pour des observables régulières et de moyenne nulle d'une application Gibbs-Markov. Cependant, pour tout $0 \leq k < M$, le système $(A_k, \pi_M, \lambda, \mu_k, T_A^M)$ est mélangeant. D'après le théorème central limite (Proposition 1.1.7),

$$\lim_{N \to +\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X_f \circ T_A^i = \sigma(f) \mathcal{N},$$

où la convergence est en distribution sur l'espace probabilisé (A_k, μ_k), où $\mathcal{N}$ est une variable aléatoire gaussienne standard, et où $\sigma(f)$ est donné par :

$$\begin{aligned}
\sigma(f)^2 &= \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i \right)^2 \mathrm{d}\mu_A \\
&= \frac{1}{M} \left[\int_A \left(\sum_{k=0}^{M-1} X_f \circ T_A^k \right)^2 \mathrm{d}\mu_A + 2 \sum_{i=1}^{+\infty} \int_A \left(\sum_{k=0}^{M-1} X_f \circ T_A^k \right) \left(\sum_{k=Mi}^{Mi+M-1} X_f \circ T_A^k \right) \mathrm{d}\mu_A \right].
\end{aligned}$$

L'expression de la variance dans le théorème 2.1.3 doit donc être changée. La convergence en loi du temps local, elle, reste inchangée : la proposition 1.4.8 n'utilise qu'une hypothèse d'ergodicité, et non de mélange.

Démonstration des théorèmes 2.1.4 et 2.3.5 sans mélange

Les modifications à apporter sont semblables à celles que l'on a apportées à la démonstration du théorèmes 2.1.3. □

2.5 Condition de régularité faible

Dans cette partie, nous allons démontrer les théorèmes 2.1.1, 2.1.3, 2.1.4 et 2.3.5 sous des conditions de régularité faibles. Pour une application Gibbs-Markov (A, π, d, μ_A, T_A) et des observables X et φ, nous demandions souvent que les conditions de régularité $\mathbb{E}(D(X)) < +\infty$ et $\mathbb{E}(D(\varphi)) < +\infty$ soient satisfaites. Ces conditions garantissent une décroissance exponentielle des corrélations. Par exemple, si $X \in \mathbb{L}^1(A, \mu_A)$ vérifie $\mathbb{E}(X) = 0$ et $\mathbb{E}(D(X)) < +\infty$, alors $\mathcal{L}X \in \mathrm{Lip}^\infty$, et donc la suite $(\mathcal{L}^n X)_{n \geq 1}$ décroît vers 0 à vitesse exponentielle en norme Lip^∞, et a fortiori en norme $\mathbb{L}^p$.

Pour θ dans $(0, 1]$, quitte à travailler avec d'autres espaces de Banach que Lip^∞, on peut démontrer tous nos résultats en supposant que $\mathbb{E}(D(X)^\theta)$ et $\mathbb{E}(D(\varphi)^\theta)$ sont finis. Cette condition garantit encore une décroissance exponentielle des corrélations, quoiqu'elle soit en générale plus lente si l'on diminue θ.

Cette amélioration est classique. Cependant, nous irons plus loin, et choisirons un critère de régularité qui ne garantit qu'une décroissance polynomiale des corrélations. Soit $\ln_+$ la partie positive du logarithme. Pour un paramètre $\theta > 0$, on pourra supposer que :

$$\mathbb{E}(\ln_+(D(X))^{1+\theta}) + \mathbb{E}(\ln_+(D(\varphi))^{1+\theta}) < +\infty, \tag{2.5.1}$$

ce qui ne garantira qu'une décroissance polynomiale des corrélations. Le prix à payer est que la condition de régularité pour X et φ d'une part, et la condition d'intégrabilité sur X d'autre part (un réel $p > 2$ tel que $X \in \mathbb{L}^p(A, \mu_A)$), ne seront plus indépendantes. Nous démontrerons avec la proposition 2.5.5 que les conclusions des théorèmes 2.1.1, 2.1.3, 2.1.4 et 2.3.5 restent valables si l'on utilise la condition de régularité (2.5.1) avec $\theta > 1/(p-1)$.

L'amélioration apportée est assez marginale, mais les résultats préliminaires de la prochaine sous-partie restent intéressants non seulement pour le problème du théorème central limite en mesure infinie, mais plus généralement pour la théorie des applications Gibbs-Markov.

2.5.1 Estimées préliminaires

Nous allons commencer par récupérer le théorème central limite et l'inégalité de Burkholder-Rosenthal sous la condition de régularité faible. Étant donné que cette condition ne garantit qu'une décroissance polynomiale des corrélations, les méthodes spectrales ne peuvent plus être employées directement. On utilisera à la place des techniques de martingales à la Gordin [29]. Notre premier résultat donne une majoration sur la vitesse à laquelle la norme de $\mathcal{L}^n X$ décroît quand n croît.

Lemme 2.5.1.
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soit $p > 1$ et soit $X \in \mathbb{L}^p(A, \mu_A)$ tel que $\int_A X \, \mathrm{d}\mu_A = 0$. Supposons que $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$ pour un certain réel $\theta > 0$.

Alors $\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p} = O(N^{-\frac{(p-1)(1+\theta)}{p}})$.

En particulier, si $\theta > 1/(p-1)$ alors la somme $\sum_{N \in \mathbb{N}} \left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p}$ est finie.

Démonstration.
Dans une première partie, on utilise des éléments déjà démontrés dans la preuve du lemme 2.2.3 pour construire un couplage bien choisi. En particulier, on reprend ici les notations de la preuve du lemme 2.2.3.

Première étape : Construction d'un couplage

Rappelons que, pour μ_A-presque tout $x \in A$ et tout $N \geq 0$, on a défini une mesure de probabilité sur A par :

$$\tilde{\mu}_x^{(N)} := \sum_{\{y : T_A^N y = x\}} g^{(N)}(y) \delta_y.$$

Ainsi, pour toute fonction $X \in \mathbb{L}^1(A, \mu_A)$ et tout entier $N \geq 0$, on a $\mathcal{L}^N X(x) = \int_A X \, \mathrm{d}\tilde{\mu}_x^{(N)}$. D'après l'équation (2.2.2), il existe des constantes $C > 0$ et $\rho \in (0, 1)$ telles que, pour tous entiers

$N, M \geq 0$, pour tout cylindre $\overline{a}$ de longueur M,

$$\left|\tilde{\mu}_x^{(N+M)}(\overline{a}) - \mu_A(\overline{a})\right| = \left|\mathcal{L}^{N+M} 1_{\overline{a}}(x) - \int_A 1_{\overline{a}} \, \mathrm{d}\mu_A\right| \leq C\mu_A(\overline{a})\rho^N. \tag{2.5.2}$$

Soient $p > 1$ et $\theta > 0$. Soit $X \in \mathbb{L}^p(A, \mu_A)$ une fonction telle que $\int_A X \, \mathrm{d}\mu_A = 0$ et que $\mathbb{E}(\ln_+(D(X))^{1+\theta})$ soit fini. Soit $x \in A$; on note $X_x^{(N+M)}$ la variable aléatoire X définie sur l'espace probabilisé $(A, \tilde{\mu}_x^{(N+M)})$, et par X^* une variable aléatoire de même loi que X définie sur l'espace probabilisé (A, μ_A). On cherche à construire un couplage entre $X_x^{(N+M)}$ et X^* de telle sorte que ces deux variables aléatoires soient proches avec grande probabilité.

Soit $\mathcal{X}^*$ une variable aléatoire à valeurs dans A et de loi μ_A, et $\mathcal{X}_x^{(N+M)}$ une variable aléatoire à valeurs dans A et de loi $\tilde{\mu}_x^{(N+M)}$. Ainsi, en loi, $X^* = X(\mathcal{X}^*)$ et $X_x^{(N+M)} = X(\mathcal{X}_x^{(N+M)})$. D'après l'équation (2.5.2), il existe un couplage entre $\mathcal{X}_x^{(N+M)}$ et $\mathcal{X}^*$ tel que, avec probabilité au moins $1 - C\rho^N$, ces deux variables aléatoire prennent leur valeur dans le même cylindre de longueur M. Ce couplage induit un couplage entre $X_x^{(N+M)}$ et X^*.

Un cylindre $\overline{a} = [a_0, \cdots, a_{M-1}]$ de longueur M a un diamètre d'au plus $\mathrm{Diam}(A)\lambda^{-M}$, donc $X(\overline{a})$ a un diamètre d'au plus $|X|_{\mathrm{Lip}(a_0)} \mathrm{Diam}(A)\lambda^{-M}$. Cette majoration nous permet de contrôler la différence entre $X_x^{(N+M)}$ et X_x^* pourvu que $\mathcal{X}_x^{(N+M)}$ et $\mathcal{X}^*$ prennent leur valeur dans le même cylindre de longueur M. Alors, pour tout $\delta > 0$:

$$\begin{aligned}
\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \delta\right) &\leq C\rho^N + \sum_{a \in \pi} \mu(a) 1\left(|X|_{\mathrm{Lip}(a)} \mathrm{Diam}(A)\lambda^{-M} > \delta\right) \\
&= C\rho^N + \mathbb{P}_{\mu_A}\left(D(X) \mathrm{Diam}(A)\lambda^{-M} > \delta\right).
\end{aligned} \tag{2.5.3}$$

Nous utilisons à présent une inégalité de Markov logarithmique. Si $\lambda^M \mathrm{Diam}(A)^{-1}\delta \geq 1$, alors :

$$\begin{aligned}
\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \delta\right) &\leq C\rho^N + \mathbb{P}_{\mu_A}\left(\ln(1 + D(X))^{1+\theta} > \ln\left(1 + \lambda^M \delta \, \mathrm{Diam}(A)^{-1}\right)^{1+\theta}\right) \\
&\leq C\rho^N + \frac{\mathbb{E}\left(\ln(1 + D(X))^{1+\theta}\right)}{\ln\left(1 + \lambda^M \delta \, \mathrm{Diam}(A)^{-1}\right)^{1+\theta}} \\
&\leq C\rho^N + \frac{\mathbb{E}\left((1 + \ln_+(D(X)))^{1+\theta}\right)}{\ln\left(\lambda^M \delta \, \mathrm{Diam}(A)^{-1}\right)^{1+\theta}} \\
&\leq C\rho^N + \frac{2^\theta\left(1 + \mathbb{E}(\ln_+(D(X))^{1+\theta})\right)}{\ln\left(\lambda^M \delta \, \mathrm{Diam}(A)^{-1}\right)^{1+\theta}}.
\end{aligned}$$

Prenons par exemple $\delta = \mathrm{Diam}(A)\lambda^{-M/2}$. Alors il existe une constante $C \geq 0$ telle que :

$$\mathbb{P}\left(\left|X_x^{(N+M)} - X^*\right| > \mathrm{Diam}(A)\lambda^{-\frac{M}{2}}\right) \leq C\rho^N + CM^{-(1+\theta)}. \tag{2.5.4}$$

Seconde étape : Obtention de la borne $\mathbb{L}^p$

Nous continuons avec les notations de la première partie. Soit $p^* > 1$ l'exposant conjugué de p. On pose $n = N + M$ et on choisit de prendre $N = M$. Soit U l'ensemble sur lequel $\left|X^* - X_x^{(N)}\right| > \mathrm{Diam}(A)\lambda^{-\frac{N}{4}}$. Étant donné que $\mathbb{E}(X^*) = 0$,

$$\begin{aligned}
\left|\mathcal{L}^N X(x)\right| &= \left|\mathbb{E}(X^*) - \mathbb{E}(X_x^{(N)})\right| \\
&\leq \mathbb{E}\left(\left|X^* - X_x^{(N)}\right|\right) \\
&\leq \mathrm{Diam}(A)\lambda^{-\frac{N}{4}} + \left\|1_U(X^* - X_x^{(N)})\right\|_{\mathbb{L}^1} \\
&\leq CN^{-\frac{(1+\theta)}{p^*}}\left(1 + \|X\|_{\mathbb{L}^p} + \left\|X_x^{(N)}\right\|_{\mathbb{L}^p}\right),
\end{aligned}$$

où la dernière majoration est une application de l'inégalité de Hölder. De plus, remarquons que $\left\| X_x^{(N)} \right\|_{\mathbb{L}^p} = \left(\mathcal{L}^N |X|^p\right)^{\frac{1}{p}}(x)$, et prenons la norme $\mathbb{L}^p$ de chaque côté de l'égalité. On rappelle que l'opérateur $\mathcal{L}$ agissant sur $\mathbb{L}^1$ est une contraction.

$$\begin{aligned}
\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p} &\leq CN^{-\frac{(1+\theta)}{p^*}} \left(1 + \|X\|_{\mathbb{L}^p} + \left\| \left\| X_x^{(N)} \right\|_{\mathbb{L}^p} \right\|_{\mathbb{L}^p}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}} \left(1 + \|X\|_{\mathbb{L}^p} + \left\| \left(\mathcal{L}^N |X|^p\right)^{\frac{1}{p}} \right\|_{\mathbb{L}^p}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}} \left(1 + \|X\|_{\mathbb{L}^p} + \left\| \left(\mathcal{L}^N |X|^p\right) \right\|_{\mathbb{L}^1}^{\frac{1}{p}}\right) \\
&\leq CN^{-\frac{(1+\theta)}{p^*}} \left(1 + \|X\|_{\mathbb{L}^p} + \left\| |X|^p \right\|_{\mathbb{L}^1}^{\frac{1}{p}}\right) \\
&= CN^{-\frac{(1+\theta)}{p^*}} \left(1 + 2\|X\|_{\mathbb{L}^p}\right) \\
&= O(N^{-\frac{(p-1)(1+\theta)}{p}}).
\end{aligned}$$

□

Grâce à des méthodes de martingales développées par Gordin [29], la sommabilité de $\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p}$ est suffisante pour démontrer un théorème central limite et l'inégalité de Burkholder-Rosenthal. Nous allons maintenant décrire comment procéder. Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov. Soit X une fonction définie sur A et satisfaisant les hypothèses du lemme 2.5.1 avec un paramètre $\theta > 1/(p-1)$. On pose :

$$C(X) := \sum_{N=1}^{+\infty} \mathcal{L}^N X, \tag{2.5.5}$$

et :

$$\tilde{X} := X + C(X) - C(X) \circ T. \tag{2.5.6}$$

Alors $\tilde{X}$ et X sont cohomologues, et d'après le lemme 2.5.1 la fonction $\tilde{X}$ appartient aussi à $\mathbb{L}^p$. De plus, un calcul rapide montre que $\mathcal{L}\tilde{X} = 0$.

Les fonctions X et $\tilde{X}$ ne diffèrent que par un cobord, dont l'influence sur les sommes partielles sera en général négligeable. Soit $\varepsilon > 0$. Les variables aléatoires $N^{-\varepsilon}C(X)$ et $N^{-\varepsilon}C(X) \circ T_A^N$ convergent vers 0 dans $\mathbb{L}^p$, et $N^{-\varepsilon}C(X)$ converge μ_A-presque sûrement vers 0. De plus, $C(X)$ appartient à $\mathbb{L}^p$, donc $|C(X)|^p$ appartient à $\mathbb{L}^1$, et d'après le théorème ergodique de Birkhoff, $N^{-1}|C(X)|^p \circ T_A^n$ converge presque sûrement vers 0. Par conséquent, $N^{-1/p}C(X) \circ T_A^n$ converge μ_A-presque sûrement vers 0. Pour résumer, le comportement asymptotique de $N^{-\varepsilon}\sum_{i=0}^{N-1} X \circ T_A^i$ et de $N^{-\varepsilon}\sum_{i=0}^{N-1} \tilde{X} \circ T_A^i$ est le même pour tout $\varepsilon > 0$ si l'on travaille en norme $\mathbb{L}^p$, et pour tout $\varepsilon \geq 1/p$ si l'on travaille en convergence presque sûre.

Pour tout $N \in \mathbb{N}$, on définit une tribu $\mathcal{G}_N := T_A^{-N}\mathcal{B}$ sur A. La suite $(\mathcal{G}_N)_{N\geq 0}$ est une filtration décroissante. Pour tout $i > N \geq 0$, on a $\mathbb{E}(\tilde{X} \circ T_A^N | \mathcal{G}_i) = (\mathcal{L}^{i-N}\tilde{X}) \circ T_A^i = 0$. La suite $(\tilde{X} \circ T_A^N)_{N\geq 0}$ est donc une suite inversée de différences de martingale pour la filtration $(\mathcal{G}_N)_{N\geq 0}$. Le théorème central limite en découle immédiatement [34, théorème 3.2] :

Proposition 2.5.2 (Théorème central limite en faible régularité).
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soient $p \geq 2$ et $\theta > 1/(p-1)$. Soit $X \in \mathbb{L}^p(A, \mu_A)$ une fonction telle que $\int_A X \,\mathrm{d}\mu_A = 0$ et $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$. Alors :

$$\lim_{N\to+\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X \circ T_A^i = \mathcal{N}(0, \sigma^2),$$

où la convergence est en loi sur l'espace probabilisé (A, μ_A), et où $\mathcal{N}(0, \sigma^2)$ est une variable aléatoire gaussienne centrée de variance :

$$\sigma^2 = \int_A X^2 \,\mathrm{d}\mu_A + 2\sum_{N=1}^{+\infty} \int_A X \cdot X \circ T_A^N \,\mathrm{d}\mu_A.$$

De plus, $\sigma^2 = 0$ si et seulement si X est un cobord.

Nous allons aussi démontrer une version de l'inégalité de Burkholder-Rosenthal.

Proposition 2.5.3 (Inégalité de Burkholder-Rosenthal en faible régularité).
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov ergodique. Soient $p \geq 2$ et $\theta > 1/(p-1)$. Soit $X \in \mathbb{L}^p(A, \mu_A)$ une fonction telle que $\int_A X \, \mathrm{d}\mu_A = 0$ et $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$. Alors il existe une constante $C \geq 0$ telle que, pour tout $N \geq 0$,

$$\left\| \sum_{i=0}^{N-1} X \circ T_A^i \right\|_{\mathbb{L}^p} \leq C N^{\frac{1}{2}}. \tag{2.5.7}$$

Démonstration.
Nous utilisons les fonctions $C(X)$ et $\tilde{X}$ définie par les équations (2.5.5) et (2.5.6) respectivement. Étant donné que $C(X) \circ T_A - C(X)$ est un cobord,

$$\lim_{N \to +\infty} N^{-\frac{1}{2}} \left\| \sum_{i=0}^{N-1} (C(X) \circ T_A - C(X)) \circ T_A^N \right\|_{\mathbb{L}^p} = 0.$$

Il suffit donc de montrer l'inégalité de Burkholder-Rosenthal en remplaçant X par $\tilde{X}$. Soit N un entier strictement positif. On applique l'inégalité de Burkholder-Rosenthal pour les martingales [17, théorème 21.1] à la suite inversée de différences de martingale $\left(\tilde{X} \circ T_A^i\right)_{0 \leq i \leq N-1}$. Il existe donc une constante $C_p > 0$, qui ne dépend que de p, telle que :

$$\left\| \sum_{i=0}^{N-1} \tilde{X} \circ T_A^i \right\|_{\mathbb{L}^p} \leq C_p \left(\left(\mathbb{E} \left(\sum_{i=0}^{N-1} \mathbb{E}(\tilde{X}^2 \circ T_A^i | \mathcal{G}_{i+1}) \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} + \left(\mathbb{E} \left(\sum_{i=0}^{N-1} \left| \tilde{X} \circ T_A^i \right|^p \right) \right)^{\frac{1}{p}} \right).$$

On utilise tout d'abord l'identité $\mathbb{E}(\tilde{X}^2 \circ T_A^i | \mathcal{G}_{i+1}) = \left(\mathcal{L}\tilde{X}^2\right) \circ T_A^{i+1}$, ce qui nous donne :

$$\left(\mathbb{E} \left(\sum_{i=0}^{N-1} \mathbb{E}(\tilde{X}^2 \circ T_A^i | \mathcal{G}_{i+1}) \right)^{\frac{p}{2}} \right)^{\frac{1}{p}} = \left\| \sum_{i=0}^{N-1} \left(\mathcal{L}\tilde{X}^2\right) \circ T_A^i \right\|_{\mathbb{L}^{\frac{p}{2}}}^{\frac{1}{2}} \leq \left\| \tilde{X} \right\|_{\mathbb{L}^p} N^{\frac{1}{2}}.$$

Le dernier membre de l'inégalité de Burkholder-Rosenthal est transformé de la façon suivante :

$$\left(\mathbb{E} \left(\sum_{i=0}^{N-1} |\tilde{X} \circ T_A^i|^p \right) \right)^{\frac{1}{p}} = \left(N \mathbb{E}(\left|\tilde{X}\right|^p) \right)^{\frac{1}{p}} = \left\| \tilde{X} \right\|_{\mathbb{L}^p} N^{\frac{1}{p}}.$$

Les constantes en jeu ne dépendent pas de l'entier N choisi. □

Le processus $(X \circ T_A^i)$ est stationnaire. Supposons que les conditions de la proposition 2.5.3 sont satisfaites avec un paramètre $p > 2$. De par une inégalité de Serfling [62, corollaire B1], utilisée avec l'inégalité de Burkholder-Rosenthal (2.5.7), il existe une constante $C \geq 0$ telle que, pour tout entier positif N,

$$\left\| \sup_{k \leq N} \sum_{i=0}^{k-1} X \circ T_A^i \right\|_{\mathbb{L}^p} \leq C N^{\frac{1}{2}}.$$

C'est en fait cette dernière inégalité que l'on utilisera le plus souvent.

2.5.2 Adaptation du couplage de Csáki-Földes

Nous allons maintenant affaiblir nos résultat principaux, les théorèmes 2.1.1 et 2.1.3, afin qu'ils ne requièrent qu'une hypothèse de régularité faible. Afin de simplifier les énoncés, nous supposerons que les applications Gibbs-Markov sont toutes mélangeantes, quand bien même une hypothèse d'ergodicité est suffisante de par la partie précédente.

Le théorème central limite et l'inégalité de Burkholder-Rosenthal démontrés précédemment seront tous deux utilisés, donc la condition $\theta > 1/(p-1)$ apparaîtra. D'autres inégalités mettant en jeu les paramètres q et ε gouvernant la construction de processus i.i.d. par morceaux devront être satisfaites ; nous devrons montrer que l'on peut encore choisir ces paramètres de façon cohérente.

Le résultat le plus difficile qu'il nous reste à adapter est le lemme 2.2.3.

Lemme 2.5.4.
Dans le lemme 2.2.3, la condition $\mathbb{E}(D(Y)) < +\infty$ peut être remplacée par la condition plus faible $\mathbb{E}(\ln_+(D(Y))^{1+\theta}) < +\infty$, pourvu que :

$$q(1+\varepsilon\theta) > 1. \tag{2.5.8}$$

Démonstration.
Nous réutilisons les notations de la preuve du lemme 2.2.3. Dans la quatrième étape de celle-ci, nous avions démontré que, si $\mathbb{E}(D(Y)) < +\infty$, alors il existe une constante C telle que, pour tous entiers n et k, il existe un couplage entre $(Y_{i,y})_{i\in I_{n,k}}$ et $(Y^*_{i,y})_{i\in I_{n,k}}$ tel que :

$$\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}}\right) \leq C\max\left\{\rho, \lambda^{-1}\right\}^{2^{q\varepsilon n-2}}.$$

C'est l'équation (2.2.8). On en déduisait alors que :

$$\sum_{n\in\mathbb{N}} 2^{(1-q)n}\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}}\right) < +\infty. \tag{2.5.9}$$

Cette dernière majoration est la clé pour démontrer le lemme 2.2.3. Si l'observable Y ne satisfait que la condition $\mathbb{E}(\ln_+(D(Y))^{1+\theta}) < +\infty$, on ne pourra pas démontrer l'inégalité (2.2.8) : le majorant ne décroît pas sur-exponentiellement en n. Cependant, on peut espérer une majoration moins bonne, mais malgré tout suffisante pour obtenir l'inégalité (2.5.9). Nous utilisons le même couplage entre $(Y_{i,y})_{i\in I_{n,k}}$ et $(Y^*_{i,y})_{i\in I_{n,k}}$ que dans la preuve du lemme 2.2.3. D'après l'inéquation (2.2.6), pour tout $\tau \in (\lambda^{-1}, 1)$, il existe des constantes positives C et C' telles que, pour tout $\delta > 0$,

$$\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \delta\right) \leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}_\gamma\left(|Y|_{\mathrm{Lip}(\gamma)} > C'(\lambda\tau)^{|I_{n,k}|-i}\lambda^{2^{q\varepsilon n-1}}\delta\right),$$

où γ est une variable aléatoire à valeurs dans π et telle que $\mathbb{P}(\gamma = a) = \mu_A(a)$ pour tout a dans π. Cette majoration est similaire à l'inégalité (2.5.3) ; de même que dans la preuve du lemme 2.5.1, nous allons employer une inégalité de Markov logarithmique. Si $C'\lambda^{2^{q\varepsilon n-1}}\delta > 1$, alors :

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i\in I_{n,k}} \left\|Y_{i,y} - Y^*_{i,y}\right\|_{\mathbb{B}} > \delta\right) \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \mathbb{P}_\gamma\left(\ln\left(1+|Y|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta} > \ln\left(1+C'\lambda^{2^{q\varepsilon n-1}}(\lambda\tau)^{|I_{n,k}|-i}\delta\right)^{1+\theta}\right) \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \frac{\mathbb{E}_\gamma\left(\ln\left(1+|Y|_{\mathrm{Lip}(\gamma)}\right)^{1+\theta}\right)}{\ln\left(C'\lambda^{2^{q\varepsilon n-1}}(\lambda\tau)^{|I_{n,k}|-i}\delta\right)^{1+\theta}} \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \sum_{i=0}^{|I_{n,k}|-1} \frac{2\left(1+\mathbb{E}\left((\ln_+ D(Y))^{1+\theta}\right)\right)}{\left(\ln\left(C'\lambda^{2^{q\varepsilon n-1}}\delta\right) + (|I_{n,k}|-i)\ln(\lambda\tau)\right)^{1+\theta}} \\
&\leq C\rho^{2^{q\varepsilon n-1}} + \frac{C''}{\ln\left(C'\lambda^{2^{q\varepsilon n-1}}\delta\right)^{\theta}}.
\end{aligned}$$

Choisissons $\delta = \lambda^{-2^{q\varepsilon n-2}}$, ce qui est possible si n est suffisamment grand. Alors :

$$\mathbb{P}\left(\sum_{i \in I_{n,k}} \left\| Y_{i,y} - Y^*_{i,y} \right\|_{\mathbb{B}} > \lambda^{-2^{q\varepsilon n-2}} \right) = O(2^{-q\varepsilon\theta n}).$$

En sommant cette inégalité pour $0 \leq k < 2^{(1-q)n}$, on voit que la borne (2.5.9) est vraie dès que $1 - q - q\varepsilon\theta < 0$, ce qui est équivalent à la condition $q(1+\varepsilon\theta) > 1$. □

Maintenant que l'on a adapté le théorème central limite, l'inégalité de Burkholder-Rosenthal et le lemme 2.2.3, les démonstrations des théorèmes 2.1.1 et 2.1.3 ne demandent que des modifications mineures. En effet, les autres ingrédients que l'on utilise n'ont rien à voir avec la régularité des observables.

Proposition 2.5.5.
Soit $\theta > 0$.

Dans l'énoncé du théorème 2.1.1, l'hypothèse selon laquelle $\mathbb{E}(D(X)) < +\infty$ et $\mathbb{E}(D(\varphi)) < +\infty$ peut être remplacée par l'hypothèse $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$ et $\mathbb{E}(\ln_+(D(\varphi))^{1+\theta}) < +\infty$, pourvu que :

$$\theta > \frac{1}{p-1}. \tag{2.5.10}$$

Dans l'énoncé du théorème 2.1.3, l'hypothèse selon laquelle $\mathbb{E}(D(X_f)) < +\infty$ peut être remplacée par l'hypothèse $\mathbb{E}(\ln_+(D(X_f))^{1+\theta}) < +\infty$, pourvu que :

$$\theta > \frac{1}{p-1}.$$

Démonstration.
Commençons par démontrer cette nouvelle version du théorème 2.1.1. Nous avons déjà adapté le lemme 2.2.3, au prix de la condition (2.5.8). La régularité des observables n'intervient pas dans la preuve du lemme 2.2.4, et n'intervient dans la preuve du lemme 2.2.5 que *via* l'inégalité de Burkholder-Rosenthal, que l'on a déjà démontrée dans le contexte d'observables de faible régularité (c'est la proposition 2.5.3). La proposition 2.2.6, qui est notre version du couplage de E. Csáki et A. Földes, est de nature purement probabiliste, et n'a donc pas besoin d'être modifiée. Cette proposition peut être employée dès que la suite $(X \circ T_A^i)$ satisfait une inégalité de Burkholder-Rosenthal.

Il reste donc à montrer qu'en appliquant le lemme 2.5.4, la condition supplémentaire (2.5.8) ne nous empêche pas de choisir les paramètre q et ε de façon cohérente. Tout d'abord, on choisit q dans l'intervalle $((1+\theta)^{-1}, 1)$. La condition plus faible $q(1+\theta) > 1$ est alors satisfaite. Ensuite, si l'on choisit ε suffisamment proche de 1, l'inégalité $q(1+\varepsilon\theta) > 1$ est elle aussi satisfaite. Par conséquent, il est bien possible de choisir les paramètre q et ε de façon cohérente. On a donc démontré la première partie de la proposition 2.5.5.

Pour démontrer notre nouvelle version du théorème 2.1.3 sachant que l'on dispose de la nouvelle version du théorème 2.1.1, il suffit de remarquer que la régularité de l'observable X_f n'a d'importance que pour démontrer le théorème central limite et l'inégalité de Burkholder-Rosenthal pour le processus $\left(\sum_{i<N} X_f \circ T_A^i\right)$. Ce sont les propositions 2.5.2 et 2.5.3 démontrées dans la sous-partie précédente. □

2.6 Observables hilbertiennes

La troisième et dernière généralisation de nos théorèmes principaux que nous allons étudier dans ce chapitre est celle des observables qui sont à valeurs non pas réelles, mais dans un espace de Hilbert. L'amélioration est marginale, mais assez facile maintenant que nous avons mis en place des méthodes de martingales. La principale difficulté est de trouver les références adéquates dans la littérature. Comme nous utilisons les mêmes méthodes que dans la partie 2.5, les conditions de régularité que nous utiliserons seront les mêmes, c'est-à-dire les conditions de régularité faibles.

Pour tout espace probabilisé (A, μ_A), pour tout espace de Hilbert $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ et pour tout $p \in [1, \infty]$, on note $\mathbb{L}^p(A, \mu_A; \mathcal{H})$ l'espace vectoriel des fonctions définies μ-presque partout sur A, à valeurs dans $\mathcal{H}$, Bochner-intégrables et avec un moment d'ordre p fini. Cet espace vectoriel est muni de la norme :

$$\|X\|_{\mathbb{L}^p(A,\mu_A;\mathcal{H})} := \left(\int_A \|X\|_{\mathcal{H}}^p \, \mathrm{d}\mu_A\right)^{\frac{1}{p}}.$$

S'il n'y a pas d'ambiguïté sur le choix de la mesure, on pourra écrire $\mathbb{L}^p(A; \mathcal{H})$ au lieu de $\mathbb{L}^p(A, \mu_A; \mathcal{H})$.

Le lecteur pourra se convaincre que, dans toutes les démonstrations que nous avons écrites jusqu'à présent, tout argument qui ne repose pas sur le théorème central limite ou l'inégalité de Burkholder-Rosenthal peut être adapté pour couvrir le cas d'observables à valeurs dans un espace de Hilbert. Il suffit de remplacer les valeurs absolues par des normes $\|\cdot\|_{\mathcal{H}}$, les normes $\|\cdot\|_{\mathbb{L}^p(A,\mu_A)}$ par des normes $\|\cdot\|_{\mathbb{L}^p(A,\mu_A;\mathcal{H})}$, etc. Certains résultats sont déjà conçus pour fonctionner dans un cadre plus général ; par exemple, le lemme 2.5.4 suppose que les observables sont à valeurs dans un espace de Banach. Notre travail dans cette partie consistera uniquement à démontrer des versions du théorème central limite et de l'inégalité de Burkholder-Rosenthal pour des observables à valeurs dans un espace de Hilbert.

Pour commencer, remarquons que le lemme 2.5.1 s'adapte aisément, et on obtient alors le lemme suivant :

Lemme 2.6.1.
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov ergodique. Soit $\mathcal{H}$ un espace de Hilbert. Soit $p > 1$ et soit $X \in \mathbb{L}^p(A, \mu_A; \mathcal{H})$ tel que $\int_A X \, \mathrm{d}\mu_A = 0$. Supposons que $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$ pour un certain réel $\theta > 0$.

Alors $\left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p(A,\mu_A;\mathcal{H})} = O(N^{-\frac{(p-1)(1+\theta)}{p}})$.

En particulier, si $\theta > 1/(p-1)$ alors la somme $\sum_{N \in \mathbb{N}} \left\|\mathcal{L}^N X\right\|_{\mathbb{L}^p(A,\mu_A;\mathcal{H})}$ est finie.

La discussion de la sous-partie 2.5.1, dans laquelle nous avons expliqué comment ajouter un cobord à X et comment définir une filtration $(\mathcal{G}_N)_{N\geq 0}$ pour obtenir des suite inversées de différences de martingales, reste valable. Nous partons de là pour démontrer tout d'abord un théorème central limite, puis une inégalité de Burkholder-Rosenthal.

Théorème 2.6.2 (Théorème central limite dans des espaces de Hilbert)**.**
Soit (A, π, d, μ_A, T_A) une application Gibbs-Markov mélangeante. Soit $\mathcal{H}$ un espace de Hilbert. Soient $p \geq 2$ et $\theta > 1/(p-1)$. Soit $X \in \mathbb{L}^p(A, \mu_A; \mathcal{H})$ une fonction telle que $\int_A X \, \mathrm{d}\mu_A = 0$ et $\mathbb{E}(\ln_+(D(X))^{1+\theta}) < +\infty$. Alors il existe un opérateur positif, symétrique et compact S sur $\mathcal{H}$ tel que :

$$\lim_{N\to+\infty} \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} X \circ T_A^i = \mathcal{N}(0, S),$$

où la convergence est en loi sur l'espace probabilisé (A, μ_A), et où $\mathcal{N}(0, S)$ est une variable aléatoire gaussienne centrée d'opérateur de covariance S. De plus, pour tous vecteurs u et v dans $\mathcal{H}$,

$$(Su, v) = \int_A (X, u)(X, v) \, \mathrm{d}\mu_A + \sum_{i=1}^{+\infty} \int_A \left(X \circ T_A^i, u\right)(X, v) \, \mathrm{d}\mu_A + \sum_{i=1}^{+\infty} \int_A (X, u)\left(X \circ T_A^i, v\right) \, \mathrm{d}\mu_A,$$

et l'opérateur S n'est pas défini positif si et seulement s'il existe un vecteur u non nul tel que (X, u) est un cobord.

Démonstration.
La référence que nous utilisons ici pour obtenir un théorème central limite pour des martingales à valeurs hilbertiennes est un article de A. Jakubowski [44, théorème 5.1]. Pour appliquer ce résultat, nous devons construire un tableau de différences de martingales, et vérifier trois conditions.

De même que dans la sous-partie 2.5.1, on pose $C(X) := \sum_{i=1}^{+\infty} \mathcal{L}^i X$ et $\tilde{X} := X + C(X) - C(X) \circ T_A$. Ensuite, pour tout entier naturel i, on écrit $\tilde{X}_i := \tilde{X} \circ T_A^i$. Alors, pour tout entier N, le processus $(\tilde{X}_{N-i})_{1\leq i\leq N}$ est une suite de différences de martingales pour la filtration $(\mathcal{G}_k)_{k\geq 0}$. Par conséquent, si l'on pose $\tilde{X}_{N,i} := N^{-\frac{1}{2}}\tilde{X}_{N-i}$, alors le processus $(\tilde{X}_{N,i})_{1\leq i\leq N}$ est un tableau de différences de martingales. Nous vérifions à présent les trois conditions présentes dans le théorème de Jakubowski.

Pour commencer, pour tout entier $N > 0$ et pour tout $x > 0$:

$$\mathbb{P}\left(\max_{1\leq i\leq N}\left\|\tilde{X}_{N,i}\right\|_{\mathcal{H}} > x\right) \leq N\mathbb{P}\left(\left\|\tilde{X}\right\|_{\mathcal{H}} > xN^{\frac{1}{2}}\right) \ ;$$

cette inégalité peut être utilisée pour majorer la norme $\mathbb{L}^2$ du maximum du processus :

$$\begin{aligned}
\mathbb{E}\left(\max_{1\leq i\leq N}\left\|\tilde{X}_{N,i}\right\|^2_{\mathcal{H}}\right) &= \int_0^{+\infty}\mathbb{P}\left(\max_{1\leq i\leq N}\left\|\tilde{X}_{N,i}\right\|_{\mathcal{H}} > x\right)\,\mathrm{d}x \\
&\leq \int_0^{+\infty}\min\left\{1, N\mathbb{P}\left(\left\|\tilde{X}\right\|_{\mathcal{H}} > xN^{\frac{1}{2}}\right)\right\}\,\mathrm{d}x \\
&= \frac{1}{\sqrt{N}}\int_0^{+\infty}\min\left\{1, N\mathbb{P}\left(\left\|\tilde{X}\right\|_{\mathcal{H}} > x\right)\right\}\,\mathrm{d}x.
\end{aligned}$$

Soit $\varepsilon > 0$, et soit $M > 0$ tel que $\mathbb{P}\left(\left\|\tilde{X}\right\|_{\mathcal{H}} > x\right) \leq \varepsilon x^{-2}$ pour tout $x \geq M$. Alors :

$$\begin{aligned}
\mathbb{E}\left(\max_{1\leq i\leq N}\left\|\tilde{X}_{N,i}\right\|^2_{\mathcal{H}}\right) &\leq \frac{M}{\sqrt{N}} + \frac{1}{\sqrt{N}}\int_M^{+\infty}\min\left\{1, \varepsilon N x^{-2}\right\}\,\mathrm{d}x \\
&= \frac{\max\{M, \sqrt{\varepsilon N}\}}{\sqrt{N}} + \frac{1}{\sqrt{N}}\int_{\max\{M,\sqrt{\varepsilon N}\}}^{+\infty}\varepsilon N x^{-2}\,\mathrm{d}x \\
&\leq \frac{\max\{M, \sqrt{\varepsilon N}\}}{\sqrt{N}} + \frac{\varepsilon\sqrt{N}}{\max\{M, \sqrt{\varepsilon N}\}},
\end{aligned}$$

de telle sorte que $\limsup_{N\to+\infty}\mathbb{E}\left(\max_{1\leq i\leq N}\left\|\tilde{X}_{N,i}\right\|^2_{\mathcal{H}}\right) \leq 2\sqrt{\varepsilon}$. Cette inégalité étant vraie pour tout $\varepsilon > 0$, la première condition est satisfaite.

Passons à la deuxième condition. Pour tous vecteurs u et v dans $\mathcal{H}$,

$$\sum_{i=1}^N\left(\tilde{X}_{N,i}, u\right)\left(\tilde{X}_{N,i}, v\right) = \frac{1}{N}\sum_{i=1}^N\left[\left(\tilde{X}, u\right)\left(\tilde{X}, v\right)\right]\circ T_A^i.$$

Par le théorème ergodique de Birkhoff, cette quantité converge presque sûrement vers une constante quand N tend vers $+\infty$.

Finalement, soit $(e_j)_{j\in\mathbb{N}}$ une base de Hilbert de $\mathcal{H}$, soit M un entier naturel et soit $\varepsilon > 0$. On calcule :

$$\begin{aligned}
\mathbb{P}\left(\sum_{1\leq i\leq N}\sum_{j\geq M}\left(\tilde{X}_{N,i}, e_j\right)^2 > \varepsilon\right) &= \mathbb{P}\left(\frac{1}{N}\sum_{1\leq i\leq N}\left(\sum_{j\geq M}\left(\tilde{X}, e_j\right)^2\right)\circ T_A^i > \varepsilon\right) \\
&\underset{N\to+\infty}{\to} 1\left[\sum_{j\geq M}\mathbb{E}\left(\left(\tilde{X}, e_j\right)^2\right) > \varepsilon\right] \underset{M\to+\infty}{\to} 0.
\end{aligned}$$

Toutes les hypothèses du théorème central limite sont donc vérifiées. Le processus $N^{-\frac{1}{2}}\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i$ converge en loi vers une variable aléatoire gaussienne quand N tend vers l'infini. L'ajout d'un cobord ne change en rien la limite du processus, donc le processus $N^{-\frac{1}{2}}\sum_{i=0}^{N-1}X\circ T_A^i$ converge en loi vers la même variable aléatoire gaussienne. Il nous reste à trouver l'expression de l'opérateur de covariance S. D'après le théorème de Jakubowski que nous avons utilisé, pour tous vecteurs u et u dans $\mathcal{H}$,

$$(Su, v) = \lim_{N\to+\infty}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i, u\right)\left(\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i, v\right)\right].$$

On commence par démontrer que l'on peut remplacer $\tilde{X}$ par X dans cette formule. En effet,

$$\begin{aligned}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i,u\right)\left(\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i,v\right)\right] &= \frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}X\circ T_A^i,u\right)\left(\sum_{i=0}^{N-1}X\circ T_A^i,v\right)\right]\\ &\quad+\mathbb{E}\left[\left(C(X)-C(X)\circ T_A^N,u\right)\left(\frac{1}{N}\sum_{i=0}^{N-1}\tilde{X}\circ T_A^i,v\right)\right]\\ &\quad+\mathbb{E}\left[\left(\frac{1}{N}\sum_{i=0}^{N-1}X\circ T_A^i,u\right)\left(C(X)-C(X)\circ T_A^N,v\right)\right].\end{aligned}$$

Étant donné que les processus $N^{-1}\sum_{i=0}^{N-1}(X,u)\circ T_A^i$ et $N^{-1}\sum_{i=0}^{N-1}(\tilde{X},v)\circ T_A^i$ convergent vers 0 dans $\mathbb{L}^2$, les deux dernières lignes tendent elles aussi vers 0 quand N tend vers l'infini. Ainsi,

$$\begin{aligned}(Su,v) &= \lim_{N\to+\infty}\frac{1}{N}\mathbb{E}\left[\left(\sum_{i=0}^{N-1}X\circ T_A^i,u\right)\left(\sum_{i=0}^{N-1}X\circ T_A^i,v\right)\right]\\ &= \lim_{N\to+\infty}\frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\mathbb{E}\left[\left(X\circ T_A^i,u\right)\left(X\circ T_A^j,v\right)\right]\\ &= \lim_{N\to+\infty}\left[\mathbb{E}\left((X,u)\,(X,v)\right)+\sum_{i=1}^{N-1}\left(1-\frac{i}{N}\right)\mathbb{E}\left(\left(X\circ T_A^i,u\right)(X,v)\right)\right.\\ &\qquad\left.+\sum_{i=1}^{N-1}\left(1-\frac{i}{N}\right)\mathbb{E}\left((X,u)\left(X\circ T_A^i,v\right)\right)\right]\\ &= \mathbb{E}\left[(X,u)\,(X,v)\right]+\sum_{i=1}^{+\infty}\mathbb{E}\left[\left(X\circ T_A^i,u\right)(X,v)\right]+\sum_{i=1}^{+\infty}\mathbb{E}\left[(X,u)\left(X\circ T_A^i,v\right)\right],\end{aligned}$$

la dernière égalité découlant du théorème de convergence dominée. Pour finir, on se ramène au cas de la dimension 1 pour montrer que l'opérateur S est dégénéré si et seulement s'il existe un vecteur u non nul tel que (X,u) est un cobord. □

Pour l'inégalité de Burkholder-Rosenthal pour des martingales à valeurs dans un espace de Hilbert, nous renvoyons le lecteur à [56, théorème 8.33].

Le théorème central limite et l'inégalité de Burkholder-Rosenthal étaient les seuls ingrédients manquants pour montrer de nouvelles versions des théorèmes 2.1.1 et 2.1.3. Les voici :

Proposition 2.6.3.

Soit (A,π,d,μ_A,T_A) une application Gibbs-Markov ergodique. Soit $\mathcal{H}$ un espace de Hilbert. Soient X et φ des fonctions mesurables définies sur A et à valeurs dans $\mathcal{H}$ et $\mathbb{R}_+$ respectivement. On pose $(X_i,\varphi_i) := (X\circ T_A^i,\varphi\circ T_A^i)$. Soient $(\tilde{X}_i)_{i\in\mathbb{N}}$ et $(\tilde{\varphi}_i)_{i\in\mathbb{N}}$ des copies indépendantes des processus $(X_i)_{i\in\mathbb{N}}$ et $(\varphi_i)_{i\in\mathbb{N}}$ respectivement.

Soient $p>2$ et $\theta>1/(p-1)$. Supposons que $X\in\mathbb{L}^p(A,\mu_A;\mathcal{H})$, que φ satisfait la condition (1.3.2) pour un paramètre $\beta\in[0,1)$ et une fonction auxiliaire ψ, et que $\mathbb{E}(\ln_+(D(X))^{1+\theta})$ et $\mathbb{E}(\ln_+(D(\varphi))^{1+\theta})$ sont finis.

Alors il existe $r\in[0,1)$ et un couplage entre $(X_i,\varphi_i)_{i\in\mathbb{N}}$ et $(\tilde{X}_i,\tilde{\varphi}_i)_{i\in\mathbb{N}}$ tel que, presque sûrement, pour tout entier N suffisamment grand,

$$\left\|\sum_{i=0}^{N-1}X_i-\sum_{i=0}^{N-1}\tilde{X}_i\right\|_{\mathcal{H}}\le N^{\frac{r}{2}},$$
$$\left|\sum_{i=0}^{N-1}\varphi_i-\sum_{i=0}^{N-1}\tilde{\varphi}_i\right|\le\psi^*(N^r).$$

Proposition 2.6.4.

Soit (Ω, μ, T) un système dynamique préservant la mesure, ergodique et conservatif qui induit une application Gibbs-Markov mélangeante sur un sous-ensemble borélien A. Supposons que la fonction $\psi \ : \ t \mapsto \mu_{|A}(\varphi_A > t)^{-1}$ est à variation régulière d'indice $\beta \in [0,1)$.

Soient $p > 2$ et $\theta > 1/(p-1)$. Soit $\mathcal{H}$ un espace de Hilbert, et soit $f \in \mathbb{L}^1(\Omega, \mu; \mathcal{H})$. Supposons que $\int_\Omega f \, \mathrm{d}\mu = 0$, que la variable aléatoire $X_{\|f\|_{\mathcal{H}}}$ appartient à $\mathbb{L}^p(A, \mu_{|A})$ et que $\mathbb{E}(\ln_+(D(X_f))^{1+\theta})$ est fini. Alors :

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(N)}} \sum_{i=0}^{N-1} f \circ T^i \to \sqrt{Y_\beta}\mathcal{N}(0,S) \text{ fortement en distribution,} \tag{2.6.1}$$

où Y_β et $\mathcal{N}(0,S)$ sont des variables aléatoires indépendantes, Y_β suit une loi de Mittag-Leffler standard de paramètre β, et $\mathcal{N}(0,S)$ est une variable aléatoire gaussienne à valeurs dans $\mathcal{H}$ d'opérateur de covariance S. De plus, pour tous vecteurs u et v dans $\mathcal{H}$,

$$(Su,v) = \int_A (X_f, u)\,(X_f, v) \ \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_A \left(X_f \circ T_A^i, u\right)(X_f, v) \ \mathrm{d}\mu + \sum_{i=1}^{+\infty} \int_A (X_f, u) \left(X_f \circ T_A^i, v\right) \ \mathrm{d}\mu.$$

L'opérateur S est dégénéré si et seulement s'il existe un vecteur non nul u dans $\mathcal{H}$ tel que (f,u) est un cobord.

Un résultat similaire peut s'obtenir dans le cas $\beta = 1$.

2.7 Exemples en temps discret

Les deux exemples que nous abordons dans ce chapitre sont les marches aléatoires (voir la sous-partie 1.5.3) et les applications de Pomeau-Manneville (voir la sous-partie 1.5.2).

Les résultats que nous fournissons pour les marches aléatoires généralisent ceux de [22] et [23]. Nous ne traitons pas le cas des $\mathbb{Z}^d$ extensions de systèmes Gibbs-Markov ; nous renvoyons le lecteur intéressé vers la partie 3.3, où nous obtenons des résultats dans le cadre plus général des $\mathbb{Z}^d$ extensions de semi-flots Gibbs-Markov. Dans le cas des marches aléatoires, traditionnellement, l'observable f ne dépend que de la position sur le réseau $\mathbb{Z}^d$ (et non du futur). La condition de régularité dans le théorème 2.1.3 est alors automatiquement satisfaite. De plus, pour des marches aléatoires, le paramètre β associé à la queue du temps de premier retour est à variation régulière d'indice $\beta \in [0, 1/2]$; le théorème 2.1.4 n'a donc aucune utilité dans ce contexte.

Pour les applications de Pomeau-Manneville, nous obtiendrons des conditions de régularité et d'intégrabilité explicites qui garantissent que l'on puisse appliquer le théorème 2.1.3. De plus, le théorème 2.1.4 sera utile.

2.7.1 Marches aléatoires

Application du théorème limite

Dans le contexte des marches aléatoires, il est plus habituel de travailler non pas avec les queues du temps de premier retour, mais avec la fonction de Green (ou, dans notre cas, avec la fonction de Green tronquée). Nous adoptons ce vocabulaire afin que la proposition 2.7.2 soit formellement plus proche des théorème démontrés par E. Csáki et A. Földes.

Définition 2.7.1 (Fonction de Green tronquée).

La fonction de Green tronquée associée à une marche aléatoire $(S_n)_{n \geq 0}$, que l'on notera g, est définie pour tout entier positif N par :

$$g(N) := \sum_{i=0}^{N-1} \mu(S_0 = 0 \text{ et } S_i = 0),$$

où μ est la mesure invariante sur $(\mathbb{Z}^d)^{\mathbb{N}}$ correspondant à la marche aléatoire.

Si (S_n) est une marche aléatoire sur $\mathbb{Z}^d$ partant de 0, alors :

$$g(N) = \sum_{i=0}^{N-1} \mathbb{P}(S_i = 0).$$

Une marche aléatoire sur $\mathbb{Z}^d$ est récurrente si et seulement si la fonction de Green tronquée tend vers l'infini en l'infini [72, proposition 1.7]. En d'autres termes, si l'on considère la suite d'évènements $(A_n) := (\{S_n = 0\})$ et si l'on suppose que la marche aléatoire part de 0, alors $\mathbb{P}(\limsup A_n) = 1$ si et seulement si $\sum_n \mathbb{P}(A_n) = +\infty$. Ce résultat est donc comparable au lemme de Borel-Cantelli, bien que les évènements (A_n) ne soient pas indépendants. Si en plus la marche aléatoire n'est pas dégénérée, alors le système dynamique défini en sous-partie 1.5.3 préserve la mesure, est ergodique et conservatif, et induit un système Gibbs-Markov. De plus, on peut déduire le comportement asymptotique des queues du temps de premier retour en 0 des asymptotiques de $(g(N))$ (et réciproquement). On peut donc appliquer le théorème 2.1.3.

Proposition 2.7.2.
Soit $d \geq 1$, et soit (S_n) une marche aléatoire sur $\mathbb{Z}^d$. Supposons que la fonction de Green tronquée est à variation régulière d'indice $\beta \in [0, 1)$ en l'infini, et tend vers $+\infty$ en $+\infty$. Supposons de plus que le noyau de transition de la marche aléatoire n'est pas supporté par un sous-groupe strict de $\mathbb{Z}^d$.

Soit (S'_n) la marche aléatoire avec le même noyau de transition, mais partant de 0. Soit φ le temps de premier retour de (S'_n) en 0.

Soit $f \in \ell^1(\mathbb{Z}^d)$ telle que $\sum_{k\in\mathbb{Z}^d} f(k) = 0$ et, pour un certain $p > 2$:

$$\mathbb{E}\left(\left(\sum_{i=0}^{\varphi-1} |f(S'_i)|\right)^p\right) < +\infty. \tag{2.7.1}$$

Alors, pour toute loi de probabilité sur $\mathbb{Z}^d$ pour le départ de la marche aléatoire,

$$\frac{1}{\sqrt{g(N)}} \sum_{i=0}^{N-1} f(S_i) \to \left\| \sum_{i=0}^{\varphi-1} f(S'_i) \right\|_{\mathbb{L}^2} \sqrt{Y_\beta}\mathcal{N}, \tag{2.7.2}$$

où la convergence est en loi, où Y_β et $\mathcal{N}$ sont indépendants, Y_β suit une loi de Mittag-Leffler standard de paramètre β et où $\mathcal{N}$ suit une loi normale standard.

Cette proposition généralise légèrement [23, théorème 2]. En effet, dans cet articles, les auteurs utilisaient des résultats déjà existants sur les lois du logarithme itéré pour en déduire l'équivalent du lemme 2.3.2. Il restait à la fin une hypothèse technique sur une loi du logarithme itéré que doit vérifier la marche aléatoire dans le théorème 2. Nous avons dû utiliser des techniques moins fines, mais qui évitent ce recours aux lois du logarithme itéré ; cette hypothèse technique n'est donc plus nécessaire.

De plus, l'espace d'état du système dynamique sous-jacent est $(\mathbb{Z}^d)^{\mathbb{N}}$ ou $\mathcal{A}^{\mathbb{N}} \times \mathbb{Z}^d$, où $\mathcal{A}$ est un alphabet dénombrable. L'observable f et la loi initiale de la marche aléatoire peuvent donc dépendre de l'état futur de la marche aléatoire, et non seulement de l'état présent. Les restrictions sont que la loi initiale doit toujours être absolument continue par rapport à la mesure correspondant à la marche aléatoire sur $(\mathbb{Z}^d)^{\mathbb{N}}$, et l'observable doit être suffisamment régulière (voir par exemple le lemme 3.3.2).

Supposons que la marche aléatoire vérifie les hypothèses de la proposition 2.7.2. Si $d = 2$ et le noyau de transition est de variance finie, alors les hypothèses de cette proposition sont vérifiées avec $\beta = 0$, et $g(N) \sim C \ln N$ (voir [43, théorème 1.1]). C'est aussi le cas si le noyau de transition est de variance infinie mais avec des queues pas trop lourdes, comme par exemple s'il existe un automorphisme M de $\mathbb{R}^2$ tel que :

$$\mathbb{P}(\langle w, X_0\rangle \geq t) \sim \frac{\|Mw\|^2}{t^2} \; \forall w \in \mathbb{R}^2 - \{0\},$$

auquel cas $g(N) \sim C \ln(\ln N)$. Ce dernier cas est très similaire aux gaz de Lorentz sur $\mathbb{R}^2$ à horizon infini.

Si $d = 1$ et le noyau de transition est de variance finie, alors les hypothèses de cette proposition sont vérifiées avec $\beta = 1/2$, et $g(N) \sim C\sqrt{N}$ (voir encore [43, théorème 1.1]). Rappelons que Y_0 est une variable aléatoire de loi exponentielle, et que $Y_{1/2}$ est la valeur absolue d'une variable aléatoire de loi normale centrée : dans ces deux cas, la limite en loi s'exprime avec des variables aléatoires usuelles.

Plus généralement, si $d = 1$ et le noyau de transition est dans le bassin d'attraction d'une loi de Lévy stable de paramètre $\alpha \in (1, 2]$, alors les hypothèses de la proposition sont vérifiées avec $\beta = 1 - 1/\alpha$ (voir [43, proposition 6.1]). Si $\alpha = 1$, il n'est pas garanti que la marche aléatoire soit récurrente, mais si elle l'est alors on peut appliquer notre proposition. Le calcul de g en fonction des données de la marche aléatoire sera discuté plus en détail dans la partie 3.2.

Nous donnerons une condition suffisante sur f pour que les hypothèses de la proposition 2.7.2 concernant cette fonction soient satisfaites : la condition (2.7.1) est satisfaite pour tout $p < +\infty$ si la marche aléatoire est ergodique, récurrente et si f est à support fini. Dans ce cas, la marche aléatoire (S_n) partant de 0 induit une chaîne de Markov irréductible sur $\{0\} \cup \mathrm{Supp}(f)$. Étant donné que l'ensemble $\{0\} \cup \mathrm{Supp}(f)$ est fini, le temps de premier retour pour la chaîne de Markov induite a donc des queues décroissant à vitesse exponentielle. Par conséquent, le nombre de fois que $|f(S_i)|$ n'est pas nul dans une excursion a aussi des queues exponentielles, tout comme $\sum_{i=0}^{\varphi_A - 1} f(S_i)$. Cette dernière variable aléatoire est donc dans $\mathbb{L}^p$ pour tout p fini.

Formule de la variance

En général, on ne dispose pas de meilleure formule pour la constante dans la loi limite que $\left\| \sum_{i=0}^{\varphi_A - 1} f(S_i) \right\|_{\mathbb{L}^2}$. Cependant, comme nous l'avons mentionné dans l'introduction, si S_n est la marche aléatoire simple sur $\mathbb{Z}$ et si f est à support fini, alors [24, théorème 2] :

$$\mathbb{E}\left(\left(\sum_{i=0}^{\varphi_A - 1} f(S_i) \right)^2 \right) = 4 \sum_{p \in \mathbb{Z}} p f^2(p) + 8 \sum_{\substack{(p,q) \in \mathbb{Z}^2 \\ p < q}} p f(p) f(q) - \sum_{p \in \mathbb{Z}} f^2(p) =: d(f)^2. \tag{2.7.3}$$

La preuve de ce théorème repose sur la structure particulière de la marche aléatoire simple sur $\mathbb{Z}$. Nous allons maintenant démontrer une formule similaire pour n'importe quelle marche aléatoire récurrente, et retrouver la formule de Dobrushin dans le cas de la marche aléatoire simple. Soit (S_n) une marche aléatoire récurrente sur $\mathbb{Z}^d$. Pour tout $q \in \mathbb{Z}^d$, on pose :

$$\left\{ \begin{array}{lll} N(q) & := & \mathrm{Card}\,\{k \ : \ 0 \leq k < \varphi_A - 1, \ S_k = q\} \\ \mathbb{P}(q) & := & \mathbb{P}(N(q) \neq 0) \end{array} \right. .$$

La variable aléatoire $N(q)$ est le temps local en q sur une excursion partant de 0, et $\mathbb{P}(q)$ est la probabilité de passer par q lors d'une excursion partant de 0. Par le théorème de Hopf, $\mathbb{E}(N(q)) = 1$ pour tout q.

Considérons une excursion partant de 0. Soit (τ_n) la suite des temps de retour de la marche aléatoire simple en $\{0, q\}$. Alors $(S_{\tau_n})_{n \geq 0}$ est une chaîne de Markov sur $\{0, q\}$, de matrice de transition :

$$\begin{pmatrix} 1 - \mathbb{P}(q) & \mathbb{P}(q) \\ \mathbb{P}(-q) & 1 - \mathbb{P}(-q) \end{pmatrix}.$$

La loi de $N(q)$ est alors donnée par :

$$\left\{ \begin{array}{lll} \mathbb{P}(N(q) = 0) & = & 1 - \mathbb{P}(q) \\ \mathbb{P}(N(q) = n) & = & \mathbb{P}(q)\mathbb{P}(-q)(1 - \mathbb{P}(-q))^{n-1} \ \forall n \geq 1 \end{array} \right. .$$

Pour tout q, la variable aléatoire $N(q)$ vaut 0 avec probabilité $1 - \mathbb{P}(q)$. De plus, conditionnée par l'évènement "$\mathbb{P}(q) \neq 0$", la variable aléatoire $N(q) - 1$ suit un loi géométrique de paramètre

$\mathbb{P}(-q)$. Le premier moment de $N(q)$ vaut donc :

$$\mathbb{E}(N(q)) = \mathbb{P}(q) \sum_{n=1}^{+\infty} \mathbb{P}(-q)(1-\mathbb{P}(-q))^{n-1} n = \frac{\mathbb{P}(q)}{\mathbb{P}(-q)} \ ;$$

or $\mathbb{E}(N(q)) = 1$, donc $\mathbb{P}(q) = \mathbb{P}(-q)$ pour tout $q \in \mathbb{Z}^d$. Le deuxième moment de $N(q)$ vaut :

$$\mathbb{E}(N(q)^2) = \mathbb{P}(q) \sum_{n=1}^{+\infty} \mathbb{P}(q)(1-\mathbb{P}(q))^{n-1} n^2 = \mathbb{P}(q) \frac{2-\mathbb{P}(q)}{\mathbb{P}(q)^2} = \frac{2-\mathbb{P}(q)}{\mathbb{P}(q)}.$$

Passons maintenant au calcul de $d(f)$. Formellement, on a :

$$\begin{aligned}\mathbb{E}\left(\left(\sum_{i=0}^{\varphi_A - 1} f(S_i)\right)^2\right) &= \mathbb{E}\left(\left(\sum_{q\in\mathbb{Z}^d} N(q) f(q)\right)^2\right) \\ &= \sum_{q\in\mathbb{Z}^d} \mathbb{E}(N(q)^2) f(q)^2 + \sum_{\substack{p,q\in\mathbb{Z}^d \\ p\neq q}} \mathbb{E}(N(p)N(q)) f(p) f(q).\end{aligned}$$

De plus, la loi limite ne dépend pas du système sur lequel on induit. La formule que l'on obtient doit donc être invariante par translation. En particulier, donnons-nous des points p et q dans $\mathbb{Z}^d$, avec $p \neq q$. Prenons la fonction f valant -1 en p, et 1 en q. Alors :

$$\mathbb{E}\left(\left(\sum_{i=0}^{\varphi_A - 1} f(S_i)\right)^2\right) = \mathbb{E}(N(p)^2) + \mathbb{E}(N(q)^2) - 2\mathbb{E}(N(p)N(q)).$$

Mais le résultat doit être le même que pour la fonction valant -1 en 0 et 1 en $q-p$. Par conséquent :

$$\mathbb{E}(N(p)^2) + \mathbb{E}(N(q)^2) - 2\mathbb{E}(N(p)N(q)) = \mathbb{E}(N(0)^2) + \mathbb{E}(N(q-p)^2) - 2\mathbb{E}(N(0)N(q-p)).$$

Remarquons que $N(0) = 1$. Le seul terme inconnu dans l'égalité ci-dessus est $\mathbb{E}(N(p)N(q))$, qui vaut donc :

$$\mathbb{E}(N(p)N(q)) = \frac{1}{2}\left(2 + \mathbb{E}(N(p)^2) + \mathbb{E}(N(q)^2) - 1 - \mathbb{E}(N(q-p)^2)\right) = \frac{1}{\mathbb{P}(p)} + \frac{1}{\mathbb{P}(q)} - \frac{1}{\mathbb{P}(q-p)}.$$

Finalement, on a donc :

$$\mathbb{E}\left(\left(\sum_{i=0}^{\varphi_A - 1} f(S_i)\right)^2\right) = \sum_{p,q\in\mathbb{Z}^d} \left(\frac{1}{\mathbb{P}(p)} + \frac{1}{\mathbb{P}(q)} - \frac{1}{\mathbb{P}(q-p)}\right) f(p) f(q). \tag{2.7.4}$$

Par le théorème de Fubini-Lebesgue, cette formule a un sens dès que $X_{|f|} \in \mathbb{L}^2$ ou que la formule de droite appliquée à $|f|$ donne un résultat fini.

Pour la marche aléatoire simple sur $\mathbb{Z}$, on a $\mathbb{P}(0) = 1$ et un argument de martingale montre que $\mathbb{P}(q) = 1/(2|q|)$ pour tout $q \neq 0$. L'équation (2.7.4) devient alors :

$$\mathbb{E}\left(\left(\sum_{i=0}^{\varphi_A - 1} f(S_i)\right)^2\right) = \sum_{q\in\mathbb{Z}^d} (4|q|-1) f(q)^2 + 4 \sum_{\substack{p,q\in\mathbb{Z} \\ p<q}} \left(|p| + |q| - |p-q|\right) f(p) f(q). \tag{2.7.5}$$

Ce n'est pas exactement la même formule que celle obtenue par Dobrushin, mais elle donne toujours le même résultat. Il existe une infinité de formules valables, toutes égales modulo la relation $\left(\sum_{\mathbb{Z}^d} f\right)^2 = 0$.

Une conjecture de R.L. Dobrushin

Le théorème de Dobrushin suppose que la fonction f dont on étudie les sommes de Birkhoff est à support fini. Après avoir énoncé ce théorème, R.L. Dobrushin conjectura la proposition suivante :

Conjecture 2.7.3 (Dobrushin, 1955).
Dans le théorème 0.0.4, on peut remplacer l'hypothèse "f est à support fini" par l'hypothèse "$d(f)$ est fini", où $d(f)$ est donné par l'équation (2.7.3).

Cette conjecture est à mettre en parallèle du théorème central limite, qui affirme la convergence en loi vers une gaussienne dès que la variance des variables aléatoires concernées est finie. Un premier problème consiste à interpréter l'expression "$d(f)$ est fini". En effet, la formule (2.7.3) peut faire apparaître des compensations de signes, qui rendent sa définition non triviale.

Nous proposons donc la conjecture plus faible suivante :

Conjecture 2.7.4.
Dans le théorème 0.0.4, on peut remplacer l'hypothèse "f est à support fini" par l'hypothèse "$d'(|f|)$ est fini", où d' est donné par le membre de droite de l'équation (2.7.5).

Cette proposition équivaut à montrer que, au moins dans le cas de la marche aléatoire simple, le théorème 2.1.3 reste valable en remplaçant l'hypothèse "$X_{|f|} \in \mathbb{L}^p$ pour un $p > 2$" par l'hypothèse plus faible "$X_{|f|} \in \mathbb{L}^2$". Elle n'est pas abordable avec les outils développés dans cette thèse. Cependant, les méthodes spectrales de la partie 3.2 pourraient permettre de résoudre cette question (voir la remarque 3.2.16).

2.7.2 Applications de Pomeau-Manneville

Nous abordons maintenant les conséquences de nos théorèmes principaux pour les applications de Pomeau-Manneville. Pour tout sous-ensemble $\mathbb{B} \subset [0,1]$ et pour tout $h \in [0,1]$, on note la semi-norme h-höldérienne sur B par $|\cdot|_{\mathrm{Hol}_h(B)}$. On va démontrer, à partir du théorème 2.1.3, le corollaire suivant :

Corollaire 2.7.5.
Soient $\alpha > 1$ et $h \in (0,1]$. Soit f une fonction définie sur $(0,1]$, à valeurs réelles, localement h-höldérienne, et telle que :
- $f(x) = O(x^{\alpha - \frac{1}{2} + \varepsilon})$ *en* 0 *pour un certain* $\varepsilon > 0$ *;*
- $|f|_{\mathrm{Hol}_h((0,x))} = O(x^{\alpha - 1 + \varepsilon'})$ *en* 0 *pour un certain* $\varepsilon' > 0$ *;*
- $\int_{(0,1]} f \, \mathrm{d}\mu_\alpha = 0$.

Alors il existe $\sigma(f) \geq 0$ tel que, pour toute mesure de probabilité ν absolument continue par rapport à la mesure de Lebesgue,

$$\frac{1}{N^{\frac{1}{2\alpha}}} \sum_{k=0}^{N-1} f(T_\alpha^k x) \to \sigma(f) \sqrt{Y_{\frac{1}{\alpha}}} \mathcal{N},$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur $((0,1],\nu)$, où $Y_{1/\alpha}$ et $\mathcal{N}$ sont des variables aléatoires indépendantes, $Y_{1/\alpha}$ suit une loi de Mittag-Leffler standard de paramètre $1/\alpha$, et $\mathcal{N}$ est une variable aléatoire gaussienne standard. De plus, $\sigma(f) = 0$ si et seulement si f est un cobord.

Comme d'habitude, le cas limite $\alpha = 1$ donne lieu à un énoncé distinct, où le théorème limite est un théorème central limite non standard :

Corollaire 2.7.6.
Soit $h \in (0,1]$. Soit f une fonction définie sur $(0,1]$, à valeurs réelles, localement h-höldérienne, et telle que :
- $f(x) = O(x^{\frac{1}{2} + \varepsilon})$ *en* 0 *pour un certain* $\varepsilon > 0$ *;*
- $|f|_{\mathrm{Hol}_h((0,x))} = O(x^{\varepsilon'})$ *en* 0 *pour un certain* $\varepsilon' > 0$ *;*
- $\int_{(0,1]} f \, \mathrm{d}\mu_1 = 0$.

Alors il existe $\sigma(f) \geq 0$ tel que, pour toute mesure de probabilité ν absolument continue par rapport à la mesure de Lebesgue,

$$\sqrt{\frac{\ln N}{N}} \sum_{k=0}^{N-1} f(T_1^k x) \to \sigma(f)\mathcal{N},$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur $((0,1],\nu)$, et où $\mathcal{N}$ suit une loi normale standard. De plus, $\sigma(f) = 0$ si et seulement si f est un cobord.

En particulier, toute fonction höldérienne à support compact dans $(0,1]$ et d'intégrale nulle vérifie les hypothèses de ces corollaires.

Démonstration.
Comme dans la sous-partie 1.5.2, on pose $A := (1/2, 1]$ et $A_n := \{x \in A : \varphi_A(x) = n+1\}$ pour tout $n \geq 0$. D'après [28], il existe une constante C_α telle que, pour tout $n \geq 1$:

$$\mathrm{Leb}(A_n) \leq \frac{C_\alpha}{n^{1+\frac{1}{\alpha}}}. \tag{2.7.6}$$

Soit f une fonction satisfaisant les hypothèses du corollaire 2.7.5.

Première étape : Condition d'intégrabilité

On commencer par démontrer qu'il existe $p > 2$ tel que $X_{|f|} \in \mathbb{L}^p(A, \mu_{\alpha|A})$. La densité de μ_α par rapport à la mesure de Lebesgue a une version localement lipschitzienne sur $(0,1]$, et donc en particulier bornée sur A. Il suffit donc de montrer que $X_{|f|} \in \mathbb{L}^p(A, \mathrm{Leb})$.

On peut toujours supposer que $\varepsilon < 1/2$. Grâce à la combinatoire des applications de Pomeau-Manneville et aux hypothèses sur f,

$$\sup_{A_n} X_{|f|} \leq \sum_{k=0}^{n} \sup_{T_\alpha A_k} |f| \leq C \sum_{k=0}^{n} \left(k^{-\frac{1}{\alpha}}\right)^{\alpha - \frac{1}{2} + \varepsilon} \leq C n^{\frac{1}{2\alpha} - \frac{\varepsilon}{\alpha}}.$$

D'après l'équation (2.7.6),

$$\left\|X_{|f|}\right\|_{\mathbb{L}^p}^p \leq \sum_{n=0}^{+\infty} \mathrm{Leb}(A_n) \sup_{A_n} X_{|f|}^p \leq C + C' \sum_{n=1}^{+\infty} n^{\frac{p}{2\alpha} - \frac{p\varepsilon}{\alpha} - 1 - \frac{1}{\alpha}} = C + C' \sum_{n=1}^{+\infty} n^{-1 - \frac{2 - p(1-2\varepsilon)}{2\alpha}},$$

et donc $X_{|f|} \in \mathbb{L}^p(A, \mathrm{Leb})$ pour tout $p \in \left(2, \frac{2}{1-2\varepsilon}\right)$.

Seconde étape : Condition de régularité

Il nous reste encore à vérifier la condition de régularité pour X_f afin d'appliquer le théorème 2.1.3. Soit $h \in (0,1]$ vérifiant les hypothèses du corollaire. Comme nous l'avons évoqué dans la sous-partie 1.1.1, pour tout $\kappa > 1$ on peut définir une distance symbolique d_κ sur A ; si κ est suffisamment proche de 1, alors $(A, d_\kappa, \mu_{\alpha|A}, T_A)$ est encore Gibbs-Markov. Quitte à prendre des valeurs de κ encore plus petites, on peut supposer que $d^h \leq C d_\kappa$. On note $|\cdot|_{\mathrm{Lip}_\kappa}$ la semi-norme lipschitzienne par rapport à d_κ.

Le calcul des bornes sur la semi-norme lipschitzienne de X_f se fait selon le même schéma que pour la norme $\left\|X_{|f|}\right\|_{\mathbb{L}^p}$. Tout d'abord, pour tout $n \geq 1$ et pour tous x et y dans A_{n-1} :

$$\begin{aligned}
|X_f(x) - X_f(y)| &\leq \sum_{k=0}^{n} \left|f(T^k x) - f(T^k y)\right| \\
&\leq \sum_{k=0}^{n} |f|_{\mathrm{Hol}_h(T_\alpha A_k)} d(T^{n-k+1}x, T^{n-k+1}y)^h \\
&\leq \sum_{k=0}^{n} |f|_{\mathrm{Hol}_h(T_\alpha A_k)} d(T_A x, T_A y)^h \\
&\leq C\kappa^{-1} \sum_{k=0}^{n} |f|_{\mathrm{Hol}_h(T_\alpha A_k)} d_\kappa(x,y).
\end{aligned}$$

Par conséquent, comme on peut supposer $\varepsilon' < 1/2$:

$$|X_f|_{\mathrm{Lip}_\kappa(A_{n-1})} \leq C\kappa^{-1} \sum_{k=1}^{n} |f|_{\mathrm{Hol}_h(T_\alpha A_k)} \leq C + C' \sum_{k=1}^{n} \left(k^{-\frac{1}{\alpha}}\right)^{\alpha-1+\varepsilon'} \leq Cn^{\frac{1-\varepsilon'}{\alpha}};$$

avec l'équation (2.7.6), on obtient :

$$\sum_{n=0}^{+\infty} \mu_{\alpha|A}(A_n)\, |X_f|_{\mathrm{Lip}_\kappa(A_n)} \leq C + C' \sum_{n=0}^{+\infty} n^{-1-\frac{\varepsilon'}{\alpha}} < +\infty.$$

La condition de régularité du théorème 2.1.3 est donc satisfaite. □

Comme nous l'avons déjà fait remarquer, on peut obtenir, grâce à des méthodes perturbatives, les propositions 1.1.7 et 1.1.8 sous l'hypothèse plus faible qu'il existe $\theta > 0$ tel que $\mathbb{E}(D(X_f)^\theta) < +\infty$. En utilisant cette condition de régularité dans le théorème 2.1.3, on peut montrer que la condition de régularité présente dans les corollaires 2.7.5 et 2.7.6 peut être remplacée par :

$$|f|_{\mathrm{Hol}_h((x,1])} = O(x^r) \text{ en } 0 \text{ pour un certain } r < 0.$$

La condition de régularité faible étudiée dans la partie 2.5, et en particulier la proposition 2.5.5, permet d'aller plus loin :

Lemme 2.7.7.
Dans l'énoncé des corollaires 2.7.5 et 2.7.6, la condition :

$$|f|_{\mathrm{Hol}_h((0,x))} = O(x^{\alpha-1+\varepsilon'}) \text{ en } 0 \text{ pour un certain } \varepsilon' > 0$$

peut être remplacée par la condition :

$$|f|_{\mathrm{Hol}_h((x,1])} = O(e^{x^{-\frac{1}{2}+\varepsilon'}}) \text{ en } 0 \text{ pour un certain } \varepsilon' > 0$$

Démonstration.
La condition de régularité (2.5.10) dépend de deux paramètres, p et θ, tels que $(p-1)\theta > 1$. Étant donné que l'on a toujours $p > 2$, il est suffisant de supposer $\theta > 1$. En reprenant la vérification de la régularité de X_f dans le corollaire précédent, on peut supposer sans perte de généralité que $\varepsilon' < 1/2$, et alors :

$$|X_f|_{\mathrm{Lip}_\kappa(A_n)} \leq C\kappa^{-1} \sum_{k=0}^{n} |f|_{\mathrm{Hol}_h(T_\alpha A_k)} \leq C \sum_{k=0}^{n} e^{C\left(k^{-\frac{1}{\alpha}}\right)^{-\frac{1}{2}+\varepsilon'}} \leq C(n+1)e^{Cn^{\frac{1}{2\alpha}-\frac{\varepsilon'}{\alpha}}} \leq e^{Cn^{\frac{1}{2\alpha}-\frac{\varepsilon'}{\alpha}}}.$$

Soit $\theta > 0$. En appliquant la fonction $\ln_+^{1+\theta}$ de chaque côté de l'inégalité puis l'équation (2.7.6), on obtient :

$$\sum_{n=0}^{+\infty} \mu_\alpha(A_n) \left(\ln_+ |X_f|_{\mathrm{Lip}_\kappa(A_n)}\right)^{1+\theta} \leq \sum_{n=0}^{+\infty} C(n+1)^{\frac{1+\theta}{2\alpha}-\frac{\varepsilon'(1+\theta)}{\alpha}-1-\frac{1}{\alpha}}.$$

La condition de régularité (2.5.10) est satisfaite pour tout paramètre θ dans $\left(1, \frac{1+2\varepsilon'}{1-2\varepsilon'}\right)$, et donc a fortiori dans $\left(\frac{1}{p-1}, \frac{1+2\varepsilon'}{1-2\varepsilon'}\right)$. □

D'après la remarque 2.3.4, le théorème limite peut être appliqué à de plus larges ensembles de fonctions. Par exemple, soit f la fonction définie sur $(0,1]$ par $f := \sum_{n=1}^{+\infty}(-1)^n 1_{T_\alpha A_n}$ sur $(0,1/2]$, telle que f soit constante sur $(1/2,1]$ et X_f soit d'intégrale nulle. Alors les conclusions des corollaires 2.7.5 et 2.7.6 sont encore valides. Il existe même des exemples de fonctions f définie sur $(0,1]$ telles que $\lim_0 |f| = +\infty$ et qui vérifient malgré tout les conclusions de ces corollaires.

Chapitre 3

Semi-flots et flots

Sommaire

Ce troisième chapitre est dédié aux systèmes en temps continu. Notre but ultime est d'obtenir une version du théorème 2.1.3 pour le flot géodésique sur des variétés riemanniennes $\mathbb{Z}$ ou $\mathbb{Z}^2$-périodiques de courbure sectionnelle strictement négative. Nous voulons donc généraliser nos résultats dans trois directions simultanément :

- Au lieu de travailler avec des systèmes en temps discret (des tours au-dessus de systèmes Gibbs-Markov), nous voulons des résultats pour des systèmes en temps continu (des semi-flots de suspension au-dessus de systèmes Gibbs-Markov) ;
- Nous voulons des résultats pour des $\mathbb{Z}$ ou $\mathbb{Z}^2$ extensions de systèmes Gibbs-Markov (parfois appelées “marches aléatoires engendrées par une application Gibbs-Markov”), pour lesquelles le système est donné non pas en terme d'un temps de premier retour, mais en termes d'un noyau de transition ;
- Nous voulons des résultats pour des systèmes inversibles, et en particulier pour des extensions naturelles d'applications Gibbs-Markov.

Le flot géodésique sur des variétés hyperboliques périodiques mélange ces trois aspects : il peut se voir par exemple comme un flot de suspension au-dessus d'une $\mathbb{Z}^d$ extension de l'extension naturelle d'une application Gibbs-Markov. De plus, les termes “flot de suspension”, “$\mathbb{Z}^d$ extension” et “extension naturelle” dans la phrase précédente, dans une certaine mesure, commutent. Si l'on prend garde à bien définir les objets en jeu, le flot géodésique sur des variétés hyperboliques périodiques est aussi, par exemple, une $\mathbb{Z}^d$ extension du flot de suspension au-dessus de l'extension naturelle d'une application Gibbs-Markov, etc.

Nous n'écrirons pas une version du théorème 2.1.3 pour chaque combinaison possible de “flot de suspension”, “$\mathbb{Z}^d$ extension” et “extension naturelle”. À la place, nous progresserons vers le cas le plus complexe (celui du flot géodésique) le long du chemin suivant. Dans la partie 3.1, nous étendons le théorème 2.1.3 au cas des semi-flots de suspension au-dessus d'un système Gibbs-Markov. Dans la partie 3.2, nous utiliserons des méthodes spectrales pour étudier les propriété de récurrence de

$\mathbb{Z}^d$ extension de semi-flots Gibbs-Markov. Dans la partie 3.3, ce travail est utilisé pour étendre le théorème 2.1.3 aux $\mathbb{Z}^d$ extension de semi-flots Gibbs-Markov. Enfin, la partie 3.4 se concentre sur le flot géodésique sur des variétés hyperboliques périodiques ; à cause du grand nombre d'hypothèses mises en jeu, nous ne formulerons pas de résultat pour des flots de suspension génériques au-dessus d'une $\mathbb{Z}^d$ extension de l'extension naturelle d'une application Gibbs-Markov.

Les parties 3.1, 3.3 et 3.4 forment la deuxième partie de l'article [70], amendée pour retirer certains lemmes redondants. La partie 3.2 est à part : elle forme l'intégralité de l'article [71], et les méthodes utilisées sont de nature très différente.

3.1 Théorèmes limites pour les semi-flots

L'objet de cette sous-partie est de mettre en place des résultats généraux pour des semi-flots Gibbs-Markov ergodiques (voir la sous-partie 1.2.3 pour la définition). En particulier, nous montrerons un équivalent du théorème 2.1.3 pour des semi-flots Gibbs-Markov. Commençons par adapter quelques définitions. Dans la suite, $(\Omega, \mu, (g_t)_{t \geq 0})$ est un semi-flot Gibbs-Markov de base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ et de hauteur φ_A.

Définition 3.1.1 (Temps de retour et temps local).
Pour tout entier $N \geq 0$, le temps de N-ième retour est la fonction $\tau_N := \sum_{k=0}^{N-1} \varphi_A \circ T_A^k$.
Pour tout réel positif t, le temps local en A au temps t est la fonction qui à presque tout $x \in A$ associe $\xi_t(x) := \mathrm{Card}\{s \in (0,t] : \ g_s(x) \in A\}$.

Avec ces définitions, on a $\xi_{\tau_N} = N$ pour tout $N \in \mathbb{N}$ et $t \leq \tau_{\xi_t}$ pour tout $t \in \mathbb{R}_+^*$. Ces fonctions peuvent être étendues à tout Ω.

La proposition 2.4.1 généralise déjà le théorème 2.1.1 de façon satisfaisante : elle nous autorise à traiter des semi-flots ergodiques, et non seulement des semi-flots dont la base est mélangeante. Notre travail consiste ici à adapter le théorème 2.1.3, dont nous donnerons la variante suivante :

Proposition 3.1.2.
Soit $(\Omega, \mu, (g_t)_{t \geq 0})$ un semi-flot Gibbs-Markov ergodique de base $(A, \pi, d, \mathcal{B}, \mu_A, T_A)$ et de hauteur φ_A. Soit f une fonction définie sur Ω et à valeurs réelles. Supposons que :
- *f est mesurable et il existe $p > 2$ tel que :*

$$X_f^* := \sup_{t \in [0, \varphi_A]} \left| \int_0^t f \circ g_s \, \mathrm{d}s \right| \in \mathbb{L}^p(\Omega, \mu_A) \ ;$$

- *$\int_A X_f \, \mathrm{d}\mu_A = 0$;*
- *La fonction $\psi(t) = \mu_A(\varphi_A \geq t)^{-1}$ est à variation régulière d'indice $\beta \in [0,1)$;*
- *$\mathbb{E}(D(X_f))$ et $\mathbb{E}(D(\varphi_A))$ sont finis.*

Alors, pour toute mesure de probabilité ν absolument continue par rapport à μ ou à μ_A,

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma(f)\sqrt{Y_\beta}\mathcal{N}, \tag{3.1.1}$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur (Ω, ν), où Y_β et $\mathcal{N}$ sont des variables aléatoires indépendantes, Y_β suit une loi de Mittag-Leffler standard de paramètre β, et $\mathcal{N}$ est une variable aléatoire gaussienne standard. De plus,

$$\sigma(f)^2 = \lim_{N \to +\infty} \frac{1}{N} \int_A \left(\sum_{i=0}^{N-1} X_f \circ T_A^i \right)^2 \mathrm{d}\mu_A, \tag{3.1.2}$$

et $\sigma(f) = 0$ si et seulement si f est un cobord.

Le cas $\beta = 1$ est, comme d'habitude, traité à part. Il suffit de changer la renormalisation ci-dessus,

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}},$$

que l'on remplace alors par :

$$\sqrt{\frac{1}{t}\int_0^t \frac{1}{\psi(s)}\,\mathrm{d}s}.$$

Une première étape consiste à analyser le comportement asymptotique du temps local pour des semi-flots Gibbs-Markov. Heureusement, la sous-partie 1.2.4 fournit la réponse dans le cadre des applications Gibbs-Markov. Il suffit d'utiliser la table 3.2.4 pour des extensions de dimension 0.

Une autre méthode aurait été de créer une version discrétisée du semi-flot, en prenant une autre fonction de hauteur suffisamment proche de φ_A mais à valeurs entières. On peut alors déduire le comportement du temps local en temps continu à partir de son comportement en temps discret. C'est cette approche qui a été adoptée dans [70], avec les lemmes 6.3 et 6.4 ; elle a l'avantage d'être plus générale. Cependant, dans le cadre des applications Gibbs-Markov, les conditions suffisantes que l'on obtient sur la hauteur du semi-flot ne sont pas meilleures qu'avec les méthodes spectrales. On évitera donc ce détour.

Nous allons maintenant démontrer la proposition 3.1.2. La plus grande partie de la preuve du théorème 2.1.3 reste inchangée. Le principal obstacle est qu'en temps discret, nous utilisions un résultat de R. Zweimüller [75] ; celui-ci n'est pas disponible en temps continu, et nous allons donc devoir en démontrer une version.

Preuve de la proposition 3.1.2.

Les deux premières étapes de la preuve du théorème 2.1.3, adaptées à des systèmes ergodiques mais non nécessairement mélangeant, peuvent être adaptées sans obstacle majeur. Quitte à supposer que $\mathbb{E}(D(\varphi_A))$ soit fini, les lemmes 2.3.1 et 2.3.2 peuvent être adaptés aisément au temps continu (dans lequel φ_A est à valeurs réelles positives, le temps local (ξ_t) est défini sur $\mathbb{R}_+$, etc.).

Ainsi, sous les mêmes hypothèses que la proposition,

$$\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}}\int_0^t f\circ g_s\,\mathrm{d}s \to \sigma(f)\sqrt{Y_\beta}\mathcal{N},$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur (Ω,μ_A), où Y_β et $\mathcal{N}$ sont des variables aléatoires indépendantes, Y_β suit une loi de Mittag-Leffler standard de paramètre β, $\mathcal{N}$ est une variable aléatoire gaussienne standard, et $\sigma(f)$ est donnée par la formule (3.1.2).

Nous allons à présent généraliser ce résultat à tout espace probabilisé de départ (Ω,ν), où ν est absolument continue par rapport à μ_A ou à μ. Cependant, nous ne pouvons pas utiliser [75, Corollaire 1], qui ne porte que sur les systèmes en temps discret. L'un des problèmes à résoudre, qui ne se pose pas en temps discret, est de borner les incréments du temps local sur une fenêtre de temps fixée. On utilisera pour cela le lemme 2.3.2.

Nous commençons par démontrer le résultat pour toute mesure initiale absolument continue par rapport à μ_A, puis nous en déduirons le cas général.

Mesures de probabilité absolument continues par rapport à μ_A

Soit f une fonction satisfaisant les hypothèses de la proposition 3.1.2.

Pour l'instant, supposons que la base (A,μ_A,T_A) est non seulement ergodique, mais aussi mélangeante. Elle est alors exacte : pour toute fonction $h\in\mathbb{L}^1(A,\mu_A)$, la suite $(\mathcal{L}^n h)_{n\geq 0}$ converge dans $\mathbb{L}^1$ vers la fonction constante égale à $\int_A h\,\mathrm{d}\mu_A$ [2, Theorem 1.3.3].

Soit $\nu = h\,\mathrm{d}\mu_A$ une mesure de probabilité absolument continue par rapport à μ_A, et soit $\varepsilon>0$. Soit n un entier tel que $\|\mathcal{L}^n h - 1\|_{\mathbb{L}^1}\leq\varepsilon$.

On sait que, pour toute fonction $G\ :\ \mathbb{R}\to\mathbb{R}$ uniformément continue et bornée,

$$\lim_{t\to+\infty}\mathbb{E}_{\mu_A}\left(G\left(\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}}\int_0^t f\circ g_s\,\mathrm{d}s\right)\right) = \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right).$$

Étant donné que $\mathcal{L}^n h\,\mathrm{d}\mu_A$ et μ_A sont à distance d'au plus ε en variation totale,

$$\limsup_{t\to+\infty}\left|\mathbb{E}_{\mathcal{L}^n h d\mu_A}\left(G\left(\frac{1}{\sqrt{\mathrm{sinc}(\beta\pi)\psi(t)}}\int_0^t f\circ g_s\,\mathrm{d}s\right)\right) - \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right)\right| \leq \varepsilon\,\|G\|_\infty.$$

De plus, si x suit la loi h dμ_A alors $T_A^n x$ suit la loi $\mathcal{L}^n h$ dμ_A. Ainsi, pour tout $t \geq 0$,

$$\mathbb{E}_{\mathcal{L}^n h d\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}}\int_0^t f\circ g_s \,\mathrm{d}s\right)\right) = \mathbb{E}_{hd\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}}\int_{\tau_n}^{t+\tau_n} f\circ g_s \,\mathrm{d}s\right)\right).$$

La quantité $\psi(t)^{-1/2}\int_0^{\tau_n} f \circ g_s$ ds tend vers 0 presque sûrement. D'après le lemme 2.3.2, pour tout $r > 0$, presque sûrement, pour tout réel t suffisamment grand, $\xi_{t+\tau_n} - \xi_t \leq \psi(t)^r$. De plus, d'après la deuxième étape de la preuve du théorème 2.1.3, pour tout $\delta > 0$, presque sûrement, pour tout réel t suffisamment grand, $X_f^* \circ T_A^{\xi_t} \leq \psi(t)^{\frac{1}{p}+\delta}$. Par conséquent, pour tous $r, \delta > 0$, presque sûrement, pour tout réel t suffisamment grand,

$$\left|\int_t^{t+\tau_n} f\circ g_s \,\mathrm{d}s\right| \leq \sum_{k=\xi_t}^{\xi_{t+\tau_n}} X_f^*\circ T_A^k \leq (\xi_{t+\tau_n}-\xi_t+1)\psi(t+\tau_n)^{\frac{1}{p}+\delta} \leq (1+\psi(t)^r)\psi(t+\tau_n)^{\frac{1}{p}+\delta}.$$

En prenant r et δ suffisamment petits, on en déduit que $\psi(t)^{-1/2}\int_t^{t+\tau_n} f\circ g_s$ ds tend vers 0 presque sûrement. Ainsi,

$$\limsup_{t\to+\infty}\left|\mathbb{E}_{hd\mu_A}\left(G\left(\frac{1}{\sqrt{\operatorname{sinc}(\beta\pi)\psi(t)}}\int_0^t f\circ g_s \,\mathrm{d}s\right)\right) - \mathbb{E}\left(G\left(\sigma(f)\sqrt{Y_\beta}\mathcal{N}\right)\right)\right| \leq \varepsilon \left\|G\right\|_\infty.$$

Ceci étant vrai pour tout $\varepsilon > 0$, la convergence en loi que l'on cherche à démontrer est vérifiée quand le point de départ est choisi selon la mesure ν, où ν est n'importe quelle mesure de probabilité absolument continue par rapport à μ_A, et tant que la base du semi-flot est mélangeante.

Supposons maintenant que la base du semi-flot est seulement ergodique. Soit M la période de la base. Alors, pour toute fonction h qui soit positive, mesurable et de masse 1, la suite des densités $(M^{-1}\sum_{k=0}^{M-1}\mathcal{L}^{n+k}h)_{n\geq 0}$ converge vers la fonction constante égale à 1 dans $\mathbb{L}^1(A,\mu_A)$, et la même stratégie fonctionne (malgré des calculs plus lourds).

Mesures de probabilité absolument continues par rapport à μ

Soit ν une mesure de probabilité absolument continue par rapport à μ, et soit ν_A sa projection sur A, donnée pour tout sous-ensemble mesurable $B \subset A$ par :

$$\nu_A(B) := \nu(\{(x,t)\in\Omega \ : \ x\in B, \ t\in[0,\varphi_A(x))\}).$$

Par construction, ν_A est absolument continue par rapport à μ_A ; en effet,

$$\frac{\mathrm{d}\nu_A}{\mathrm{d}\mu_A} = X_{\frac{\mathrm{d}\nu}{\mathrm{d}\mu}}.$$

Enfin, les mesures ν et ν_A sont naturellement couplées par la projection sur la base. Soit $(f_t)_{t\geq 0}$ le processus $(f\circ g_t)_{t\geq 0}$ où le point de départ est choisi selon la mesure ν, et $(\tilde{f}_t)_{t\geq 0}$ le même processus mais où le point de départ est choisi selon la mesure ν_A. On dispose alors d'un couplage entre ces deux processus, et suivant ce couplage ils ne diffèrent que par un décalage aléatoire du temps :

$$\left|\int_0^t f_s \,\mathrm{d}s - \int_0^t \tilde{f}_s \,\mathrm{d}s\right| \leq \int_0^{\varphi_A}\left|\tilde{f}_s\right| \,\mathrm{d}s + \int_t^{t+\varphi_A}\left|\tilde{f}_s\right| \,\mathrm{d}s.$$

En utilisant la même méthode que dans la première étape, on montre que le processus stochastique $\psi(t)^{-1/2}\left|\int_0^t f_s \,\mathrm{d}s - \int_0^t \tilde{f}_s \,\mathrm{d}s\right|$ converge presque sûrement vers 0. Il suffit donc d'étudier le comportement asymptotique de $(\tilde{f}_t)_{t\geq 0}$. Mais on le connaît déjà, car $\nu_A \ll \mu_A$.

Finalement, supposons que $\sigma(f) = 0$. Alors, de même qu'en temps discret, X_f est un cobord. Soit $\tilde{u}$ une fonction mesurable sur A et à valeurs réelles telle que $X_f = \tilde{u}\circ T_A - \tilde{u}$. Pour tout $x \in A$ et pour tout $t \in [0, \varphi_A(x))$, on pose :

$$u(x,t) := \tilde{u}(T_A x) - \int_t^{\varphi_A(x)} f\circ g_s(x,0) \,\mathrm{d}s.$$

Alors $\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_s - u$ pour tout $t \geq 0$, de telle sorte que f est un cobord.

A l'inverse, supposons que f est un cobord. Alors il existe une fonction mesurable u telle que, pour tout $t \geq 0$:

$$\int_0^t f \circ g_s \, \mathrm{d}s = u \circ g_t - u,$$

et alors $X_f = u \circ T_A - u$ est aussi un cobord. □

Pour démontrer le cas $\beta = 1$, il suffit de changer les renormalisations utilisées.

Remarque 3.1.3.

L'énoncé de la proposition 3.1.2 donne un résultat de convergence en loi pour toute mesure de probabilité ν absolument continue par rapport à μ ou μ_A. On voit en fait dans la preuve que ce résultat est vrai dès que la projection sur la base de la mesure ν est absolument continue par rapport à μ_A.

De plus, f n'est pas nécessairement une fonction définie sur Ω, mais peut être une mesure. On pose alors :

$$\int_0^t f \circ g_s(x) \, \mathrm{d}s := f(\{g_s(x) \ : \ s \in (0,t]\}).$$

L'ensemble des objets et notions utilisés, que ce soient les fonctions induites X_f et $X_{|f|}$, la fonction "supremum sur une excursion" X_f^* on encore la notion de cobord, se généralisent sans problème à ces objets ; en effet, ils sont tous définis sous forme intégrale.

Le temps local peut s'exprimer sous cette forme ; avec des $\mathbb{Z}^d$ extensions de semi-flots, si $\xi_t(q)$ est le temps local en $A \times \{q\} \times \{0\}$, alors $\xi_t(q) - \xi_t(q')$ est d'intégrale nulle pour tous q, $q' \in \mathbb{Z}^d$. On vérifiera dans la partie suivante qu'une telle fonction satisfait toujours les hypothèses de la proposition 3.1.2.

Plus généralement, on peut obtenir des théorèmes limites pour le nombre de fois où f passe à travers une hypersurface transverse au flot géodésique, ou pour le nombre de fois où une particule touche un obstacle donné pour un billard, sans avoir à forcer l'apparition de vraies fonctions sur Ω.

3.2 Temps local et de premier retour pour des $\mathbb{Z}^d$-extensions

Une importante classe de systèmes dynamiques munis d'une mesure infinie est constituée de $\mathbb{Z}$ et $\mathbb{Z}^2$ extensions de systèmes Gibbs-Markov, ou de semi-flots Gibbs-Markov. Comme annoncé, nous n'allons pas travailler sur les $\mathbb{Z}^d$ extensions de systèmes Gibbs-Markov, mais directement sur les $\mathbb{Z}^d$ extension de semi-flots Gibbs-Markov. Les résultats en temps discret peuvent se récupérer facilement en prenant des temps de transition constants ; nous en laissons le soin au lecteur intéressé.

Dans cette partie, nous nous intéressons aux propriétés de récurrence de $\mathbb{Z}^d$ extensions de flots de suspension au-dessus de systèmes dynamiques. Exceptionnellement, nous mènerons une partie de l'analyse dans un cadre plus général que celui des applications Gibbs-Markov : des hypothèses spectrales sur les systèmes considérés suffisent.

Ces objets ont un certain nombre de propriétés en commun avec les marches aléatoires. Les sauts ne sont pas indépendants, car le hasard est engendré par un système dynamique sous-jacent. De plus, au lieu de faire un saut à chaque unité de temps, le processus attend un certain temps avant de procéder à un saut. L'étude du flot géodésique sur des surfaces hyperboliques périodiques, ou sur des billards périodiques, peut se ramener à de tels systèmes.

Dans cette partie, nous nous intéressons plus particulièrement au temps de premier retour et au temps local en l'origine. Nous voulons répondre à deux questions : quelles sont les queues du temps de premier retour ? Quelle est le comportement distributionnel asymptotique du temps local ?

Afin de répondre à ces questions, une possibilité est de chercher les réponses dans la littérature sur les systèmes en temps discret, qui est beaucoup plus étendue, et d'en déduire des résultats pour des semi-flots. Ainsi, par [4] nous disposons d'un théorème central limite local pour des $\mathbb{Z}^d$ extension d'applications Gibbs-Markov, et la série d'articles [64] [65] [66] [26] donne un théorème central limite local pour des gaz de Lorentz planaires à horizon fini ou infini en temps discret. Certains de ces résultats se transposent en temps continu.

Il y a deux problèmes avec cette approche. Premièrement, elle est inélégante : la majeure partie de l'algèbre employée avec des systèmes discrets peut aussi être utilisée pour des systèmes continus. Le passage par des systèmes discrets peut donc sembler être un détour superflu. Deuxièmement, pour passer du temps discret au temps continu, il faut contrôler les déviations du processus $(\sum_{k=0}^{n-1} r \circ T_A^k)_{n \geq 0}$ par rapport à son espérance $(n\mathbb{E}(r))_{n \geq 0}$, ce qui engendre des restrictions assez artificielles sur l'intégrabilité de r. En particulier, certains arguments nécessitent un temps de saut r de carré intégrable [70, lemme 6.10], ce qui n'est pas le cas pour des gaz de Lorentz à horizon infini. De tels arguments ont de plus peu de chances de fonctionner si r n'est pas intégrable. Ces restrictions peuvent être évitées si l'on travaille directement avec des systèmes en temps continu.

Nous utilisons dans cette partie une combinaison de méthodes spectrales et de théorèmes taubériens, assez proche des méthodes utilisées pour compter les géodésiques fermées sur des variétés périodiques de courbure sectionnelle strictement négative [47] [63] [57] [8] [9]. De plus, les opérateurs de transfert tordus ont déjà été employés, de façon différente, pour démontrer des propriétés statistiques de semi-flots [25].

Nous commençons par présenter les méthodes perturbatives en théorie spectrale, ainsi que quelques définitions supplémentaires, dans la sous-partie 3.2.1. Les lemmes-clés sont démontrés dans la sous-partie 3.2.2 sous des conditions abstraites portant sur les opérateurs de transfert et les espaces de Banach. Ces résultats restent valides pour de vastes classes de systèmes dynamiques hyperboliques, comme par exemple les applications uniformément dilatantes de l'intervalle [49]. Les deux sous-parties suivantes sont le lieu de calculs assez lourds, pour lesquels nous travaillerons uniquement avec des applications Gibbs-Markov. Dans la sous-partie 3.2.3, nous obtenons un développement asymptotique pour la valeur propre principale des opérateurs de transfert perturbés. Dans la sous-partie 3.2.4, nous utilisons ce développement conjointement avec les résultats de la sous-partie 3.2.2 pour répondre aux deux questions que nous avons posées sur le temps de premier retour et le temps local en l'origine.

Dans l'ensemble de la partie 3.2, (A, μ_A, T_A) est un système dynamique qui préserve une mesure de probabilité. Nous désignerons par $\mathcal{B}$ un espace de Banach complexe de fonctions définies μ_A-presque partout sur A. Nous supposerons de plus que les fonctions constantes appartiennent à $\mathcal{B}$, que $\mathcal{B} \subset \mathbb{L}^1$, et que la transformation $T_A : A \to A$ a un nombre dénombrable de branches. Enfin, nous utiliserons la notation $S_n f := \sum_{k=0}^{n-1} f \circ T_A^k$, et la notation $S_n^{T_A} f$ s'il y a ambiguïté sur la transformation appliquée.

3.2.1 Quelques définitions supplémentaires

$\mathbb{Z}^d$ extensions et système induit

La principale différence avec la partie précédente est qu'au lieu de disposer directement d'une base et d'une hauteur explicite, le système est donné en termes d'une transformation de transition, d'une fonction de saut et d'un temps de saut. La définition d'une $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov est très proche à la fois de la définition d'un semi-flot, et de la définition d'une $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov de la sous-partie 1.5.3.

Définition 3.2.1 ($\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov).

Soit $(A, \pi, d_A, \mathcal{B}, \mu_A, T)$ une application Gibbs-Markov ergodique. Soit r une fonction mesurable et μ_A-presque sûrement strictement positive sur A. Soit d un entier positif, et soit F une fonction définie sur A, à valeurs dans $\mathbb{Z}^d$ et $\sigma(\pi)$-mesurable (c'est-à-dire presque sûrement constante sur chaque élément de la partition π).

La $\mathbb{Z}^d$ extension de base $(A, \pi, d_A, \mathcal{B}, \mu_A, T_A)$, de fonction de saut F et de temps de saut r est le système dynamique $(\Omega, \mu, (g_t)_{t \geq 0})$ défini de la façon suivante :

- *Ω est l'espace quotient de l'espace topologique $A \times \mathbb{Z}^d \times \mathbb{R}_+$ par la relation d'équivalence $\sim$ définie par $(x, q, r(x) + t) \sim (T_A(x), q + F(x), t)$ pour tous (x, q, t) dans $A \times \mathbb{Z}^d \times \mathbb{R}_+$;*
- *Le semi-flot $(g_t)_{t \geq 0}$ agit sur $A \times \mathbb{Z}^d \times \mathbb{R}_+$ par translation sur la troisième coordonnée, et sur Ω par la projection canonique de cette action (qui est bien définie) ;*

- $\mu = \mu_A \otimes \text{Leb} \otimes \text{Leb}$ *sur* $A \times \mathbb{Z}^d \times \mathbb{R}_+$, *et est définie sur* Ω *par restriction au domaine fondamental* $\{(x,q,t) \in A \times \mathbb{Z}^d \times \mathbb{R}_+ : \ t < \varphi_A(x)\}$.

Nous utiliserons souvent l'identification $\Omega \simeq A \times \mathbb{Z}^d \times [0,r)$.

Une telle extension est dite dégénérée si F *peut être écrite comme la somme d'une fonction à valeurs dans un translaté d'un sous-réseau propre de* $\mathbb{Z}^d$ *et d'un cobord.*

Cette notion de dégénérescence recouvre à la fois le cas où le processus de la $\mathbb{Z}^d$ extension vit dans un hyperplan, auquel cas la dimension du processus est strictement inférieure à d (par exemple, une marche aléatoire de dimension 1 dans $\mathbb{Z}^2$), et le cas où la dimension d est adaptée, mais où le processus a une période (par exemple, la marche aléatoire simple sur $\mathbb{Z}$ ou $\mathbb{Z}^2$). Quitte à ajouter un cobord à F et à itérer la transformation sous-jacente, on peut supposer qu'une $\mathbb{Z}^d$ extension dégénérée prend ses valeurs dans un sous-réseau de $\mathbb{Z}^d$, qui est isomorphe à $\mathbb{Z}^{d'}$ pour un certain $d' \leq d$. On ne perd donc rien en généralité en se restreignant aux processus non dégénérés ; cependant, cette hypothèse élimine des causes évidentes de non-ergodicité, et simplifie notablement les énoncés des théorèmes.

Considérons une $\mathbb{Z}^d$ extension ergodique de base (A, μ_A, T_A), de fonction de saut F et de temps de saut r. On identifie A avec $A_0 := A \times \{0\} \times \{0\}$, et on définit :

- Le temps de premier retour φ en A_0 par $\varphi(x) := \inf\{t > 0 : \ g_t(x,0,0) \in A_0\}$ (cette fonction peut être étendue à Ω) ;
- Le temps local en l'origine par $\xi_t(x) := \text{Card}\{s \in (0,t] : \ g_s(x,0,0) \in A_0\}$;
- Le temps de premier retour discrétisé $\overline{\varphi}$ en A_0 comme étant le temps de premier retour en A_0 pour la $\mathbb{Z}^d$ extension avec la même base, la même fonction de saut, mais un temps de saut constant et égal à 1 ;
- La position au temps t comme étant l'unique élément $S_t \in \mathbb{Z}^d$ tel qu'il existe $y \in A$ et $s < r(y)$ tels que $g_t(x,0,0) = (y, S_t, s)$.

Le semi-flot étant supposé ergodique, les fonctions φ et $\overline{\varphi}$ sont presque sûrement finies. La fonction $\overline{\varphi}$ est le temps de premier retour en $A \times \{0\}$ pour le système dynamique ergodique et préservant la mesure $T(x,q) := (T_A(x), q + F(x))$ sur $A \times \mathbb{Z}^d$.

Grâce au travail de la partie 3.1, on dispose déjà des fonctions "temps de N-ième retour en A_0" τ_N et "temps local en A_0" (ξ_t). En prenant $r \equiv 1$, on définit aussi des version discrétisées de ces quantités, que l'on distinguera par une barre : $(\overline{S}_t)$, $\overline{\tau}_N$, $(\overline{\xi}_t)$, $\overline{\psi}(t)$...

Perturbation de l'opérateur de transfert

Dans la sous-partie 1.1.2, nous avons présenté quelques fondamentaux de théorie spectrale appliquée à l'opérateur de transfert d'un système dynamique Gibbs-Markov. Dans ce qui suit, nous nous intéressons aux perturbations de l'opérateur de transfert pour un système dynamique mesuré général, avant de regarder de plus près le cas des applications Gibbs-Markov.

Rappelons que l'opérateur de transfert $\mathcal{L}$ associé au système dynamique (A, μ_A, T_A) peut être défini par :

$$\mathcal{L}f(x) = \sum_{\{y \ : \ T_A y = x\}} g(y) f(y) \quad \forall \ f \in \mathbb{L}^1(A, \mu_A),$$

où g est l'inverse du jacobien de la transformation T_A par rapport à la mesure μ_A.

Donnons nous une fonction potentiel $\phi \ : \ A \to \mathbb{C}$ mesurable et telle que $\Re(\phi)$ soit bornée supérieurement. Alors on peut définir un nouvel opérateur $\mathcal{L}_\phi$, qui agit continûment sur $\mathbb{L}^1(A, \mu_A)$ par :

$$\mathcal{L}_\phi f(x) := \sum_{y \in T_A^{-1}(\{x\})} e^{\phi(y)} g(y) f(y) = \mathcal{L}(e^\phi f)(x) \quad \forall \ f \in \mathbb{L}^1(A, \mu_A).$$

Soit $\mathbb{T}^d \simeq (\mathbb{R}/2\pi\mathbb{Z})^d$ le tore de dimension d. Dans cette partie, on utilisera plus particulièrement la famille à deux paramètres d'opérateurs perturbés $\left(\mathcal{L}_{-sr+i\langle w,F\rangle}\right)_{s\geq 0, w \in \mathbb{T}^d}$. Cette notation étant un peu lourde, nous l'allégeons ainsi :

$$\mathcal{L}_{s,w} := \mathcal{L}_{-sr+i\langle w,F\rangle}.$$

Suivant les propriétés des fonctions r et F, ainsi que de l'espace de Banach $\mathcal{B}$ choisi, il se peut que les opérateurs $\mathcal{L}_{s,w}$ agissent toujours continûment sur $\mathcal{B}$. Si elle existe et est unique, on notera $\rho_{s,w}$ la valeur propre de module maximal correspondant à cette action. Si de plus le sous-espace propre associé à la valeur propre $\rho_{s,w}$ n'est pas inclus dans l'hyperplan des fonctions d'intégrale nulle, on notera $f_{s,w}$ l'unique fonction propre associée qui soit d'intégrale 1.

Si le rayon spectral essentiel de $\mathcal{L}$ agissant sur $\mathcal{B}$ est strictement inférieur à 1, et si le système (A, μ_A, T_A) est faiblement mélangeant, alors 1 est une valeur propre isolée de multiplicité 1, et le reste du spectre de $\mathcal{L}$ est inclus dans un disque de rayon strictement inférieur à 1. Dans ce cas, tout opérateur $\mathcal{L}'$ suffisamment proche de $\mathcal{L}$ en norme d'opérateur a une unique valeur propre de module maximale, qui est de multiplicité 1 et est proche de 1 [46, Part IV.3.5],

Pour la plupart des applications, ils sera utile de supposer que la famille d'opérateurs $(\mathcal{L}_{s,w})$ est continue pour de petites valeurs de s et de w. Entre autres, cette hypothèse implique que la valeur propre principale $\rho_{s,w}$ et la fonction propre associée $f_{s,w}$ sont bien définies et sont continues en tant que fonctions de s et de w près de l'origine [46, Part IV.3.5].

Hypothèse 3.2.2 (Continuité de la perturbation).
Il existe un voisinage U de 0 dans $\mathbb{R}_+$ et un voisinage V de 0 dans $\mathbb{T}^d$ tels que la famille d'opérateurs $(\mathcal{L}_{s,w})_{(s,w)\in U\times V}$ dépend continûment de (s,w) pour la norme d'opérateur.
De plus, $\mathcal{L}$ a une valeur propre simple isolée en 1.

Système induit

Considérons une $\mathbb{Z}^d$ extension d'un flot de suspension de base (A, μ_A, T_A), de temps de saut r et de fonction de saut F. Supposons que ce système est ergodique et conservatif. Pour μ_A-presque tout $x \in A$, définissons $\widetilde{T}x \in A$ par :

$$g_{\varphi(x)}(x, 0, 0) = (\widetilde{T}x, 0, 0),$$

ou, de façon équivalente, pour presque tout x dans A,

$$\widetilde{T}x = T_A^{\overline{\varphi}(x)}x.$$

Alors $\widetilde{T}$ est une transformation de A qui préserve μ_A et qui est ergodique. De plus, T_A^n a un nombre dénombrable de branches pour tout $n \geq 0$, donc $\widetilde{T}$ a aussi un nombre dénombrable de branches. L'opérateur de transfert $\widetilde{\mathcal{L}}$ associé à $\widetilde{T}$ est bien défini ; pour tout $f \in \mathbb{L}^1$,

$$\widetilde{\mathcal{L}}f(x) = \sum_{y \in \widetilde{T}^{-1}(\{x\})} \widetilde{g}(y) f(y),$$

où :

$$\widetilde{g}(x) = \prod_{k=0}^{\overline{\varphi}(x)-1} g(T_A^k x) = e^{S_{\overline{\varphi}(x)}^{T_A} \ln g(x)}.$$

Nous pouvons définir une famille d'opérateurs perturbés $(\widetilde{\mathcal{L}}_s)$ qui agissent sur $\mathbb{L}^1(A, \mu_A)$ par $\widetilde{\mathcal{L}}_s f = \widetilde{\mathcal{L}}(e^{-s\varphi} f)$. Cependant, même si l'action de $\mathcal{L}$ sur un espace de Banach $\mathcal{B}$ vérifie de bonnes propriétés, ce n'est en général par le cas de l'action de $\widetilde{\mathcal{L}}$ sur le même espace de Banach $\mathcal{B}$.

Armés de la notion de système induit, nous pouvons énoncer un nouvel ensemble d'hypothèses. Bien qu'il puisse pour l'instant sembler obscur, il sera très utile par la suite ; nous renvoyons le lecteur curieux au lemme 3.2.5 pour voir comment un tel ensemble d'hypothèses peut apparaître.

Hypothèse 3.2.3.
Il existe un réel $\beta \in [0, 1]$ et une fonction G, définie sur un voisinage de 0 dans $\mathbb{R}_+$ et à valeurs réelles, tels que :

- $\lim_{0^+} G = +\infty$;
- *G est à variation régulière d'indice $-\beta$ en 0 ;*

- *Pour toute fonction $f \in \mathcal{B}$ positive,*

$$\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f = G(s) \left(\int_A f \, \mathrm{d}\mu_A + o(1) \right) \quad \textit{dans } \mathcal{B} \textit{ quand } s \textit{ tend vers } 0,$$

où le terme $o(1)$ est négligeable dans $\mathcal{B}$ et uniforme en f sur tout sous-ensemble borné de $\mathcal{B}$.

Remarquons que, sous cette hypothèse, pour toute fonction $f \in \mathcal{B}$ positive, pour tout $s > 0$ suffisamment petit,

$$\left(\mathrm{Id} - \widetilde{\mathcal{L}}_s \right) \sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f = f. \tag{3.2.1}$$

En effet, $\mathcal{B} \subset \mathbb{L}^1$, donc $\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f$ est une somme intégrable de termes positifs. La suite des sommes partielles $\left(\sum_{k=0}^{n-1} \widetilde{\mathcal{L}}_s^k f \right)_{n \geq 0}$ converge donc dans $\mathbb{L}^1$. Comme l'opérateur $\widetilde{\mathcal{L}}_s$ agit continûment sur $\mathbb{L}^1$, pour toute fonction $f \in \mathcal{B}$ positive,

$$\begin{aligned}
\left(\mathrm{Id} - \widetilde{\mathcal{L}}_s \right) \sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f &= \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s \right) \lim_{N \to +\infty} \sum_{n=0}^{N-1} \widetilde{\mathcal{L}}_s^n f \\
&= \lim_{N \to +\infty} \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s \right) \sum_{n=0}^{N-1} \widetilde{\mathcal{L}}_s^n f \\
&= \lim_{N \to +\infty} \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s^N \right) f \\
&= f,
\end{aligned}$$

où la limite est prise dans $\mathbb{L}^1$.

L'un des buts de cette partie est de trouver une fonction G ayant les propriétés décrites dans l'hypothèse 3.2.3. L'outil principal pour ce faire est le lemme 3.2.5, mais le calcul de G aura lieu dans la partie 1.5.

Perturbation pour des applications Gibbs-Markov

Une première remarque importante est que, si l'application à la base est Gibbs-Markov, alors l'application induite peut aussi être munie d'une structure Gibbs-Markov. Supposons que la base est une application Gibbs-Markov $(A, \pi, \lambda, \mu_A, T_A)$. Revenons à la sous-partie 1.2.5. Pour tout $n \geq 0$, l'évènement $\{\overline{\varphi} \leq n\}$ ne dépend que de la valeur de $(F, F \circ T, \cdots, F \circ T^{n-1})$; la fonction F étant $\mathcal{F}_1$-mesurable, cet évènement est $\mathcal{F}_n$-mesurable. Par conséquent, $\overline{\varphi}$ est un temps d'arrêt pour la filtration $(\mathcal{F}_n)_{n \geq 0}$ qui est presque sûrement strictement positif et fini. On utilise alors le lemme 1.2.6 pour définir une nouvelle application Gibbs-Markov $(A_0, \pi_{\overline{\varphi}}, \lambda, \mathcal{B}, \mu_A, T_{\overline{\varphi}})$. La $\mathbb{Z}^d$ extension initiale est alors un facteur du semi-flot Gibbs-Markov de base $(A_0, \pi_{\overline{\varphi}}, \lambda, \mathcal{B}, \mu_A, T_{\overline{\varphi}})$ et de hauteur φ. Nous noterons $D := D_{\pi, \lambda}$ et $D_{\overline{\varphi}} := D_{\pi_{\overline{\varphi}}, \lambda}$. Si l'opérateur $\mathcal{L}$ agit bien sur Lip_d^∞, ce n'est pas le cas de $\widetilde{\mathcal{L}}$; l'espace de Banach naturel sur lequel on peut faire agir $\widetilde{\mathcal{L}}$ est $\mathrm{Lip}_{d_{\overline{\varphi}}}^\infty$.

Pour démontrer le lemme 3.2.15, qui est essentiel pour le calcul des asymptotiques du temps local, nous aurons besoin d'un contrôle très précis du comportement des opérateurs de transfert perturbés. C'est à ce niveau, principalement, qu'interviendra l'hypothèse selon laquelle la base de la $\mathbb{Z}^d$ extension est Gibbs-Markov. Le lemme suivant, qui permettra d'avoir un tel contrôle, est construit sur les modèles de [31, lemme 3.5] et [31, corollaire 3.6].

Lemme 3.2.4.

Soit $(A, \pi, \lambda, \mu_A, T_A)$ une application Gibbs-Markov mélangeante. Soit d un entier positif. Soient r et F des fonctions mesurables sur A, à valeurs dans $\mathbb{R}_+$ et dans $\mathbb{R}^d$ respectivement. Supposons que $\mathbb{E}(D(r))$ est fini, que F est $\sigma(\pi)$-mesurable, et que $F \in \mathbb{L}^p(A, \mu_A)$ pour un certain $p > 1$. Alors il existe $\kappa \in (1, \lambda]$ et une constante $C > 0$ tels que :

- *La famille d'opérateur $(\mathcal{L}_{s,w})_{(s,w) \in \mathbb{R}_+ \times \mathbb{R}^d}$ est continue en tant que fonction de (s, w) pour la norme d'opérateur sur $\mathrm{Lip}_\kappa^\infty$;*

- $$\|\mathcal{L} - \mathcal{L}_{s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} \leq C \int_A 1 - e^{-sr} \, \mathrm{d}\mu_A;$$
- *Pour tout s suffisamment petit et pour tout w,*
$$\|\mathcal{L}_{s,w} - \mathcal{L}_{s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} \leq C \|w\| .$$

Démonstration.

Le premier point est vérifié si l'on suppose que $\mathbb{E}(D(r)^2)$ est fini, de par la remarque suivant le lemme 4.1.2 dans [32]. Cependant, d'après la même remarque, la condition $\mathbb{E}(D(r)) < +\infty$ est suffisante si l'on prend $\kappa \leq \sqrt{\lambda}$.

Le second point se montre en modifiant légèrement la preuve des lemmes 4.1.1 et 4.1.2 dans [32]. On obtient alors :

$$\|\mathcal{L} - \mathcal{L}_{s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} \leq C \max \left\{ s, \int_A 1 - e^{-sr} \, \mathrm{d}\mu_A \right\},$$

mais si r n'est pas uniformément nulle alors $s = O(\int_A 1 - e^{-sr} \, \mathrm{d}\mu_A)$.

La démonstration du troisième point est plus compliquée. Soit $\eta \in (0,1]$. Dans ce qui suit, nous posons $\kappa := \lambda^\eta$. Pour tout $(s,w) \in \mathbb{R}_+ \times \mathbb{T}^d$ et $a \in \pi$, définissons une fonction $\psi_{s,w}$ sur A par $\psi_{s,w}(x) = e^{-sr(y)+i\langle w,F\rangle(y)}$ dès qu'il existe un $y \in a$ tel que $T_A y = x$, et $\psi_{s,w}(x) = 0$ sinon. Définissons ensuite un opérateur $M_{a,s,w}$ agissant sur $\mathrm{Lip}_\kappa^\infty$ par :

$$M_{a,s,w} f(x) := \sum_{y:\ T_A y = x} 1_a(y) e^{-sr(y)+i\langle w,F\rangle(y)} g(y) u(y).$$

Fixons un $s \geq 0$. D'après la preuve du lemme 4.1.2 dans [32], il existe une constante C qui ne dépend pas de a et telle que, pour tout $w \in \mathbb{T}^d$:

$$\|M_{a,s,w} - M_{a,s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} \leq C \mu_A(a) \|\psi_{s,w} - \psi_{s,0}\|_{\mathrm{Lip}_\kappa^\infty} .$$

Soient x et x' dans le même cylindre $a' \in \pi$. L'image d'un cylindre par une application Gibbs-Markov est une union de cylindres, donc ou bien x et x' ont tous les deux une pré-image dans a, ou bien aucun des deux n'a de pré-image dans a. Supposons que les deux ont une pré-image dans a (l'autre possibilité est triviale), et notons y et y' respectivement ces pré-images. La fonction F est supposée être $\sigma(\pi)$-mesurable, donc $F(y) = F(y')$, et :

$$\begin{aligned}
|\psi_{s,w}(x) & -\psi_{s,0}(x) - \psi_{s,w}(x') + \psi_{s,0}(x')| \\
& = \left| e^{-sr(y)+i\langle w,F\rangle(y)} - e^{-sr(y)+i\langle 0,F\rangle(y)} - e^{-sr(y')+i\langle w,F\rangle(y')} + e^{-sr(y')+i\langle 0,F\rangle(y')} \right| \\
& = \left| 1 - e^{i\langle w,F\rangle(y)} \right| \left| e^{-sr(y)} - e^{-sr(y')} \right| \\
& \leq C s^\eta |r(y) - r(y')|^\eta \frac{1}{\mu_A(a)} \int_a \left| 1 - e^{i\langle w,F\rangle} \right| \, \mathrm{d}\mu_A.
\end{aligned}$$

On en tire une majoration de la semi-norme lipschitzienne de la fonction $\psi_{s,w} - \psi_{s,0}$:

$$\sup_{a' \in \pi} |\psi_{s,w} - \psi_{s,0}|_{\mathrm{Lip}_\kappa(a')} \leq C s^\eta |r|_{\mathrm{Lip}_\kappa(a)}^\eta \frac{1}{\mu_A(a)} \int_a |\langle w,F\rangle| \, \mathrm{d}\mu_A.$$

Il reste à estimer $\|\psi_{s,w} - \psi_{s,0}\|_{\mathbb{L}^\infty}$. Pour tout $a' \in \pi$ et tout $y \in a$,

$$\|\psi_{s,w} - \psi_{s,0}\|_{\mathbb{L}^\infty(a')} \leq \left| 1 - e^{i\langle w,F\rangle(y)} \right| \leq \frac{1}{\mu_A(a)} \int_a |\langle w,F\rangle| \, \mathrm{d}\mu_A.$$

Finalement, on obtient :

$$\|M_{a,s,w} - M_{a,s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} \leq C \left(1 + s^\eta |r|_{\mathrm{Lip}_\kappa(a)}^\eta \right) \int_a |\langle w,F\rangle| \, \mathrm{d}\mu_A.$$

On somme ces inégalités pour tout $a \in \pi$:

$$\begin{aligned}\|\mathcal{L}_{s,w} - \mathcal{L}_{s,0}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} &\leq C\|w\| \sum_{a\in\pi} \left(1 + s^\eta |r|_{\mathrm{Lip}_\kappa(a)}^\eta\right) \int_a \|F\| \, \mathrm{d}\mu_A \\ &= C\|w\| \left(\int_A \|F\| \, \mathrm{d}\mu_A + s^\eta \sum_{a\in\pi} |r|_{\mathrm{Lip}_\kappa(a)}^\eta \int_a \|F\| \, \mathrm{d}\mu_A\right).\end{aligned}$$

Par l'inégalité de Hölder,

$$\begin{aligned}&\sum_{a\in\pi} \mu_A(a) |r|_{\mathrm{Lip}_\kappa(a)}^\eta \frac{1}{\mu_A(a)} \int_a \|F\| \, \mathrm{d}\mu_A \\ &\leq \left(\sum_{a\in\pi} \mu_A(a) |r|_{\mathrm{Lip}_\kappa(a)}^{\frac{p\eta}{p-1}}\right)^{1-\frac{1}{p}} \left(\sum_{a\in\pi} \mu_A(a) \left(\frac{1}{\mu_A(a)} \int_a \|F\| \, \mathrm{d}\mu_A\right)^p\right)^{\frac{1}{p}} \\ &= \left(\sum_{a\in\pi} \mu_A(a) |r|_{\mathrm{Lip}_\kappa(a)}^{\frac{p\eta}{p-1}}\right)^{1-\frac{1}{p}} \|\|F\|\|_{\mathbb{L}^p}.\end{aligned}$$

Choisissons $\eta = 1 - 1/p$. Alors la borne ci-dessus est finie, donc $\|\mathcal{L}_{s,0} - \mathcal{L}_{s,w}\|_{\mathrm{Lip}_\kappa^\infty \to \mathrm{Lip}_\kappa^\infty} = O(\|w\|)$ uniformément sur tout compact de $\mathbb{R}_+$. □

Prenons un peu d'avance sur l'écoulement de la thèse. Par le lemme 3.3.1 (qui peut être démontré indépendamment), si $\mathbb{E}(D(r))$ est fini, alors $\mathbb{E}(D_{\overline{\varphi}}(\varphi))$ l'est aussi. Le lemme ci-dessus implique alors que la famille à un paramètre d'opérateurs $(\widetilde{\mathcal{L}}_s)_{s\geq 0}$ est continue. Si la $\mathbb{Z}^d$ extension est ergodique, on peut définir la valeur propre principale $\widetilde{\rho}_s$ de $\widetilde{\mathcal{L}}_s$ de sorte qu'elle soit continue pour de petites valeurs de s, et égale à 1 si $s = 0$. Le lemme 3.2.14 pourra même être utilisé pour estimer le comportement de $\widetilde{\rho}_s$ au voisinage de 0, et montrer que $(\mathrm{Id} - \widetilde{\mathcal{L}}_s)$ est inversible pour de petite valeurs de s.

Nous n'aurons pas besoin d'utiliser ces propriété additionnelles des applications Gibbs-Markov, mais elles rendent une partie de l'algèbre plus intuitive. Par exemple, elles permettent d'écrire $(\mathrm{Id} - \widetilde{\mathcal{L}}_s)^{-1} = \sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n$ pour de petite valeurs de s.

Hypothèses d'intégrabilité

Nous allons démontrer des théorèmes limite sous diverses conditions sur le temps de saut r et la fonction de saut F, qui s'expriment aisément en termes de fonctions à variation régulière. Nous en profitons pour définir des fonctions H, I et P sur un voisinage de 0 dans $\mathbb{R}_+$, $\mathbb{T}^d$ et $\mathbb{R}_+ \times \mathbb{T}^d$ respectivement. Les fonctions H et I sont associées à r et à F respectivement. Ces fonctions apparaissent naturellement quand l'on étudie le comportement de la valeur propre principale $\rho_{s,w}$ des opérateurs perturbés pour de petites valeurs de s et de w. Nous nous servirons aussi des fonctions H et I pour donner le résultat de nos calculs à la fin de cette partie.

Nous supposerons que le temps de saut r vérifie l'une des hypothèses suivantes :

- r est intégrable. Nous posons alors $H(s) := s \int_A r \, \mathrm{d}\mu_A$;
- r n'est pas intégrable, mais $\mathbb{P}_{\mu_A}(r > t)$ est à variation régulière d'indice -1. Nous posons alors $H(s) := s \int_0^{1/s} \mathbb{P}_{\mu_A}(r > t) \, \mathrm{d}t$;
- La fonction $\mathbb{P}_{\mu_A}(r > t)$ est à variation régulière d'indice $-\beta \in (-1, 0]$. Nous posons alors $H(s) := \Gamma(1-\beta)\mathbb{P}_{\mu_A}(r > s^{-1})$.

Dans ces trois cas, la fonction H peut être étendue par continuité en 0 par $H(0) := 0$. Cette fonction est définie de telle sorte que $\mathbb{E}(1 - e^{-sr}) \sim H(s)$ en 0.

Nous supposerons que la fonction de saut F vérifie l'une des hypothèses suivantes :

- F est dans $\mathbb{L}^2(A, \mu_A)$ et satisfait les hypothèses d'un théorème central limite, c'est-à-dire que $N^{-1/2} S_N F$ converge en loi vers une variable aléatoire gaussienne d'opérateur de covariance S non dégénéré. Nous posons alors $I(w) := \left\|\sqrt{S}w\right\|^2 / 2$;
- F n'est pas dans $\mathbb{L}^2(A, \mu_A)$, et il existe une fonction J à variation régulière d'indice $\gamma \in (1, 2]$ en 0 et un automorphisme M de $\mathbb{R}^d$ tels que $\mathbb{E}(1 - e^{i\langle w, F\rangle}) \sim J(\|Mw\|)$ en 0. Nous posons alors $I(w) := J(\|Mw\|)$.

Ces deux hypothèses impliquent que F est non seulement intégrable, mais aussi qu'il existe un $q > 1$ tel que $F \in \mathbb{L}^q(A, \mu_A)$. De plus, si F vérifie l'une de ces hypothèses, alors F est dans le bassin d'attraction d'une loi de Lévy stable symétrique de paramètre dans $(1, 2]$. Nous n'étudions pas le cas des fonctions de saut dans le bassin d'attraction d'une loi de Cauchy. D'après les résultats de la sous-partie 1.3.4, en dimension 1, le second cas est vérifié si F a des queues à variation régulière. Plus précisément, si la fonction F n'est pas dans $\mathbb{L}^2$, et $\mathbb{P}_{\mu_A}(|F| > x)$ est à variation régulière d'indice -2, alors nous choisissons $I(w) = w^2/2 \cdot \int_A F^2 1_{\{-|w|^{-1}, |w|^{-1}\}}(F) \, \mathrm{d}\mu_A$ [41, théorème 2.6.2]. Si $\mathbb{P}_{\mu_A}(F > x)$ et $\mathbb{P}_{\mu_A}(F < -x)$ sont équivalents et à variation régulière d'indice $\gamma \in (-2, -1)$, alors nous choisissons $I(w) = \mathbb{P}_{\mu_A}(F > |w|^{-1})$ [41, théorème 2.6.5].

En dimension 1, nous supposerons toujours que $M = \mathrm{Id}$.

Soit $K > 0$ tel que $1/J$ est bien défini et localement intégrable sur $(0, K)$. Nous posons alors $\widetilde{J}(x) = \int_x^K t/J(t) \, \mathrm{d}t$ pour tout $x \in (0, K)$. Par exemple, si $J(x) \sim x^2/2$, alors $\widetilde{J}(x) \sim -2\ln(x)$ en 0. La fonction $\widetilde{J}$ apparaîtra lors de l'étude des semi-flots $\mathbb{Z}^2$-périodiques.

Finalement, nous posons $P(s, w) := H(s) + I(w)$ dès que cette fonction est bien définie.

3.2.2 Propriétés statistiques de semi-flots $\mathbb{Z}^d$-périodiques

Une équation de renouvellement

Le cœur de cette partie est que, pour contrôler $\widetilde{\mathcal{L}}_s$ pour de petites valeurs de s, il suffit de contrôler $\mathcal{L}_{s,w}$ pour de petites valeurs de s et de w. L'outil principal permettant ce contrôle est le lemme suivant, que nous énonçons sous des hypothèses assez générales.

Lemme 3.2.5 (Équation de renouvellement)**.**
Supposons que le rayon spectral de $\mathcal{L}_{s,0}$ est strictement inférieur à 1 pour tout s strictement positif et suffisamment proche de 0. Alors, pour tout s strictement positif et suffisamment proche de 0, pour toute fonction $f \in \mathcal{B}$ positive,

$$\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \sum_{k=0}^{+\infty} \mathcal{L}_{s,w}^k f \, \mathrm{d}w, \tag{3.2.2}$$

où les séries de fonctions convergent dans $\mathbb{L}^1(A, \mu_A)$.

Cette équation est connue - sous différentes formulations, et avec des opérateurs différents - au moins depuis les travaux de M. Pollicott et R. Sharp [57], et a été utilisée dans des articles subséquents [9] [8]. Elle reste vraie, à des modifications mineures près, si l'on remplace l'opérateur de transfert par des opérateurs de Ruelle, sans le jacobien ; on parle alors de “séries de Poincaré tordues”.

Démonstration.
Soit $s > 0$ et suffisamment petit. Soit $f \in \mathcal{B}$ une fonction positive. La fonction $\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f$ est bien définie sur A, si l'on admet qu'elle puisse prendre la valeur $+\infty$. L'équation (3.2.2) repose sur le fait que $\varphi = S_n^{T_A} r$, où n est le nombre de pas qu'il faut au semi-flot pour revenir en $A \times \{0\}$, et que $\ln \widetilde{g} = S_n^{T_A} \ln g$ pour le même entier n. Par conséquent, pour tout $x \in A$:

$$\begin{aligned}
\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f(x) &= \sum_{n=0}^{+\infty} \sum_{\{y:\ \widetilde{T}^n(y) = x\}} e^{-sS_n^{\widetilde{T}} \varphi(y)} e^{S_n^{\widetilde{T}} \ln \widetilde{g}(y)} f(y) \\
&= \sum_{k=0}^{+\infty} \sum_{\{y:\ T_A^k(y) = x\}} e^{-sS_k^{T_A} r(y)} e^{S_k^{T_A} \ln g(y)} 1_{\{0\}}(S_k^{T_A} F(y)) f(y).
\end{aligned}$$

L'étape suivant consiste à utiliser de l'analyse de Fourier, et en particulier la formule d'inversion de Fourier, pour éliminer la fonction $1_{\{0\}}$. Remarquons que $\widehat{1_{\{0\}}} \equiv 1$ sur $\mathbb{T}^d$.

$$\begin{aligned}\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n f(x) &= \frac{1}{(2\pi)^d} \sum_{k=0}^{+\infty} \sum_{\{y:\ T_A^k(y)=x\}} \int_{\mathbb{T}^d} e^{-sS_k^{T_A} r(y)} e^{i\langle w, S_k^{T_A} F(y)\rangle} e^{S_k^{T_A} \ln g(y)} f(y) \, \mathrm{d}w \\ &= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \sum_{k=0}^{+\infty} \mathcal{L}_{s,w}^k f(x) \, \mathrm{d}w.\end{aligned}$$

Par le théorème de Fubini-Lebesgue, on peut inverser la somme et l'intégrale :

$$\frac{1}{(2\pi)^d} \sum_{k=0}^{+\infty} \sum_{\{y:\ T_A^k(y)=x\}} \int_{\mathbb{T}^d} \left| e^{-sS_k^{T_A} r(y)} e^{i\langle w, S_k^{T_A} F(y)\rangle} e^{S_k^{T_A} \ln g(y)} f(y) \right| \, \mathrm{d}w = \sum_{k=0}^{+\infty} \mathcal{L}_{s,0}^k f(x). \qquad (3.2.3)$$

Comme on a supposé que le rayon spectral de $\mathcal{L}_{s,0}$ est strictement plus petit que 1, la suite $\left\| \mathcal{L}_{s,0}^k \right\|_{\mathcal{B} \to \mathcal{B}}$ décroît à vitesse exponentielle, et le membre de droite est bien défini. □

Remarque 3.2.6 (Équation de renouvellement pour d'autres potentiels).
Soit h une fonction mesurable, à valeurs complexe et bornée sur A. Alors on peut définir d'autres opérateurs de Ruelle par :

$$\mathcal{L}_{h,s,w} f(x) = \sum_{y \in T_A^{-1}(\{x\})} e^{h(y) - sr(y) + i\langle w, F\rangle(y)} f(y).$$

Si $h = \ln g$, on retombe sur l'opérateur de transfert. Posons $\widetilde{h}(x) := S_{\varphi(x)}^{T_A} h(x)$ pour presque tout $x \in A$, et :

$$\widetilde{\mathcal{L}}_{\widetilde{h},s} f(x) = \sum_{y \in \widetilde{T}^{-1}(\{x\})} e^{\widetilde{h}(y) - s\varphi(y)} f(y).$$

Sous des hypothèses bien choisies, on peut obtenir une variante de l'équation de renouvellement, qui relie $\widetilde{\mathcal{L}}_{\widetilde{h},s}$ et $\mathcal{L}_{h,s,w}$. En particulier, pour tout $s \in [0,1]$, le cas $h = s \ln g$ est intéressant. En effet, on a $\widetilde{h} = s \ln \widetilde{g}$. Les articles [57], [9] et [8] utilisent ces opérateur avec $s = 0$ pour compter les géodésiques fermées sur des variétés hyperboliques de courbure sectionnelle strictement négative. Nous travaillons ici avec $s = 1$.

Remarque 3.2.7 (Utilisation de l'équation de renouvellement).
Informellement, l'équation (3.2.2) peut être lue comme étant :

$$\left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right)^{-1} = \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \left(\mathrm{Id} - \mathcal{L}_{s,w}\right)^{-1} \, \mathrm{d}w, \qquad (3.2.4)$$

de telle sorte que :

$$\frac{1}{1 - \widetilde{\rho}_s} \sim \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} \frac{1}{1 - \rho_{s,w}} \, \mathrm{d}w, \qquad (3.2.5)$$

où la croissance de l'intégrale quand s tend vers 0 est due aux petites valeurs de w. Informellement, l'équation (3.2.2) relie donc le comportement de la valeur propre principale $\widetilde{\rho}_s$ (qui peut ne pas être définie) de $\widetilde{\mathcal{L}}_s$ pour de petites valeurs de s et le comportement de $\rho_{s,w}$ pour de petites valeurs de s et w.

D'après le travail de la sous-partie 1.2.5, si la base (A, μ_A, T_A) est Gibbs-Markov, si la fonction de saut F est $\sigma(\pi)$-mesurable, si la $\mathbb{Z}^d$ extension est ergodique et conservative, et si la fonction r est lisse, alors l'équation (3.2.4) est rigoureuse.

Temps de premier retour

Notre première application de l'équation de renouvellement consiste à obtenir un équivalent de la queue du temps de premier retour φ en $A \times \{0\} \times \{0\}$. On cherche tout d'abord un équivalent de la transformée de Laplace de ces queues $\int_0^{+\infty} \mathbb{P}(\varphi > t) e^{-st}\, \mathrm{d}t$ pour s proche de 0, puis on utilise un théorème taubérien.

Proposition 3.2.8 (Queues du temps de premier retour).
Sous l'hypothèse 3.2.3, supposons que $\mathcal{B} \subset \mathbb{L}^\infty(A, \mu_A)$ et que $\beta \in [0,1)$. Alors :

$$\mathbb{P}_{\mu_A}(\varphi > t) \sim \frac{1}{\Gamma(1-\beta) G(1/t)} \text{ quand } t \text{ tend vers } +\infty. \tag{3.2.6}$$

Démonstration.
Avant de pouvoir appliquer des résultats sur les opérateurs de transfert, on doit les forcer à apparaître. D'après le théorème de Fubini, pour tout $s > 0$,

$$\int_0^{+\infty} \mathbb{P}_{\mu_A}(\varphi > t) e^{-st}\, \mathrm{d}t = \int_A \int_0^{\varphi(x)} e^{-st}\, \mathrm{d}t\, \mathrm{d}\mu_A(x) = \frac{1 - \int_A e^{-s\varphi}\, \mathrm{d}\mu_A}{s}.$$

C'est ici que l'hypothèse 3.2.3 intervient. Soient β et G le paramètre et la fonction donnés par l'hypothèse 3.2.3. Alors, d'après le commentaire qui la suit :

$$\begin{aligned} 1 &= \int_A \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right) \sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n 1\, \mathrm{d}\mu_A \\ &= \int_A \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right) G(s)(1 + o(1))\, \mathrm{d}\mu_A \\ &= G(s) \int_A (1 - e^{-s\varphi})(1 + o(1))\, \mathrm{d}\mu_A. \end{aligned} \tag{3.2.7}$$

Comme la norme de l'espace $\mathcal{B}$ domine la norme du supremum, la fonction $o(1)$ décroît uniformément, et :

$$\left| \int_A (1 - e^{-s\varphi}) o(1)\, \mathrm{d}\mu_A \right| \leq \int_A |1 - e^{-s\varphi}|\, \mathrm{d}\mu_A \left\| o(1) \right\|_{\mathbb{L}^\infty} = \int_A 1 - e^{-s\varphi}\, \mathrm{d}\mu_A \cdot o(1).$$

On insère cette inégalité dans l'équation (3.2.7), pour finalement obtenir :

$$1 = G(s) \int_A 1 - e^{-s\varphi}\, \mathrm{d}\mu_A (1 + o(1)),$$

et donc :

$$\int_0^{+\infty} \mathbb{P}(\varphi > t) e^{-st}\, \mathrm{d}t \sim \frac{1}{sG(s)}.$$

La fonction G est à variation régulière d'indice $\beta \in [0,1)$ en 0, donc la fonction $s \to (sG(s))^{-1}$ est à variation régulière d'indice $1 - \beta > 0$ en 0. D'après le théorème taubérien de Karamata 1.3.6 :

$$\mathbb{P}(\varphi > t) \sim \frac{1}{\Gamma(1-\beta) G(1/t)}. \qquad \square$$

Remarque 3.2.9 (Cas Gibbs-Markov).
Dans le cas où la $\mathbb{Z}^d$ extension du semi-flot a une base Gibbs-Markov et où les fonction F et r ont de bonnes propriétés de régularité, on a plus simplement :

$$\int_0^{+\infty} \mathbb{P}_{\mu_A}(\varphi > t) e^{-st}\, \mathrm{d}t = \frac{1}{s} \int_A \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right) 1\, \mathrm{d}\mu_A \sim \frac{1 - \tilde{\rho}_s}{s} \ ;$$

l'équation de renouvellement permet de connaître le comportement asymptotique de $(1 - \tilde{\rho}_s)^{-1}$ en 0, et de là celui de $(1 - \tilde{\rho}_s)$.

Temps locaux

Notre deuxième application de l'équation de renouvellement consiste à trouver des asymptotiques en loi pour le temps local ξ_t. Nous employons la méthode des moments.

Lemme 3.2.10 (Moments du temps local).
Supposons que l'hypothèse 3.2.3 est vérifiée. Soit p un entier positif. Soit ν une mesure de probabilité sur A_0 qui est absolument continue par rapport à μ_A, et dont la densité est dans $\mathcal{B}$. Alors :

$$\mathbb{E}_\nu(\xi_t^p) \sim \frac{p!G(1/t)^p}{\Gamma(1+p\beta)} \text{ quand } t \text{ tend vers } +\infty. \tag{3.2.8}$$

Démonstration.
Posons $h := \mathrm{d}\nu/\mathrm{d}\mu_A$. Le cas $p=0$ étant trivial, supposons que p est strictement positif. On cherche d'abord un équivalent quand s tend vers 0 de :

$$\int_0^{+\infty} e^{-st} \,\mathrm{d}\mathbb{E}_\nu(\xi_t^p) = \int_A \int_0^{+\infty} e^{-st} \,\mathrm{d}\xi_t^p(x) h(x) \,\mathrm{d}\mu_A(x).$$

À son n-ième retour, la fonction ξ_t^p croît de $(n-1)^p$ à n^p, donc, pour tout $x \in A$,

$$\int_0^{+\infty} e^{-st} \,\mathrm{d}\xi_t^p(x) = \sum_{n=1}^{+\infty} (n^p - (n-1)^p) e^{-sS_n^{\widetilde{T}}\varphi(x)}.$$

En utilisant le théorème de Fubini-Tonelli, on force les opérateurs de transfert à apparaître :

$$\begin{aligned}
\int_0^{+\infty} e^{-st} \,\mathrm{d}\mathbb{E}_\nu(\xi_t^p) &= \sum_{n=1}^{+\infty} (n^p - (n-1)^p) \int_A e^{-sS_n^{\widetilde{T}}\varphi(x)} h(x) \,\mathrm{d}\mu_A(x) \\
&= \sum_{n=1}^{+\infty} (n^p - (n-1)^p) \int_A \widetilde{\mathcal{L}}^n e^{-sS_n^{\widetilde{T}}\varphi} h(x) \,\mathrm{d}\mu_A(x) \\
&= \sum_{n=1}^{+\infty} (n^p - (n-1)^p) \int_A \sum_{y:\ \widetilde{T}^n y = x} \widetilde{g}^{(n)}(y) e^{-sS_n^{\widetilde{T}}\varphi(y)} h(y) \,\mathrm{d}\mu_A(x) \\
&= \sum_{n=1}^{+\infty} (n^p - (n-1)^p) \int_A \widetilde{\mathcal{L}}_s^n h \,\mathrm{d}\mu_A.
\end{aligned}$$

En manipulant des séries entières, on a alors :

$$\begin{aligned}
\sum_{n=1}^{+\infty} (n^p - (n-1)^p) X^n &= (1-X) \sum_{n=1}^{+\infty} n^p X^n \\
&= \sum_{k=0}^{p} \frac{A(p,k)}{(1-X)^k},
\end{aligned}$$

où les $A(p,k)$ sont des entiers tels que $A(p,p) = p!$ (ils sont liés aux nombres eulériens). Par conséquent,

$$\int_0^{+\infty} e^{-st} \,\mathrm{d}\mathbb{E}_\nu(\xi_t^p) = \int_A \sum_{k=0}^{p} A(p,k) \left(\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n \right)^k h \,\mathrm{d}\mu_A.$$

L'hypothèse 3.2.3 implique que $\left(\sum_{n=0}^{+\infty} \widetilde{\mathcal{L}}_s^n\right)^k h = G(s)^k(1+o(1))$. Comme $A(p,p) = p!$, le terme dominant dans cette expression est $p!G(s)^p$:

$$\int_0^{+\infty} e^{-st} \,\mathrm{d}\mathbb{E}_\nu(\xi_t^p) \sim p!G(s)^p \text{ quand } s \text{ tend vers } 0.$$

La fonction G^p est à variation régulière d'indice $\beta p \in [0,p)$. On conclut en utilisant le théorème taubérien de Karamata 1.3.5. $\square$

A partir de la limite des moments, on déduit la limite en loi du temps local renormalisé.

Proposition 3.2.11 (Asymptotique en loi du temps local).
Supposons que l'hypothèse 3.2.3 est vérifiée. Soit ν une mesure de probabilité sur A_0 qui est absolument continue par rapport à μ_A, et dont la densité est dans $\mathcal{B}$. Alors :

$$\frac{\Gamma(1+\beta)\xi_t}{G(1/t)} \to Y_\beta \text{ quand } t \text{ tend vers } +\infty, \tag{3.2.9}$$

où la convergence est en loi et où Y_β suit une loi de Mittag-Leffler standard de paramètre β.

Démonstration.
La fonction caractéristique d'une loi de Mittag-Leffler a un rayon de convergence strictement positif en 0. Les lois de Mittag-Leffler sont donc caractérisées par leurs moments. Par [10, théorème 30.2], si un processus stochastique $(X_t)_{t\geq 0}$ est tel que tous les moments de X_t existent et convergent vers les moments correspondants d'une variable aléatoire Y_β qui suit une loi de Mittag-Leffler standard de paramètre β, alors X_t converge en loi vers Y_β.

Soit p un entier positif. D'après le lemme 3.2.10,

$$\mathbb{E}_\nu\left(\left(\frac{\Gamma(1+\beta)\xi_t}{G(1/t)}\right)^p\right) \to \frac{p!\Gamma(1+\beta)^p}{\Gamma(1+p\beta)} = \mathbb{E}(Y_\beta^p) \text{ quand } t \text{ tend vers } +\infty,$$

où Y_β suit une loi de Mittag-Leffler standard de paramètre β. □

Remarque 3.2.12 (Cas Gibbs-Markov).
Dans le cas où la $\mathbb{Z}^d$ extension du semi-flot a une base Gibbs-Markov et où les fonction F et r ont de bonnes propriétés de régularité, on a plus simplement :

$$\int_0^{+\infty} e^{-st}\, \mathrm{d}\mathbb{E}_\nu(\xi_t^p) = \int_A \sum_{k=0}^{p} A(p,k) \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right)^{-k} h\, \mathrm{d}\mu_A \sim p! \int_A \left(\mathrm{Id} - \widetilde{\mathcal{L}}_s\right)^{-p} h\, \mathrm{d}\mu_A.$$

Quand le système sous-jacent est Gibbs-Markov, la proposition 3.2.11 peut être renforcée :

Corollaire 3.2.13 (Limite en loi de sommes de Birkhoff).
Supposons que le système dynamique (A, μ_A, T_A) peut être muni d'une structure Gibbs-Markov par rapport à laquelle $\mathbb{E}(D(r))$ est fini et F est $\sigma(\pi)$-mesurable. Supposons que la $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov $(\Omega, \mu, (g_t))$ de base (A, μ_A, T_A), de temps de saut r et de fonction de saut F est ergodique et conservative. Supposons que l'hypothèse 3.2.3 soit vérifiée.

Alors, pour toute fonction $f \in \mathbb{L}^1(\Omega, \mu)$, pour toute mesure de probabilité $\nu \ll \mu$,

$$\frac{\Gamma(1+\beta)}{G(1/t)} \int_0^t f \circ g_s \, \mathrm{d}s \to \int_\Omega f \, \mathrm{d}\mu \cdot Y_\beta \text{ en loi quand } t \text{ tend vers } +\infty, \tag{3.2.10}$$

où le point de départ est choisi suivant la loi ν et Y_β suit une loi de Mittag-Leffler standard de paramètre β.

Démonstration.
Le temps de saut r est presque sûrement strictement positif et lipschitzien sur chaque élément de la partition. On peut donc trouver un cylindre sur lequel il est borné inférieurement par un réel strictement positif. Quitte à induire sur ce cylindre, on peut supposer sans perte de généralité que r est borné inférieurement par un réel $\varepsilon > 0$.

Soit f_ε la fonction qui vaut ε^{-1} sur $A \times [0, \varepsilon] \times \{0\}$, et 0 ailleurs. Alors, pour tout $x \in A_0$, on a :

$$\xi_t(x) \leq \int_0^t f_\varepsilon \circ g_s(x) \, \mathrm{d}s \leq \xi_t(x) + 1.$$

Par la proposition 3.2.11, pour toute fonction $h \in \mathrm{Lip}^\infty$, sous la loi $h\, \mathrm{d}\mu_A$, la convergence (3.2.10) a lieu. Or Lip^∞ est dense dans $\mathbb{L}^1$. Donc, sous toute mesure de probabilité $\nu \ll \mu_A$, la convergence (3.2.10) a lieu.

Supposons que le point de départ x est choisi selon une mesure de probabilité $\nu \ll \mu$. Par l'argument utilisé dans la démonstration de la proposition 3.1.2, quitte à ajouter un décalage de temps $\varphi(x)$, on peut se ramener au cas précédent. Or, comme r est borné inférieurement, la variation du temps local est bornée sur tout intervalle de temps borné. Ce décalage de temps n'a donc aucune influence sur le comportement asymptotique du temps local. Le corollaire 3.2.13 est donc démontré pour la fonction f_ε. Pour le cas général, il suffit d'appliquer le théorème de Hopf pour les semi-flots. □

3.2.3 Perturbation de la valeur propre principale

Comme nous l'avons fait remarquer dans la sous-partie précédente, nous aurons besoin d'un contrôle précis de la valeur propre principale de $\mathcal{L}_{s,w}$ pour pouvoir démontrer des propriétés relatives au temps de premier retour et au temps local de $\mathbb{Z}^d$ extensions de semi-flots Gibbs-Markov. Nous y parviendrons grâce à trois lemmes. Le premier lemme donne un développement de la valeur propre principale de $\mathcal{L}_{s,0}$ au voisinage de 0, sous des hypothèses spectrales sur le système. Le second lemme améliore le premier et permet d'approcher la valeur propre principale de $\mathcal{L}_{s,w}$ au voisinage de 0. Finalement, le troisième lemme traite des valeurs de w loin de l'origine. Nous supposerons, pour le second et le troisième lemme, que le système sous-jacent est Gibbs-Markov ; cette hypothèse est cruciale dans le deuxième lemme, et pratique pour le troisième.

Lemme 3.2.14.
Supposons que l'hypothèse 3.2.2 est vérifiée, que $\mathcal{B} \subset \mathbb{L}^\infty(A, \mu_A)$, et que r satisfait l'une des hypothèses de la sous-sous-partie 3.2.1. Alors la fonction $s \to \rho_{s,0}$ est continue sur un voisinage de 0 dans $\mathbb{R}_+$, et :

$$1 - \rho_{s,0} \sim H(s). \tag{3.2.11}$$

Si en plus le système (A, μ_A, T_A) peut être muni d'une structure Gibbs-Markov pour laquelle $\mathbb{E}(D(r))$ est fini, alors :

$$\|\mathcal{L} - \mathcal{L}_{s,0}\|_{\mathcal{B}\to\mathcal{B}} = O(H(s)), \tag{3.2.12}$$

et :

$$\|1 - f_{s,0}\|_{\mathcal{B}} = O(H(s)). \tag{3.2.13}$$

Démonstration.
On commence par démontrer l'équation (3.2.11). Soit s suffisamment proche 0. Par définition,

$$\rho_{s,0} f_{s,0} = \mathcal{L}\left(e^{-sr} f_{s,0}\right),$$

ce qui donne, en intégrant sur A,

$$\rho_{s,0} = \int_A e^{-sr} f_{s,0} \,\mathrm{d}\mu_A.$$

La famille de fonctions propres $(f_{s,0})$ est continue pour la norme $\mathcal{B}$ [46, théorème 3.16], qui contrôle la norme $\mathbb{L}^\infty$ d'où :

$$1 - \rho_{s,0} = \int_A (1 - e^{-sr}) f_{s,0} \,\mathrm{d}\mu_A = \left(\int_A 1 - e^{-sr} \,\mathrm{d}\mu_A\right)(1 + o(1)).$$

Maintenant, on réécrit l'équation ci-dessus :

$$\frac{1}{s}\int_A 1 - e^{-sr} \,\mathrm{d}\mu_A = \int_A \int_0^{r(x)} e^{-st} \,\mathrm{d}t \,\mathrm{d}\mu_A(x) = \int_0^{+\infty} e^{-st} \mathbb{P}_{\mu_A}(r > t) \,\mathrm{d}t.$$

Si r est intégrable, l'intégrale converge de façon monotone vers $\int_A r \,\mathrm{d}\mu_A$, est nous avons fini. Suppose que r n'est pas intégrable. Si la fonction $t \mapsto \mathbb{P}_{\mu_A}(r > t)$ est à variation régulière d'indice $-\beta \in (-1, 0]$, on utilise le théorème taubérien 1.3.6. Si la fonction $t \mapsto \mathbb{P}_{\mu_A}(r > t)$ est à variation régulière d'indice -1, on utilise le théorème taubérien 1.3.7.

Comme nous avons obtenu un équivalent de $\int_A 1 - e^{-sr} \,\mathrm{d}\mu_A$, dans le cas Gibbs-Markov, les équations (3.2.12) et (3.2.13) découlent du lemme 3.2.4. □

Passons maintenant au développement asymptotique de la valeur propre principale $\rho_{s,w}$.

Lemme 3.2.15.
Soit (A, π, d, μ_A, T) une application Gibbs-Markov mélangeante. Soit $d \geq 0$. Soient F et r, à valeurs dans $\mathbb{Z}^d$ et $\mathbb{R}_+$ respectivement, tels que $\mathbb{E}(D(r))$ soit fini et F soit $\sigma(\pi)$-mesurable, et les fonctions r et F satisfassent l'une des hypothèses correspondantes de la sous-sous-partie 3.2.1.

Supposons que $F \in \mathbb{L}^2$, ou que $r \in \mathbb{L}^q$ pour un certain $q > 1$. Alors, la famille d'opérateurs $(\mathcal{L}_{s,w})$ agissant sur Lip^∞ satisfait :

$$\rho_{s,w} = 1 - P(s,w) + o(P(s,w)) \text{ au voisinage de } 0. \tag{3.2.14}$$

Démonstration.
La démonstration est divisée en quatre parties, avec des sous-parties correspondant aux deux situations possibles dans l'énoncé du lemme. Les deux dernière parties concernent exclusivement le cas $F \in \mathbb{L}^2$.

Sur un voisinage de 0, la valeur propre principale de $\mathcal{L}_{s,w}$ peut s'exprimer de la façon suivante :

$$\begin{aligned}\rho_{s,w} &= \int_A e^{-sr+i\langle w,F\rangle} f_{s,w} \, \mathrm{d}\mu_A \\ &= \int_A e^{-sr+i\langle w,F\rangle} \, \mathrm{d}\mu_A + \int_A \left(e^{-sr+i\langle w,F\rangle} - 1\right)(f_{s,w} - 1) \, \mathrm{d}\mu_A. \end{aligned} \tag{3.2.15}$$

On pose :

$$Q(s,w) := \int_A e^{-sr+i\langle w,F\rangle} \, \mathrm{d}\mu_A.$$

Dans la première partie de la démonstration, on calcule un développement de la fonction Q au voisinage de 0.

Étude de la fonction Q

Découpons Q en morceaux plus faciles à manipuler :

$$Q(s,w) = \int_A e^{i\langle w,F\rangle} \, \mathrm{d}\mu_A + \int_A (e^{-sr} - 1) \, \mathrm{d}\mu_A + \int_A \left(e^{-sr} - 1\right)\left(e^{i\langle w,F\rangle} - 1\right) \, \mathrm{d}\mu_A. \tag{3.2.16}$$

De plus, posons $\overline{I} := I$ si $F \notin \mathbb{L}^2$, et $\overline{I}(w) = \int_A \langle w,F\rangle^2 \, \mathrm{d}\mu_A$ sinon. Posons aussi $\overline{P}(s,w) := H(s) + \overline{I}(w)$.

Maintenant, montrons que le terme croisé dans l'équation (3.2.16) est négligeable. C'est ici que la non-indépendance des hypothèses concernant r et F intervient. Premièrement, remarquons que :

$$\left| \int_A \left(e^{-sr} - 1\right)\left(e^{i\langle w,F\rangle} - 1\right) \, \mathrm{d}\mu_A \right| \leq \|w\| \int_A \|F\| \left(1 - e^{-sr}\right) \, \mathrm{d}\mu_A.$$

A ce niveau, séparons la démonstration en deux sous-parties, correspondant aux deux hypothèses possibles sur le couple (r, F).

Premier cas : Supposons que $F \in \mathbb{L}^2$. Soit $\varepsilon > 0$. Soit $M \geq 0$ tel que, pour tout $t > M$,

$$\mathbb{P}(\|F\| \geq t) \leq \frac{\varepsilon}{t^2}.$$

Soit $t > M$. nous disposons d'une majoration :

$$\|F\| \left(1 - e^{-sr}\right) \leq t \left(1 - e^{-sr}\right) + (\|F\| - t)1_{\|F\| \geq t},$$

d'où :

$$\begin{aligned}\int_A \|F\| \left(1 - e^{-sr}\right) \, \mathrm{d}\mu_A &\leq t \int_A \left(1 - e^{-sr}\right) \, \mathrm{d}\mu_A + \int_t^{+\infty} \mathbb{P}(\|F\| \geq s) \, \mathrm{d}s \\ &\leq t \int_A \left(1 - e^{-sr}\right) \, \mathrm{d}\mu_A + \frac{\varepsilon}{t}.\end{aligned}$$

Pour tout s suffisamment petit, posons :

$$t := \sqrt{\frac{\varepsilon}{\int_A (1-e^{-sr}) \, \mathrm{d}\mu_A}} > M.$$

Alors, pour tout s suffisamment petit :

$$\int_A \|F\| \left(1-e^{-sr}\right) \, \mathrm{d}\mu_A \leq 2\sqrt{\varepsilon \int_A (1-e^{-sr}) \, \mathrm{d}\mu_A}.$$

Cette propriété étant valable pour tout $\varepsilon > 0$, finalement :

$$\|w\| \int_A \|F\| \left(1-e^{-sr}\right) \, \mathrm{d}\mu_A = o(\|w\| \sqrt{H(s)}) = o(P(s,w)).$$

Second cas : Supposons que $r \in \mathbb{L}^q$ pour un certain $q > 1$ et que F est dans le bassin d'attraction d'une loi stable de Lévy de paramètre $p \in (1,2]$ symétrique. Alors, pour tout $\varepsilon \in [0, q-1]$:

$$\int_A \left(1-e^{-sr}\right)^{1+\varepsilon} \, \mathrm{d}\mu_A \leq \int_A (sr)^{1+\varepsilon} \, \mathrm{d}\mu_A = s^{1+\varepsilon} \|r\|_{\mathbb{L}^{1+\varepsilon}}^{1+\varepsilon} = O(s^{1+\varepsilon}).$$

Donc, pour tout $\kappa \in (0, p-1)$, pour tout $\varepsilon \in [0, \min(q-1, (p-1)^{-1})]$:

$$\begin{aligned}
\int_A \|F\| \left(1-e^{-sr}\right) \, \mathrm{d}\mu_A &\leq \|F\|_{\mathbb{L}^{p-\kappa}} \left(\int_A \left(1-e^{-sr}\right)^{1+\frac{1}{p-1-\kappa}} \, \mathrm{d}\mu_A\right)^{1-\frac{1}{p-\kappa}} \\
&\leq \|F\|_{\mathbb{L}^{p-\kappa}} \left(\int_A \left(1-e^{-sr}\right)^{1+\varepsilon} \, \mathrm{d}\mu_A\right)^{1-\frac{1}{p-\kappa}} \\
&= O\left(s^{(1+\varepsilon)\left(1-\frac{1}{p-\kappa}\right)}\right).
\end{aligned}$$

Soit $\delta > 0$. Par l'inégalité de Young,

$$\|w\| \int_A \|F\| \left(1-e^{-sr}\right) \, \mathrm{d}\mu_A = O\left(\|w\|^{p+\delta} + s^{(1+\varepsilon)\left(1-\frac{\kappa+\delta}{(p+\delta-1)(p-\kappa)}\right)}\right).$$

Si κ et δ sont choisis suffisamment petits, alors ε peut être choisi tel que l'exposant de s est strictement plus grand que 1. De plus, $\|w\|^{p+\delta} = o(I(w))$ pour tout $\delta > 0$. Par conséquent :

$$\|w\| \int_A \|F\| \left(1-e^{-sr}\right) \, \mathrm{d}\mu_A = o(P(s,w)),$$

ce qu'il fallait démontrer.

Dans les deux cas, nous avons prouvé que $1 - Q(s,w) - \overline{P}(s,w) = o(P(s,w))$.

Étude d'un terme de second ordre

Nous étudions maintenant le dernier membre de l'équation (3.2.15), c'est-à-dire :

$$\begin{aligned}
&\int_A \left(e^{-sr+i\langle w,F\rangle} - 1\right) (f_{s,w} - 1) \, \mathrm{d}\mu_A \\
&\qquad = \int_A e^{-sr} \left(e^{i\langle w,F\rangle} - 1\right) (f_{s,w} - 1) \, \mathrm{d}\mu_A + \int_A \left(e^{-sr} - 1\right) (f_{s,w} - 1) \, \mathrm{d}\mu_A \\
&\qquad = \int_A e^{-sr} \left(e^{i\langle w,F\rangle} - 1\right) (f_{s,w} - f_{s,0}) \, \mathrm{d}\mu_A \\
&\qquad\quad + \int_A e^{-sr} \left(e^{i\langle w,F\rangle} - 1\right) (f_{s,0} - 1) \, \mathrm{d}\mu_A \\
&\qquad\quad + \int_A \left(e^{-sr} - 1\right) (f_{s,w} - f_{s,0}) \, \mathrm{d}\mu_A + \int_A \left(e^{-sr} - 1\right) (f_{s,0} - 1) \, \mathrm{d}\mu_A.
\end{aligned}$$

Étant donné que F est intégrable et que $\mathrm{Lip}^\infty \subset \mathbb{L}^\infty$, d'après le lemme 3.2.14,

$$\int_A e^{-sr}\left(e^{i\langle w,F\rangle}-1\right)(f_{s,0}-1)\ \mathrm{d}\mu_A = O(H(s)\|w\|) = o(P(s,w)).$$

De par le lemme 3.2.4, nous obtenons la borne $\|f_{s,w}-f_{s,0}\|_{\mathcal{B}} = O(\|w\|)$ uniformément en s dans un voisinage de 0, de telle sorte que :

$$\int_A \left(e^{-sr}-1\right)(f_{s,w}-f_{s,0})\ \mathrm{d}\mu_A = O(H(s)\|w\|) = o(P(s,w))\ ;$$

finalement,

$$\int_A \left(e^{-sr}-1\right)(f_{s,0}-1)\ \mathrm{d}\mu_A = O(H(s)^2) = o(P(s,w)).$$

Il nous suffit donc de contrôler le premier terme de la décomposition. C'est un terme en $O(\|w\|^2)$. Si F n'appartient pas à $\mathbb{L}^2$, alors ce terme est négligeable par rapport à $I(w)$, et nous en avons terminé. Dans le reste de la démonstration, nous supposons que $F \in \mathbb{L}^2$. Il nous faut alors un développement plus précis de $f_{s,w}-f_{s,0}$. C'est l'objet de la prochaine partie de la preuve.

Contrôle de $f_{s,w}-f_{s,0}$

Le but de cette étape est d'étudier plus en détail $f_{s,w}-f_{s,0}$. La borne $O(\|w\|)$ n'étant plus suffisante, nous allons d'abord dériver la fonction $w \mapsto f_{s,w}-f_{s,0}$ en 0, avant de contrôler le comportement de cette dérivée quand s tend vers 0. Premièrement, remarquons que :

$$\begin{aligned}
\rho_{s,w}-\rho_{s,0} &= \int_A \mathcal{L}_{s,w}f_{s,w}-\mathcal{L}_{s,0}f_{s,0}\ \mathrm{d}\mu_A \\
&= \int_A \mathcal{L}_{s,0}(f_{s,w}-f_{s,0})\ \mathrm{d}\mu_A + \int_A (\mathcal{L}_{s,w}-\mathcal{L}_{s,0})f_{s,0}\ \mathrm{d}\mu_A \\
&\quad + \int_A (\mathcal{L}_{s,w}-\mathcal{L}_{s,0})(f_{s,w}-f_{s,0})\ \mathrm{d}\mu_A \\
&= \int_A e^{-sr}(f_{s,w}-f_{s,0})\ \mathrm{d}\mu_A + \int_A e^{-sr}\left(e^{i\langle w,F\rangle}-1\right)f_{s,0}\ \mathrm{d}\mu_A + O(\|w\|^2).
\end{aligned}$$

Notons $\widetilde{P}_{s,0}$ l'opérateur sur Lip^∞ défini, pour tout $h \in \mathcal{B}$, par :

$$\widetilde{P}_{s,0}h := \int_A e^{-sr}h\ \mathrm{d}\mu_A \cdot f_{s,0}.$$

Alors :

$$\begin{aligned}
\rho_{s,0}(f_{s,w}-f_{s,0}) &= \mathcal{L}_{s,0}(f_{s,w}-f_{s,0}) + (\mathcal{L}_{s,w}-\mathcal{L}_{s,0})f_{s,0} \\
&\quad + (\rho_{s,0}-\rho_{s,w})f_{s,w} + (\mathcal{L}_{s,w}-\mathcal{L}_{s,0})(f_{s,w}-f_{s,0}) \\
&= \mathcal{L}_{s,0}(f_{s,w}-f_{s,0}) + \mathcal{L}_{s,0}\left[\left(e^{i\langle w,F\rangle}-1\right)f_{s,0}\right] \\
&\quad + (\rho_{s,0}-\rho_{s,w})f_{s,0} + O(\|w\|^2) \\
&= \mathcal{L}_{s,0}(f_{s,w}-f_{s,0}) + \mathcal{L}_{s,0}\left[\left(e^{i\langle w,F\rangle}-1\right)f_{s,0}\right] \\
&\quad - \widetilde{P}_{s,0}(f_{s,w}-f_{s,0}) - \widetilde{P}_{s,0}\left(\left(e^{i\langle w,F\rangle}-1\right)f_{s,0}\right) + O(\|w\|^2),
\end{aligned}$$

et donc :

$$\left(\rho_{s,0}\,\mathrm{Id}-\mathcal{L}_{s,0}+\widetilde{P}_{s,0}\right)(f_{s,w}-f_{s,0}) = \left(\mathcal{L}_{s,0}-\widetilde{P}_{s,0}\right)\left[\left(e^{i\langle w,F\rangle}-1\right)f_{s,0}\right] + O(\|w\|^2). \qquad (3.2.17)$$

La quantité $\mathbb{E}(D(\langle w,F\rangle f_{s,0}))$ est finie pour tous $s \geq 0$ et $w \in \mathbb{T}^d$ suffisamment petits. De plus, pour tout $s \geq 0$, l'opérateur $\mathcal{L}_{s,0}-\widetilde{P}_{s,0}$ est continu de l'espace de Banach Lip^1_d des fonctions intégrables f telles que $\mathbb{E}(D(f))$ est fini dans Lip^∞_d. Enfin, l'opérateur $\widetilde{P}_{0,0}$ est la projection canonique sur le sous-espace propre correspondant à la valeur propre 1 de $\mathcal{L}$; l'opérateur $\mathrm{Id} - \mathcal{L} + \widetilde{P}_{0,0}$ est donc inversible. La famille d'opérateurs $\left(\rho_{s,0}\,\mathrm{Id} - \mathcal{L}_{s,0} + \widetilde{P}_{s,0}\right)_{s \geq 0}$ est continue, donc ces opérateurs sont eux aussi inversibles si s est suffisamment petit. Par conséquent, on peut poser :

$$h_{s,w} := \left(\rho_{s,0}\,\mathrm{Id} - \mathcal{L}_{s,0} + \widetilde{P}_{s,0}\right)^{-1} \left(\mathcal{L}_{s,0} - \widetilde{P}_{s,0}\right) \left(\langle w, F\rangle f_{s,0}\right). \tag{3.2.18}$$

D'après l'équation (3.2.17), on a alors :

$$f_{s,w} - f_{s,0} = h_{s,w} + \left(\rho_{s,0}\,\mathrm{Id} - \mathcal{L}_{s,0} + \widetilde{P}_{s,0}\right)^{-1} \left(\mathcal{L}_{s,0} - \widetilde{P}_{s,0}\right) \left(\left(e^{i\langle w,F\rangle} - 1 - \langle w, F\rangle\right) f_{s,0}\right) + O(\|w\|^2).$$

Mais $\left(e^{i\langle w,F\rangle} - 1 - \langle w, F\rangle\right) f_{s,0} = O(\|w\|^2)$ dans Lip^1_d. Finalement, on a donc $f_{s,w} - f_{s,0} = h_{s,w} + O(\|w\|^2)$, où le terme $O(\|w\|^2)$ est uniforme en s.

Il nous reste à démontrer que la fonction $(s, w) \mapsto h_{s,w}$ se comporte bien au voisinage de 0. On calcule :

$$\begin{aligned}\left(\mathcal{L}_{s,0} - \widetilde{P}_{s,0}\right) \left(\langle w, F\rangle f_{s,0}\right) &= \left(\mathcal{L} - \widetilde{P}_{0,0}\right) \left(e^{-sr}\langle w, F\rangle f_{s,0}\right) + \int_A e^{-sr}\langle w, F\rangle f_{s,0}\, \mathrm{d}\mu_A (1 - f_{s,0}) \\ &= \left(\mathcal{L} - \widetilde{P}_{0,0}\right) \left(e^{-sr}\langle w, F\rangle f_{s,0}\right) + o(H(s)),\end{aligned}$$

et l'opérateur $\mathcal{L} - \widetilde{P}_{0,0}$ est continu de Lip^1_d dans Lip^∞_d. Comme $e^{-sr}\langle w, F\rangle f_{s,0}$ tend vers $\langle w, F\rangle$ dans Lip^1_d, le membre de droite de l'équation (3.2.18) tend vers $\mathcal{L}(\langle w, F\rangle)$ dans Lip^∞. Finalement,

$$h_{s,w} - h_{0,w} = O(\|w\|) o_s(1),$$

où :

$$h_{0,w} = (\mathrm{Id} - \mathcal{L})^{-1} \mathcal{L}\left(\langle w, F\rangle\right) \in \mathrm{Lip}^\infty.$$

Fin de la démonstration

D'après la partie précédente, pour tous w et s suffisamment proches de 0, on a $e^{i\langle w,F\rangle} - 1 = \langle w, F\rangle + o(\|w\|)$ dans $\mathbb{L}^2$, et $f_{s,w} - f_{s,0} = h_{s,w} + o(\|w\|)$. Par conséquent :

$$\begin{aligned}&\int_A e^{-sr}\left(e^{i\langle w,F\rangle} - 1\right)\left(f_{s,w} - f_{s,0}\right)\, \mathrm{d}\mu_A \\ &\qquad = \int_A e^{-sr}\langle w, F\rangle h_{s,w}\, \mathrm{d}\mu_A + o(\|w\|^2) \\ &\qquad = \int_A \langle w, F\rangle h_{s,w}\, \mathrm{d}\mu_A + \int_A \left(e^{-sr} - 1\right)\langle w, F\rangle h_{s,w}\, \mathrm{d}\mu_A + o(\|w\|^2).\end{aligned}$$

Remarquons que la famille de fonctions $(h_{s,w})_{s \geq 0}$ est uniformément bornée sur un voisinage de 0. L'étude de $\int_A \left(e^{-sr} - 1\right)\langle w, F\rangle h_{s,w}\, \mathrm{d}\mu_A$ peut être menée comme dans la première partie de la démonstration. De plus,

$$\int_A \langle w, F\rangle \left(h_{s,w} - h_{0,w}\right)\, \mathrm{d}\mu_A = O(\|w\|^2) o(1).$$

Pour finir,

$$\int_A e^{-sr}\left(e^{i\langle w,F\rangle} - 1\right)\left(f_{s,w} - f_{s,0}\right)\, \mathrm{d}\mu_A = \int_A \langle w, F\rangle (\mathrm{Id} - \mathcal{L})^{-1} \mathcal{L}\left(\langle w, F\rangle\right)\, \mathrm{d}\mu_A + o(P(s, w)).$$

Comme $(w, Sw) = \mathbb{E}(\langle w, F\rangle (\mathrm{Id} - \mathcal{L})^{-1} \mathcal{L}(\langle w, F\rangle))$ est la covariance dans le théorème central limite [32, théorème 4.1.4], nous retrouvons bien l'expression de $P(s, w) - H(s)$, et :

$$1 - \rho_{s,w} - P(s, w) = o(P(s, w)).$$

□

Remarque 3.2.16.

Si les variables aléatoires r et F sont indépendantes, alors la technique de la démonstration fonctionne même si r et F ont des queues épaisses. On peut même supporter une certaine mesure de non-indépendance ; il suffit de montrer que :

$$\|w\| \int_A \|F\| \left(1 - e^{-sr}\right) \mathrm{d}\mu_A = o(P(s, w)).$$

Pouvoir approcher $\rho_{s,w}$ par $1 - P(s, w)$ revient à dire qu'au premier ordre, le développement de la valeur propre principale se comporte comme si $S_n r$ et $S_n F$ étaient indépendants. C'est une forme d'indépendance asymptotique. En particulier, le lemme 3.2.15 couvre le cas où $F \in \mathbb{L}^2$ et où r a des queues lourdes. Si l'on se place dans le contexte du chapitre 2, on peut poser $F = X_f$ et $r = \varphi_A$. On dispose donc, par des méthodes spectrales, d'une forme d'indépendance asymptotique différente du théorème 2.1.1. Cela suggère la possibilité d'exploiter ce phénomène pour obtenir une autre démonstration du théorème 2.1.3. En particulier, il serait alors possible d'affaiblir l'hypothèse "$X_{|f|} \in \mathbb{L}^p$ pour un certain $p > 2$" en "$X_{|f|} \in \mathbb{L}^2$".

Remarquons que, de façon fondamentalement similaire, la preuve du théorème 2.1.1 et la preuve du lemme 3.2.15 reposent toutes deux sur la structure Gibbs-Markov du système dynamique sous-jacent.

Maintenant que l'on connaît le comportement de $\rho_{s,w}$ pour de petites valeurs de s et de w, il nous faut un contrôle du rayon spectral de $\mathcal{L}_{s,w}$ pour de petites valeurs de s et de grandes valeurs de w. Nous supposerons pour cela que la $\mathbb{Z}^d$ extension de base (A, μ_A, T_A) et de fonction de saut F n'est pas dégénérée.

Lemme 3.2.17.

Plaçons-nous sous les hypothèses du lemme 3.2.15, et supposons que F n'est pas la somme d'un cobord d'une fonction à valeurs dans un translaté d'un sous-réseau propre de $\mathbb{Z}^d$. Alors, pour tout $(s, w) \neq (0, 0)$, le rayon spectral de $\mathcal{L}_{s,w}$ est strictement plus petit que 1.

Démonstration.

Pour commencer, remarquons que, pour tout $s \geq 0$ et tout $w \in \mathbb{T}^d$, le rayon spectral essentiel de $\mathcal{L}_{s,w}$ est strictement plus petit que 1 [32, corollaire 4.1.3]. Ceci nous permet de travailler avec des valeurs propres et des fonctions propres.

Supposons que $s > 0$. Soit h une fonction propre non nulle de $\mathcal{L}_{s,w}$, et soit ρ la valeur propre associée. Alors, sur $\{h \neq 0\}$,

$$|\rho h| = |\mathcal{L}_{s,w} h| = \left|\mathcal{L}(e^{-sr + i\langle w, F\rangle} e^{i\theta} h)\right| \leq \mathcal{L}(e^{-sr}|h|) < \mathcal{L}|h|.$$

Donc $|\rho||h| < \mathcal{L}|h|$ sur $\{h \neq 0\}$. En intégrant, on obtient $|\rho| < 1$.

Supposons maintenant que $s = 0$ et $w \neq 0$.

Supposons que le spectre de $\mathcal{L}_{0,w}$ contient un point $\rho_{0,w}$ de module supérieur ou égal à 1. Comme le rayon spectral essentiel de $\mathcal{L}_{0,w}$ est strictement plus petit que 1, un tel point est une valeur propre de $\mathcal{L}_{0,w}$; soit $f_{0,w}$ une des fonctions propres non nulles correspondantes.

Par l'inégalité triangulaire,

$$\|\mathcal{L}_{0,w} f_{0,w}\|_{\mathbb{L}^1} = |\rho_{0,w}| \, \|f_{0,w}\|_{\mathbb{L}^1} \leq \|\mathcal{L}|f_{0,w}|\|_{\mathbb{L}^1} = \|f_{0,w}\|_{\mathbb{L}^1} .$$

L'opérateur $\mathcal{L}$ est une contraction sur $\mathbb{L}^1$, donc $|\rho_{0,w}| = 1$. Alors :

$$|f_{0,w}| = |\rho_{0,w} f_{0,w}| = |\mathcal{L}_{0,w} f_{0,w}| \leq \mathcal{L}|f_{0,w}|.$$

Comme $\|f_{0,w}\|_{\mathbb{L}^1} = \|\mathcal{L}|f_{0,w}|\|_{\mathbb{L}^1}$, cette dernière inégalité est en fait une égalité. La fonction $|f_{0,w}| \in \mathbb{L}^1$ est donc une fonction propre non triviale de l'opérateur $\mathcal{L}$, correspondant à la valeur propre 1. Le système dynamique Gibbs-Markov est supposé être mélangeant, donc toute fonction propre correspondant à la valeur propre 1 est constante. La fonction $|f_{0,w}|$ est donc constante. Sans perte de généralité, on peut supposer que $|f_{0,w}| \equiv 1$.

Notons $\theta_{0,w} : A \to \mathbb{T}^1$ la phase de $f_{0,w}$, de telle sorte que $f_{0,w} = e^{i\theta_{0,w}}$. Soit $P_{0,w} \in \mathbb{T}^1$ tel que $\rho_{0,w} = e^{iP_{0,w}}$. Étant donné que $|f_{0,w}| = |\mathcal{L}_{0,w} f_{0,w}|$, pour presque tout x, la quantité $\theta_{0,w} + \langle w, F\rangle$ doit être constante sur $\{y : T_A y = x\}$. Par conséquent :

$$\theta_{0,w} \circ T - \theta_{0,w} = P_{0,w} + \langle w, F\rangle \ [2\pi]. \tag{3.2.19}$$

Pour tout $n \in \mathbb{Z}$, on peut multiplier cette équation par n, et obtenir alors :

$$\mathcal{L}_{0,nw} f_{0,w}^n = \rho_{0,w}^n f_{0,w}^n.$$

L'espace vectoriel Lip^∞ muni de l'opération de multiplication de fonctions est une algèbre, donc $f_{0,w}^n$ appartient aussi à Lip^∞ et est non nul. Le complexe $\rho_{0,w}^n$ est donc une valeur propre de $\mathcal{L}_{0,nw}$. Si w a une coordonnée rationnellement indépendante de π, alors 0 est un point d'accumulation de la suite $(nw)_{n\in\mathbb{Z}}$. Cela contredit le fait que, par le lemme 3.2.15, si w' est suffisamment proche de 0 mais n'est pas nul, alors les valeurs propres de $\mathcal{L}_{0,w'}$ sont toutes de module strictement inférieur à 1.

Toutes les coordonnées de w sont donc des multiples rationnels de π. Soit $q > 0$ tel que $qw = 0$. Alors, par l'équation (3.2.19),

$$q\theta_{0,w} \circ T - q\theta_{0,w} = qP_{0,w} + \langle qw, F\rangle \ [2\pi q],$$

et, quitte à ajouter un cobord, la fonction F est à valeurs dans un translaté d'un sous-réseau propre de $\mathbb{Z}^d$, ce qui entre en contradiction avec les hypothèses du lemme. □

3.2.4 Perturbations de la valeur propre principale de l'opérateur induit

Pour démontrer la prochaine proposition, nous partons du développement asymptotique de la valeur propre principale des opérateurs $(\mathcal{L}_{s,w})$, obtenu dans la sous-partie précédente, et l'équation de renouvellement, pour obtenir une fonction G vérifiant les propriétés de l'hypothèse 3.2.3. Ceci requiert des restrictions supplémentaires sur la dimension d et la fonction F, car il ne faut pas que le système soit transient. En particulier, on aura toujours $d \leq 2$.

Proposition 3.2.18.
Soit $(A, \pi, d_A, \mu_A, T_A)$ une application Gibbs-Markov mélangeante. Soit $d \in \{1, 2\}$. Soient F et r des fonctions à valeurs dans $\mathbb{Z}^d$ et dans $\mathbb{R}_+$ respectivement, telles que F est $\sigma(\pi)$-mesurable, r est presque sûrement strictement positive, $\mathbb{E}(D(r))$ est fini, les deux fonctions r et F satisfont une des hypothèses respectives de la sous-sous-partie 3.2.1, et à cobord près F ne prend pas ses valeurs dans un translaté d'un sous-réseau strict de $\mathbb{Z}^d$. Supposons de plus que $F \in \mathbb{L}^2$ ou que $r \in \mathbb{L}^q$ pour un certain $q > 1$.

Si $d = 1$, pour tout $f \in \mathrm{Lip}^\infty$,

$$\frac{1}{2\pi} \int_{\mathbb{T}} (\mathrm{Id} - \mathcal{L}_{s,w})^{-1} \, \mathrm{d}w f = \frac{1}{p \sin(\pi/p)} \frac{I^*(H(s))}{H(s)} \left(\int_A f \, \mathrm{d}\mu_A + o(1) \right) + O(1). \tag{3.2.20}$$

Si $d = 2$, pour tout $f \in \mathrm{Lip}^\infty$,

$$\frac{1}{(2\pi)^2} \int_{\mathbb{T}^2} (\mathrm{Id} - \mathcal{L}_{s,w})^{-1} \, \mathrm{d}w f = \frac{1}{2\pi \det(M)} \widetilde{J}(J^*(H(s))) \left(\int_A f \, \mathrm{d}\mu_A + o(1) \right) + O(1). \tag{3.2.21}$$

Démonstration.
Soit $V = [0, s_0) \times U$ un voisinage de l'origine dans $\mathbb{R}_+ \times \mathbb{T}^d$ suffisamment petit pour que la valeur propre principale $\rho_{s,w}$ soit bien définie et continue sur V. Pour tout $(s, w) \in V$, notons $Q_{s,w}$ la projection canonique sur le sous-espace propre correspondant à la valeur propre $\rho_{s,w}$ de $\mathcal{L}_{s,w}$.

Quitte à prendre un voisinage encore plus petit, on peut supposer que la famille d'opérateurs $(\mathrm{Id} - \mathcal{L}_{s,w})^{-1}$ restreints au noyau de $Q_{s,w}$ est uniformément bornée pour $(s, w) \in V$. Par le lemme 3.2.17, en dehors de V, les opérateurs $(\mathrm{Id} - \mathcal{L}_{s,w})^{-1}$ sont aussi uniformément bornés. Donc, pour tout $f \in \mathrm{Lip}^\infty$,

$$\begin{aligned}\frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} (\mathrm{Id} - \mathcal{L}_{s,w})^{-1}\, \mathrm{d}w f &= \frac{1}{(2\pi)^d}\int_U (\mathrm{Id} - \mathcal{L}_{s,w})^{-1} Q_{s,w} f\, \mathrm{d}w + O(1)\\ &= \frac{1}{(2\pi)^d}\int_U \frac{1}{1-\rho_{s,w}}\, \mathrm{d}w \left(\int_A f\, \mathrm{d}\mu_A + o(1)\right) + O(1).\end{aligned}$$

Définissons, pour tout $s \in [0, s_0)$,

$$G(s) := \frac{1}{(2\pi)^d}\int_U \frac{1}{1-\rho_{s,w}}\, \mathrm{d}w.$$

Quitte à choisir un voisinage V encore plus petit, la fonction $h(s,w) := 1 - \rho_{s,w} - P(s,w)$ est bien définie sur V. Par le lemme 3.2.15,

$$\lim_{(s,w)\to 0} \frac{h(s,w)}{P(s,w)} = 0.$$

Soit $\delta \in (0,1)$. Quitte à prendre une valeur de s_0 encore plus petite, on peut choisir un paramètre $\varepsilon > 0$ tel que, si U est l'ouvert $\{w \in \mathbb{R}^d : \ I(w) < \varepsilon\}$, alors :

$$\sup_V \frac{|h(s,w)|}{P(s,w)} \le \delta.$$

Alors, sur V :

$$\left|\frac{1}{1-\rho_{s,w}} - \frac{1}{P(s,w)}\right| \le \frac{\delta}{1-\delta}\frac{1}{P(s,w)}.$$

En intégrant sur U, on obtient pour tout $s \in [0, s_0)$:

$$\left|G(s) - \frac{1}{(2\pi)^d}\int_U \frac{1}{P(s,w)}\, \mathrm{d}w\right| \le \frac{\delta}{1-\delta}\frac{1}{(2\pi)^d}\int_U \frac{1}{P(s,w)}\, \mathrm{d}w. \tag{3.2.22}$$

Si $G(s)$ tend vers $+\infty$ quand s tend vers 0, alors la contribution de l'extérieur de n'importe quel voisinage de 0 est négligeable. En d'autres termes, le comportement asymptotique de G est une propriété locale de la fonction $(s,w) \mapsto (\mathcal{L}_{s,w})$ en 0. Par conséquent, toutes les valeurs admissibles de ε donnent la même valeur de $(2\pi)^{-d}\int_U P(s,w)^{-1}\, \mathrm{d}w$ à équivalence asymptotique près. Comme ε peut être choisi aussi petit que souhaité, si $\lim_{s\to 0}(2\pi)^{-d}\int_U P(s,w)^{-1}\, \mathrm{d}w = +\infty$ alors :

$$G(s) \sim \frac{1}{(2\pi)^d}\int_U \frac{1}{P(s,w)}\, \mathrm{d}w. \tag{3.2.23}$$

En dimension 1

Soient $\varepsilon, s > 0$. Notons I^* l'inverse généralisé de I, qui est bien défini sur un voisinage de 0. On peut supposer, sans perte de généralité, que I est continue et strictement croissante, de telle sorte que I^* est le vrai inverse de I. Posons $\gamma := I^*(H(s))$. Alors :

$$\begin{aligned}\frac{1}{2\pi}\int_U \frac{1}{P(s,w)}\, \mathrm{d}w &= \frac{1}{2\pi}\int_{B(0,I^*(\varepsilon))} \frac{1}{H(s)+I(w)}\, \mathrm{d}w\\ &= \frac{1}{\pi H(s)}\int_0^{I^*(\varepsilon)} \frac{1}{1+\frac{I(w)}{I(\gamma)}}\, \mathrm{d}w\\ &= \frac{\gamma}{\pi H(s)}\int_0^{\frac{I^*(\varepsilon)}{\gamma}} \frac{1}{1+\frac{I(\gamma v)}{I(\gamma)}}\, \mathrm{d}v.\end{aligned}$$

Remarquons que $\gamma v \leq I^*(\varepsilon)$ dans cette intégrale. Soit $\kappa \in (0, p-1)$. D'après le théorème de Potter 1.3.3, si ε et s sont suffisamment petits, $I(\gamma v)/I(\gamma) \geq \min(v^{p-\kappa}, v^{p+\kappa})/2$, et :

$$\frac{1}{1+\frac{I(\gamma v)}{I(\gamma)}} 1_{\gamma v \leq I^*(\varepsilon)} \leq \frac{1}{1+\frac{1}{2}\min(v^{p-\kappa}, v^{p+\kappa})}.$$

Par le théorème de convergence dominée,

$$\lim_{\gamma \to 0} \int_0^{\frac{I^*(\varepsilon)}{\gamma}} \frac{1}{1+\frac{I(\gamma v)}{I(\gamma)}} \,\mathrm{d}v = \int_0^{+\infty} \frac{1}{1+x^p} \,\mathrm{d}x = \frac{1}{\operatorname{sinc}(\pi/p)},$$

où la dernière égalité découle d'équations fonctionnelle impliquant la fonction Bêta, la fonction Gamma et le sinus. Finalement :

$$\frac{1}{2\pi} \int_U \frac{1}{P(s,w)} \,\mathrm{d}w \sim \frac{1}{p \sin(\pi/p)} \frac{I^*(H(s))}{H(s)}.$$

Remarquons que $\lim_{s \to 0} H(s) = 0$. De plus, la fonction I est à variation régulière d'indice strictement plus grand que 1, donc I^* est à variation régulière d'indice strictement plus petit que 1. Par conséquent, $\lim_{s \to 0} I^*(H(s))/H(s) = +\infty$, et :

$$G(s) \sim \frac{1}{p \sin(\pi/p)} \frac{I^*(H(s))}{H(s)}. \tag{3.2.24}$$

En dimension 2

Rappelons que l'on peut supposer, grâce aux hypothèses ou par le lemme 3.2.15, qu'il existe une fonction J à variation régulière et un automorphisme M de $\mathbb{R}^2$ tels que $I(w) := J(\|Mw\|)$. Si $F \in \mathbb{L}^2$, alors on peut choisir M comme étant la racine carrée de la matrice de covariance S dans le théorème central limite pour les sommes de Birkhoff de F, et $J(v) = v^2/2$.

Posons $\gamma := J^*(H(s))$.

$$\begin{aligned}
\frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \,\mathrm{d}w &= \frac{1}{(2\pi)^2} \int_{M^{-1}B(0,J^*(\varepsilon))} \frac{1}{H(s)+J(\|Mw\|)} \,\mathrm{d}\,\mathrm{Leb}(w) \\
&= \frac{1}{2\pi \det(M)} \int_0^{J^*(\varepsilon)} \frac{v'}{H(s)+J(v')} \,\mathrm{d}v' \\
&= \frac{1}{2\pi \det(M) H(s)} \int_0^{J^*(\varepsilon)} \frac{v'}{1+\frac{J(v')}{J(\gamma)}} \,\mathrm{d}v' \\
&= \frac{(J^*(H(s)))^2}{2\pi \det(M) H(s)} \int_0^{\frac{J^*(\varepsilon)}{\gamma}} \frac{v}{1+\frac{J(\gamma v)}{J(\gamma)}} \,\mathrm{d}v
\end{aligned}$$

Si $p < 2$, l'intégrale ci-dessus converge quand s tend vers 0, donc les conclusions de la proposition sont trivialement vraies. Si $F \in \mathbb{L}^2$, alors on peut supposer que $J(v) = v^2/2$, et :

$$\frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \,\mathrm{d}w \sim -\frac{1}{2\pi \det(M)} \ln((J^*(H(s)))^2) \sim \frac{1}{2\pi \det(M)} \widetilde{J}(J^*(H(s))).$$

Supposons que $p = 2$, et que $F \notin \mathbb{L}^2$. Soit $N > 0$. Par le théorème de convergence uniforme [11, théorème 1.5.2], $J(\gamma v)/J(\gamma)$ converge uniformément vers v^2 sur $[0, N]$. Ainsi, quand s tends vers 0 :

$$\int_0^{\frac{J^*(\varepsilon)}{\gamma}} \frac{v}{1+\frac{J(\gamma v)}{J(\gamma)}} \,\mathrm{d}v = \frac{\ln(1+N^2)}{2}(1+o(1)) + \frac{H(s)}{(J^*(H(s)))^2} \int_{NJ^*(H(s))}^{J^*(\varepsilon)} \frac{v}{H(s)+J(v)} \,\mathrm{d}v.$$

Comme F n'appartient pas à $\mathbb{L}^2$, on a $J(x) \gg x^2$ pour x proche de 0, et donc $J^*(x) \ll \sqrt{x}$ pour x proche de 0. La fonction $x/J^*(x)^2$ tend vers $+\infty$ quand x tend vers 0, et :

$$\frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \, \mathrm{d}w \sim \frac{1}{2\pi \det(M)} \int_{NJ^*(H(s))}^{J^*(\varepsilon)} \frac{v}{H(s)+J(v)} \, \mathrm{d}v.$$

Remarquons que $J(NJ^*(H(s))) \sim \sqrt{N} H(s)$ quand s tend vers 0. Soit $\varepsilon' > 0$. Choisissons N suffisamment grand, de telle sorte que, pour tout s suffisamment petit, $\varepsilon' J(v) \geq H(s)$ dès que $v \geq NJ^*(H(s))$. Alors :

$$\left| \frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \, \mathrm{d}w - \frac{1}{2\pi \det(M)} \int_{NJ^*(H(s))}^{J^*(\varepsilon)} \frac{v}{J(v)} \, \mathrm{d}v \right| \leq \varepsilon' \int_{NJ^*(H(s))}^{J^*(\varepsilon)} \frac{v}{J(v)} \, \mathrm{d}v.$$

Si la fonction $v/J(v)$ n'est pas intégrable sur un voisinage de 0, alors, par le même argument que nous avons utilisé pour démontrer l'équivalence (3.2.23),

$$\frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \, \mathrm{d}w \sim \frac{1}{2\pi \det(M)} \widetilde{J}(NJ^*(H(s))) - \widetilde{J}(J^*(\varepsilon)) \sim \frac{1}{2\pi \det(M)} \widetilde{J}(J^*(H(s))). \quad (3.2.25)$$

□

Si F appartient à $\mathbb{L}^2$, l'équation (3.2.25) devient :

$$\frac{1}{(2\pi)^2} \int_U \frac{1}{P(s,w)} \, \mathrm{d}w \sim -\frac{1}{2\pi\sqrt{\det(S)}} \ln(H(s)).$$

Si $d = 0$, c'est-à-dire pour des semi-flots de suspension, $\mathcal{L}_{\lambda,0} = \widetilde{\mathcal{L}}_\lambda$, et nous n'avons donc pas besoin de manipuler la famille d'opérateurs $\mathcal{L}_{\lambda,w}$. Par conséquent, la proposition 3.2.11 est vérifiée sous les hypothèses du lemme 3.2.14 : il suffit que $\mathcal{B} \subset \mathbb{L}^\infty(A, \mu_A)$, et que la famille d'opérateurs $\mathcal{L}_{\lambda,0}$ soit continue.

Finalement, nous allons décrire une application curieuse de ces travaux. Soit (S_n) une marche aléatoire sur $\mathbb{Z}$ dont le noyau de transition est à variance finie, et qui est ergodique sur $\mathbb{Z}$. Soit φ son temps de premier retour en l'origine. Soit $f : \mathbb{Z} \mapsto \mathbb{Z}$ une fonction non nulle, de somme nulle, et telle que $X_f \in \mathbb{L}^2$. Pour tout $i \geq 0$, soit $X_{f,i}$ la sommes de $f(S_k)$ le long de la i-ème excursion, et φ_i la longueur de la i-ème excursion.

Considérons le processus $(S_n, \sum_{k=0}^{n-1} f(S_k))_{n \geq 0}$. Si celui-ci vaut $(0,0)$, alors en particulier S_n vaut 0, donc il y a eu un nombre entier d'excursions : on a donc aussi $\sum_{k=0}^{n-1} f(S_k) = \sum_{i=0}^{\tau_n - 1} X_{f,i}$. Le processus $(S_n, \sum_{k=0}^{n-1} f(S_k))_{n \geq 0}$ revient donc en $(0,0)$ exactement aux mêmes instants que le processus faisant des sauts de taille $X_{f,i}$, puis attendant un temps φ_i. Si l'on prend $r := \varphi$ et $F := X_f$ dans le tableau 3.2.4 (ligne $d = 1$, le temps de saut r est à variation régulière d'indice $1/2 \in [0,1)$, et la fonction de saut X_f est dans $\mathbb{L}^2$), alors on obtient des estimées sur les temps de retour en $(0,0)$ du processus $(S_n, \sum_{k=0}^{n-1} f(S_k))$. Par exemple, il existe des constantes C et C' strictement positives telles que :

$$\mathbb{P}\left(\inf \left\{ n > 0 \; : \; S_n = 0 \text{ et } \sum_{k=0}^{n-1} f(S_k) = 0 \right\} > t \right) \sim \frac{C}{t^{\frac{1}{4}}},$$

et :

$$\frac{C'}{n^{\frac{1}{4}}} \sum_{k=0}^{n-1} 1 \left(S_k = 0 \text{ et } \sum_{\ell=0}^{k-1} f(S_\ell) = 0 \right) \to Y_{1/4},$$

où la convergence est en loi et où $Y_{1/4}$ suit une loi de Mittag-Leffler standard de paramètre $1/4$.

Le résultat de nos calculs est résumé dans la table ci-dessous. Voici comment la lire :

- Le système considéré est une $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov mélangeant, de fonction de saut F et de temps de saut r ;
- La fonction de saut F est supposée $\sigma(\pi)$-mesurable et non dégénérée ;

- Le temps de saut r est supposé presque sûrement strictement positif et tel que $\mathbb{E}(D(r))$ est fini ;
- La première colonne, d, est la dimension de l'extension. Le cas $d = 0$ correspond à un semi-flot Gibbs-Markov ;
- Pour tout $\beta \in [0, 1]$, dire que l'"hypothèse sur r" est "VR d'indice β" signifie que la fonction $t \mapsto \mu_A(r > t)^{-1}$ est à variation régulière d'indice β en l'infini. La "variation lente" correspond à $\beta = 0$;
- Pour tout $p \in (1, 2]$, dire que l'"hypothèse sur F" est "VR d'indice p" signifie que la fonction J est à variation régulière d'indice p en 0 ;
- L'"indice de variation" est un réel $\gamma \in [0, 1]$ tel que $\mathbb{P}_{\mu_A}(\varphi > t)^{-1}$ est à variation régulière d'indice γ en l'infini ;
- La "fonction de renormalisation" est une fonction a définie sur un voisinage de l'infini dans $\mathbb{R}_+$, croissante, divergente, et à variation régulière d'indice γ donné par la colonne précédente. Elle est définie de telle sorte que, pour toute fonction f intégrable, $a(t)^{-1} \int_0^t f \circ g_s \, \mathrm{d}s$ converge fortement en loi vers $\int_\Omega f \, \mathrm{d}\mu \cdot Y_\gamma$, où Y_γ suit une loi de Mittag-Leffler standard de paramètre γ.

TABLE 3.1 – Résumé des résultats

d	Hypothèse sur r	Hypothèse sur F	Équivalent de $\mathbb{P}_{\mu_A}(\varphi > t)$	Indice de variation	Fonction de renormalisation
0	VR d'indice 1	-	$\mathbb{P}_{\mu_A}(r > t)$	1	$\frac{t}{\int_0^t \mathbb{P}_{\mu_A}(r > s)\,\mathrm{d}s}$
0	VR d'indice $\beta \in [0,1)$	-	$\mathbb{P}_{\mu_A}(r > t)$	β	$\frac{\mathrm{sinc}(\beta\pi)}{\mathbb{P}_{\mu_A}(r > t)}$
1	$r \in \mathbb{L}^1$	$F \in \mathbb{L}^2$	$\sqrt{\frac{2\int_A r\,\mathrm{d}\mu_A \det S}{\pi}}\frac{1}{\sqrt{t}}$	1/2	$\sqrt{\frac{2}{\pi\int_A r\,\mathrm{d}\mu_A \det S}}\sqrt{t}$
1	VR d'indice 1	$F \in \mathbb{L}^2$	$\sqrt{\frac{2\det S}{\pi}}\sqrt{\frac{\int_0^t \mathbb{P}_{\mu_A}(r > s)\,\mathrm{d}s}{t}}$	1/2	$\sqrt{\frac{2}{\pi\det S}}\sqrt{\frac{t}{\int_0^t \mathbb{P}_{\mu_A}(r > s)\,\mathrm{d}s}}$
1	VR d'indice $\beta \in [0,1)$	$F \in \mathbb{L}^2$	$\frac{\sqrt{\Gamma(1-\beta)}}{\Gamma\left(1-\frac{\beta}{2}\right)}\sqrt{2\det S \cdot \mathbb{P}_{\mu_A}(r > t)}$	$\beta/2$	$\frac{1}{\Gamma\left(1+\frac{\beta}{2}\right)\sqrt{2\Gamma(1-\beta)\det S}}\frac{1}{\sqrt{\mathbb{P}_{\mu_A}(r > t)}}$
1	$r \in \mathbb{L}^q$, $q > 1$	VR d'indice $p \in (1,2]$	$\frac{p\sin(\pi/p)\left(\int_A r\,\mathrm{d}\mu_A\right)^{1-\frac{1}{p}}}{\Gamma(1/p)}\frac{1}{tI^*(1/t)}$	$1-1/p$	$\frac{tI^*(1/t)}{p\sin(\pi/p)\Gamma(2-1/p)\left(\int_A r\,\mathrm{d}\mu_A\right)^{1-\frac{1}{p}}}$
2	$r \in \mathbb{L}^1$ ou VR d'indice $\beta \in (0,1]$	$F \in \mathbb{L}^2$	$\frac{2\pi\sqrt{\det S}}{\beta\ln(t)}$	0	$\frac{\beta\ln(t)}{2\pi\sqrt{\det S}}$
2	Variation lente	$F \in \mathbb{L}^2$	$\frac{2\pi\sqrt{\det S}}{\lvert\ln(\mathbb{P}_{\mu_A}(r > t))\rvert}$	0	$\frac{\lvert\ln(\mathbb{P}_{\mu_A}(r > t))\rvert}{2\pi\sqrt{\det S}}$
2	$r \in \mathbb{L}^q$, $q > 1$	VR d'indice 2 et $\widetilde{J}$ diverge	$\frac{2\pi\det(M)}{\widetilde{J}(J^*(1/t))}$	0	$\frac{\widetilde{J}(J^*(1/t))}{2\pi\det(M)}$

3.3 Théorème limite pour les $\mathbb{Z}^d$-extensions

Nos différents résultats portent sur des processus conservatifs. Implicitement, on suppose donc que les extensions avec lesquelles nous travaillons sont récurrentes. Cette hypothèse est très restrictive : entre autres, elle élimine de notre champ d'étude toutes les $\mathbb{Z}^d$ extensions non dégénérées de dimension $d > 2$.

Notre but est d'appliquer la proposition 3.1.2. Pour cela, il faut pouvoir décrire une $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov ergodique directement comme un semi-flot Gibbs-Markov. La méthode est la même qu'avec les marches aléatoires récurrentes : on s'intéresse au processus de décalage sur les excursions à partir de l'origine. Un outil clef est l'accélération, décrite dans la sous-partie 1.2.5.

Commençons par étudier la régularité des fonctions φ et X_f pour un système et une observable f donnés.

3.3.1 Conditions de régularité

On cherche à employer la proposition 3.1.2. Les hypothèses de cette proposition faisant intervenir des conditions de régularité sur φ et sur X_f, il va nous falloir trouver des conditions suffisantes sur la $\mathbb{Z}^d$ extension et l'observable f pour que ces conditions de régularité soient vérifiées.

Lemme 3.3.1 (Régularité du temps de premier retour).
Pour toute $\mathbb{Z}^d$ extension ergodique d'un semi-flot Gibbs-Markov de base $(A, \pi, \lambda, \mathcal{B}, \mu_A, T)$, de fonction de saut F et de temps de saut r,

$$\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) \leq \frac{\lambda}{\lambda - 1} \mathbb{E}\left(D(r)\right).$$

Démonstration.
Soit $a \in \pi_{\overline{\varphi}}$, et soient x, y dans a. De par la définition de la partition $\pi_{\overline{\varphi}}$, pour tout $k < \overline{\varphi}(x)$, les points $T^k(x)$ et $T^k(y)$ sont dans le même élément de la partition π, et donc $d_A(T_{\overline{\varphi}}(x), T_{\overline{\varphi}}(y)) = \lambda^{\overline{\varphi}(x)-k} d_A(T^k(x), T^k(y))$. Pour tout $0 \leq k < \overline{\varphi}(x)$, notons $a_k(a)$ l'unique élément de π dans lequel sont les points $T^k(x)$ et $T^k(y)$. Alors :

$$\begin{aligned}
|\varphi(x) - \varphi(y)| &= \left| \sum_{k=0}^{\overline{\varphi}(x)-1} \left(r(T^k(x)) - r(T^k(y))\right) \right| \\
&\leq \sum_{k=0}^{\overline{\varphi}(x)-1} |r|_{\mathrm{Lip}_{d_A}(a_k(a))} d_A(T^k(x), T^k(y)) \\
&\leq d_A(T_{\overline{\varphi}}(x), T_{\overline{\varphi}}(y)) \sum_{k=0}^{\overline{\varphi}(x)-1} \lambda^{k-\overline{\varphi}(x)} |r|_{\mathrm{Lip}_{d_A}(a_k(a))} \\
&\leq d_{\overline{\varphi}}(x, y) \sum_{k=0}^{\overline{\varphi}(x)-1} \lambda^{k-\overline{\varphi}(a)+1} |r|_{\mathrm{Lip}_{d_A}(a_k(a))}.
\end{aligned}$$

Cette inégalité est valable pour tous x et y dans a et pour tout $a \in \pi_{\overline{\varphi}}$. On en déduit donc l'inégalité suivante, plus succincte :

$$D_{\overline{\varphi}}(\varphi) \leq \sum_{k=0}^{\overline{\varphi}-1} \lambda^{-k} D(r) \circ T^{\overline{\varphi}-1-k}. \tag{3.3.1}$$

Nous allons maintenant utiliser la structure des $\mathbb{Z}^d$ extensions pour exploiter cette inégalité. Soit M un entier strictement positif. Soit $(q_0, \cdots, q_M) \in (\mathbb{Z}^d)^M$. Supposons qu'il existe un entier $0 \leq k < M$, tel que $q_k = q_M$. Alors le chemin initial $(q_0, \cdots, q_M)$ est la concaténation de deux chemins $(q_0, \cdots, q_{M-N-1})$ et $(q_{M-N}, \cdots, q_M)$ tels que $q_{M-N} = q_M$ et $q_k \neq q_M$ pour tout $M-N < k < M$, et cette partition est unique. Réciproquement, si l'on se donne deux chemins vérifiant ces propriétés, par concaténation, on peut créer un chemin $(q_0, \cdots, q_M)$ tel que $q_k = q_M$ pour un certain $0 \leq k < M$. En appliquant cette observation au chemin $(\overline{S}_0(x), \cdots, \overline{S}_M(x))$ pour tout $x \in A$, on voit que les deux ensembles suivants sont identiques :

- $\{x \in A : \exists\, 0 \leq k < M,\ \overline{S}_k(x) = \overline{S}_M(x)\}$;
- $\bigsqcup_{N=1}^{M} \{x \in A : T^{M-N}(x) \in \{\overline{\varphi} = N\}\}$,

où les ensembles $\{x \in A : T^{M-N}(x) \in \{\overline{\varphi} = N\}\}$ pour $1 < N \leq M$ sont disjoints. Par conséquent, pour tout $M \geq 1$:

$$\sum_{N=1}^{M} 1_{\{\overline{\varphi}=N\}} \circ T^{M-N} \leq 1. \tag{3.3.2}$$

Prenons l'espérance de chaque membre de l'inéquation (3.3.1) :

$$\begin{aligned}\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) &\leq \mathbb{E}\left(\sum_{k=0}^{\overline{\varphi}-1} \lambda^{-k} D(r) \circ T^{\overline{\varphi}-1-k}\right) \\ &= \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(1_{\{\overline{\varphi}>k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right).\end{aligned}$$

Fixons un entier $k \geq 0$. Alors, de par l'inéquation (3.3.2), pour tout $M > k$,

$$\begin{aligned}\mathbb{E}\left(1_{\{\overline{\varphi}\leq M\}} 1_{\{\overline{\varphi}>k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right) &= \sum_{N=k+1}^{M} \mathbb{E}\left(1_{\{\overline{\varphi}=N\}} \cdot D(r) \circ T^{N-1-k}\right) \\ &= \sum_{N=k+1}^{M} \mathbb{E}\left(1_{\{\overline{\varphi}=N\}} \circ T^{M-N} \cdot D(r) \circ T^{M-1-k}\right) \\ &\leq \mathbb{E}\left(D(r) \circ T^{M-1-k}\right) \\ &= \mathbb{E}\left(D(r)\right).\end{aligned}$$

Le système étant supposé récurrent, la fonction $\overline{\varphi}$ est presque sûrement finie, de telle sorte que $\lim_{M\to+\infty} 1_{\{\overline{\varphi}\leq M\}} = 1$ presque sûrement. Par le théorème de convergence monotone,

$$\mathbb{E}\left(D_{\overline{\varphi}}(\varphi)\right) \leq \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(1_{\{\overline{\varphi}>k\}} \cdot D(r) \circ T^{\overline{\varphi}-1-k}\right) \leq \sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{E}\left(D(r)\right) = \frac{\lambda}{\lambda-1} \mathbb{E}\left(D(r)\right). \qquad \square$$

Soit $f \in \mathbb{L}^1(\Omega, \mu)$. On définit sur A (ou, de façon équivalente, sur A_0) les fonctions suivantes :

- Pour tout $q \in \mathbb{Z}^d$, la fonction $X_{f,q}(x) := \int_0^{r(x)} f \circ g_s(x, q, 0)\, \mathrm{d}s$;
- $X_f(x) := \int_0^{\varphi(x)} f \circ g_s(x, 0, 0)\, \mathrm{d}s$.

Alors $X_f(x) = \sum_{k=0}^{\overline{\varphi}(x)-1} X_{f,\overline{S}_k(x)}(T^k x)$. En utilisant le même argument que dans la démonstration du lemme 3.3.1, on obtient :

Lemme 3.3.2 (Régularité de l'observable).
Pour toute $\mathbb{Z}^d$ extension ergodique d'un semi-flot Gibbs-Markov de base $(A, \pi, \lambda, \mathcal{B}, \mu_A, T)$, de fonction de saut F et de temps de saut r, pour toute observable $f \in \mathbb{L}^1(\Omega, \mu)$,

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \frac{\lambda}{\lambda-1} \mathbb{E}\left(\sup_{q\in\mathbb{Z}^d} D\left(X_{f,q}\right)\right). \tag{3.3.3}$$

Démonstration.
La différence entre ce lemme et le lemme 3.3.1 est que la fonction $D\left(X_{f,q}\right)$ dépend de q, donc sa somme pondérée le long d'une excursion dépend de l'excursion en question. Ce problème est résolu en majorant toutes les fonctions $D\left(X_{f,q}\right)$ par $\sup_{q\in\mathbb{Z}^d} D\left(X_{f,q}\right)$, qui ne dépend pas de q. $\square$

En particulier, si $\mathbb{E}\left(D\left(X_{f,q}\right)\right)$ est fini pour tout q et $X_{f,q} \equiv 0$ pour tout q sauf un nombre fini d'entre eux, alors $\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right)$ est aussi fini.

Remarque 3.3.3 (Majorations alternatives).

On peut, en adaptant plus finement la preuve du lemme 3.3.1, obtenir d'autres bornes supérieures sur $\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right)$. Par exemple, en utilisant l'inégalité de Hölder, on peut démontrer que, pour tout $\theta \in [0,1]$,

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \sum_{q \in \mathbb{Z}^d} \left(\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ et } \overline{S}_{\overline{\varphi}-1-k} = q)^{1-\theta} \right) \left\| D\left(X_{f,q}\right) \right\|_{\mathbb{L}^{1/\theta}}.$$

Pour $\theta = 1$, c'est une version plus faible du lemme 3.3.2. Cependant, si l'on prend par exemple $\theta = 0$, cette inégalité devient :

$$\mathbb{E}\left(D_{\overline{\varphi}}(X_f)\right) \leq \sum_{q \in \mathbb{Z}^d} \left(\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ et } \overline{S}_{\overline{\varphi}-1-k} = q) \right) \left\| D\left(X_{f,q}\right) \right\|_{\infty}. \tag{3.3.4}$$

Cette majoration est particulièrement intéressante si la partition π est finie, car dans ce cas les semi-normes $\|D(\cdot)\|_\infty$ et $\mathbb{E}(D(\cdot))$ sont équivalentes, d'où :

$$\mathbb{E}\left(\sup_{q \in \mathbb{Z}^d} D\left(X_{f,q}\right)\right) \simeq \sup_{q \in \mathbb{Z}^d} \left\| D\left(X_{f,q}\right) \right\|_{\infty}.$$

La majoration (3.3.4) *est une amélioration de la borne* (3.3.3) *par l'addition des poids sommables $\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ et } \overline{S}_{\overline{\varphi}-1-k} = q)$. Cette fonction est une fonction de Green pour une sorte de "marche aléatoire inversée". Si la partition π est finie, la fonction F prend un nombre fini de valeurs, et donc est bornée. Les poids $\left(\sum_{k=0}^{+\infty} \lambda^{-k} \mathbb{P}(\overline{\varphi} > k \text{ et } \overline{S}_{\overline{\varphi}-1-k} = q)\right)$ décroissent alors exponentiellement en $\|q\|$. Grossièrement, la régularité de f peut dans une certaine mesure empirer à vitesse exponentielle quand on s'éloigne de l'origine.*

Nous allons maintenant discuter brièvement de la condition d'intégrabilité sur les observables.

3.3.2 Conditions d'intégrabilité

Malheureusement, alors qu'il existe des conditions suffisantes simples sur f qui garantissent que la fonction X_f est suffisamment régulière, on ne dispose pas encore de telles conditions pour garantir que $X_{|f|}$ est dans $\mathbb{L}^p$ pour un certain $p > 1$. Le problème qui se pose est essentiellement le même qu'avec les marches aléatoires : contrôler l'intégrabilité de $X_{|f|}$ à partir de f demande une connaissance précise du comportement des excursions de la marche aléatoire. On se contentera ici du critère général suivant : pour tout $p \geq 1$,

$$\left\| X_{|f|} \right\|_{\mathbb{L}^p}^p \leq \mathbb{E}\left(\left(\sum_{k=0}^{\overline{\varphi}-1} \left\| X_{|f|, \overline{S}_k} \right\|_{\mathbb{L}^\infty} \right)^p \right).$$

Nous donnerons aussi une condition suffisante pour que $X_{|f|}$ soit dans $\mathbb{L}^p(A, \mu_A)$ pour un certain $p > 2$; cette condition est hélas très forte.

Lemme 3.3.4.

Soit $(\Omega, \mu, (g_t)_{t \geq 0})$ une $\mathbb{Z}^d$ extension ergodique et conservative d'un semi-flot Gibbs-Markov de base $(A, \pi, \lambda, \mathcal{B}, \mu_A, T)$, de fonction de saut F et de temps de saut r. Soit $f \in \mathbb{L}^1(\Omega, \mu)$.

Pour tous $1 \leq p^* < p \leq +\infty$, si $X_{|f|,q} \equiv 0$ pour tout $q \in \mathbb{Z}^d$ excepté un nombre fini d'entre eux, et si $X_{|f|,q} \in \mathbb{L}^p(A, \mu_A)$ pour tout $q \in \mathbb{Z}^d$, alors $X_{|f|} \in \mathbb{L}^{p^*}(A, \mu_A)$.

Démonstration.

Soient $1 \leq p^* < p \leq +\infty$, et soit f une fonction satisfaisant les hypothèses du lemme. Étant donné que $X_{|f|,q} \equiv 0$ pour tout $q \in \mathbb{Z}^d$ excepté un nombre fini d'entre eux, on peut trouver un parallélépipède $B \subset \mathbb{Z}^d$ tel que $0 \in B$ et $X_{|f|,q} \equiv 0$ pour tout $q \notin B$. Ensuite, on induit le système sur $A \times \{0\}$ en deux étapes : premièrement nous induisons le système de $A \times \mathbb{Z}^d$ sur $A \times B$, et ensuite de $A \times B$ sur $A \times \{0\}$.

Pour le système induit sur $A \times B$, l'intégrale de $|f|$ le long d'une excursion est égale à l'intégrale de $|f|$ pendant le premier saut, c'est-à-dire $\int_0^{r(x)} |f| \circ g_t(x, 0) \, \mathrm{d}t$: ensuite, le processus vit en dehors de la boîte B jusqu'à l'excursion suivante, donc $f \circ g_s \equiv 0$. La fonction induite sur $A \times B$ est donc donnée par $\tilde{X}_{|f|}(x, q) := X_{|f|,q}(x)$ pour tout $(x, q) \in A \times B$.

De par le lemme 1.2.6, quitte à diviser la mesure restreinte à $A \times B$ par $|B|$, le système dynamique induit sur $A \times B$ est Gibbs-Markov, le temps d'arrêt utilisé étant le premier temps de retour en $A \times B$. Le sous-ensemble $A \times \{0\} \subset A \times B$ est une union d'éléments de la partition Gibbs-Markov de $A \times B$. Soit τ le temps de premier retour en $A \times \{0\}$ pour l'application Gibbs-Markov sur $A \times B$. Alors $\mathbb{P}_{\mu_A}(\tau \geq n)$ décroît à vitesse exponentielle.

La fonction $X_{|f|}$ est la somme de $\tilde{X}_{|f|}$ le long d'une excursion partant de $A \times \{0\}$ dans $A \times B$, et donc :

$$
\begin{aligned}
\mathbb{E}\left(X_{|f|}^{p^*}\right) &\leq \mathbb{E}\left(\left(\sum_{k=0}^{\tau-1} (\max_{q \in B} X_{|f|,q}) \circ T_A^k\right)^{p^*}\right) \\
&= \sum_{n=1}^{+\infty} \mathbb{E}\left(1_{\tau=n}\left(\sum_{k=0}^{n-1} (\max_{q \in B} X_{|f|,q}) \circ T_A^k\right)^{p^*}\right) \\
&\leq \sum_{n=1}^{+\infty} n^{p^*} \mathbb{E}\left(1_{\tau=n} \sum_{k=0}^{n-1} (\max_{q \in B} X_{|f|,q})^{p^*} \circ T_A^k\right).
\end{aligned}
$$

Soit $r > 1$ tel que $rp^* < p$. Par l'inégalité de Hölder,

$$
\begin{aligned}
\mathbb{E}\left(X_{|f|}^{p^*}\right) &\leq \sum_{n=1}^{+\infty} n^{p^*} \left\|\max_{q \in B} X_{|f|,q}^{p^*}\right\|_{\mathbb{L}^r} \|1_{\tau=n}\|_{\mathbb{L}^{\frac{r}{r-1}}} \\
&\leq \sum_{n=1}^{+\infty} n^{p^*} \left\|\max_{q \in B} X_{|f|,q}\right\|_{\mathbb{L}^{rp^*}}^{p^*} \mathbb{P}(\tau = n)^{1-\frac{1}{r}} \\
&\leq \left\|\max_{q \in B} X_{|f|,q}\right\|_{\mathbb{L}^p}^{p^*} \sum_{n=1}^{+\infty} n^{p^*} \mathbb{P}(\tau = n)^{1-\frac{1}{r}}.
\end{aligned}
$$

La queue de τ décroît à vitesse exponentielle et $r > 1$, donc la suite $n^{p^*} \mathbb{P}(\tau = n)^{1-\frac{1}{r}}$ est sommable et la norme $\|X_{|f|}\|_{\mathbb{L}^{p^*}}$ est finie. □

En particulier, si $p > 2$, on peut choisir $p^* > 2$ dans le lemme 3.3.4 et appliquer la proposition 3.1.2.

3.3.3 Résultats principaux

Parmi les hypothèses de la proposition 3.1.2, nous avons adapté les conditions de régularité, ainsi que la condition d'intégrabilité de $X_{|f|}$, au cadre des $\mathbb{Z}^d$ extensions. La condition $\int_\Omega f \, \mathrm{d}\mu = 0$ est naturelle.

Il ne reste à adapter que la condition de variation régulière sur les queues du temps de premier retour en l'origine. On pourrait étudier une version modifiée du semi-flot, où le temps de saut r serait remplacé par sa moyenne $\mathbb{E}(r)$. C'est l'approche qui a été adoptée dans [70], avec le lemme 6.10. Cependant, cette méthode est un peu maladroite, et en outre ajoute artificiellement l'hypothèse $r \in \mathbb{L}^2$. Ce défaut est particulièrement gênant si l'on cherche des résultats sur le gaz de Lorentz à horizon infini en une ou deux dimensions, dans lequel r n'est pas de carré intégrable.

Heureusement, le travail de la partie 3.2 porte ses fruits. Nous pouvons donc énoncer des version du théorème central limite généralisé dans le cadre des $\mathbb{Z}^d$ extensions de semi-flots Gibbs-Markov. Nous pourrions donner une version de ce théorème limite par couple valable de fonction de saut et de temps de saut présent dans la table 3.2.4, mais, par un souci de simplicité, on se contentera des deux cas les plus importants : on supposera que F est de carré intégrable, que r est intégrable, et que la dimension d vaut 1 ou 2.

Corollaire 3.3.5.

Soit $(\Omega, \mu, (g_t)_{t\geq 0})$ une $\mathbb{Z}$ extension d'un semi-flot Gibbs-Markov de base $(A, \pi, \lambda, \mathcal{B}, \mu_A, T)$ mélangeante, de fonction de saut F et de temps de saut r. Supposons que :

- *F n'est pas dégénérée ;*
- *$\mathbb{E}(D(r))$ est fini ;*
- *$F \in \mathbb{L}^2(A, \mu_A)$ et $r \in \mathbb{L}^1(A, \mu_A)$;*

Alors la quantité $\sigma(F)$ est strictement positive et finie, où :

$$\sigma(F)^2 = \int_A F^2 \, \mathrm{d}\mu_A + 2 \sum_{n=1}^{+\infty} \int_A F \cdot F \circ T^n \, \mathrm{d}\mu_A.$$

Soit f une fonction à valeurs réelles mesurable sur Ω. Supposons qu'en plus des hypothèses ci-dessus,

- *Il existe $p > 2$ tel que $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$;*
- *$\int_\Omega f \, \mathrm{d}\mu = 0$;*
- *$\mathbb{E}(\sup_{q\in\mathbb{Z}^d} D(X_{f,q}))$ est fini.*

Alors, pour toute mesure de probabilité ν absolument continue par rapport à μ :

$$\left(\frac{\pi\sigma(F)^2}{2} \cdot \frac{\int_A r \, \mathrm{d}\mu_A}{t} \right)^{\frac{1}{4}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma_{\overline{\varphi}}(f) \sqrt{|\mathcal{N}'|} \mathcal{N},$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur (Ω, ν), où $\mathcal{N}$ et $\mathcal{N}'$ sont des variables aléatoires indépendantes, $\mathcal{N}$ est une variable aléatoire gaussienne standard, et $\mathcal{N}'$ est une variable aléatoire gaussienne centrée de variance $\pi/2$ (de telle sorte que $\mathbb{E}(|\mathcal{N}'|) = 1$). La quantité $\sigma_{\overline{\varphi}}(f)$ est donnée par l'équation (2.1.2) en remplaçant T_A par $\widetilde{T}$, et vaut 0 si et seulement si f est un cobord.

Corollaire 3.3.6.

Soit $(\Omega, \mu, (g_t)_{t\geq 0})$ une $\mathbb{Z}^2$ extension d'un semi-flot Gibbs-Markov de base $(A, \pi, \lambda, \mathcal{B}, \mu_A, T)$ mélangeante, de fonction de saut F et de temps de saut r. Supposons que :

- *F n'est pas dégénérée ;*
- *$\mathbb{E}(D(r))$ est fini ;*
- *$F \in \mathbb{L}^2(A, \mu_A)$ et $r \in \mathbb{L}^1(A, \mu_A)$;*

Alors l'opérateur de covariance $\Sigma(F)$ est défini positif, où pour tous u et v dans $\mathbb{R}^2$:

$$(u, \Sigma(F)v) = \int_A (F,u)(F,v) \, \mathrm{d}\mu_A + \sum_{n=1}^{+\infty} \int_A (F,u) \cdot (F,v) \circ T^n \, \mathrm{d}\mu_A + \sum_{n=1}^{+\infty} \int_A (F,v) \cdot (F,u) \circ T^n \, \mathrm{d}\mu_A.$$

Soit f une fonction à valeurs réelles mesurable sur Ω. Supposons qu'en plus des hypothèses ci-dessus,

- *Il existe $p > 2$ tel que $X_{|f|} \in \mathbb{L}^p(A, \mu_A)$;*
- *$\int_\Omega f \, \mathrm{d}\mu = 0$;*
- *$\mathbb{E}(\sup_{q\in\mathbb{Z}^d} D(X_{f,q}))$ est fini.*

Alors, pour toute mesure de probabilité ν absolument continue par rapport à μ :

$$\left(\frac{2\pi\sqrt{\det(\Sigma(F))}}{\ln(t)} \right)^{\frac{1}{2}} \int_0^t f \circ g_s \, \mathrm{d}s \to \sigma_{\overline{\varphi}}(f) X,$$

où la convergence est en loi quand le membre de gauche est vu comme variable aléatoire sur (Ω, ν), et où X suit une loi de Laplace centrée de paramètre $1/\sqrt{2}$. La quantité $\sigma_{\overline{\varphi}}(f)$ est donnée par l'équation (2.1.2) en remplaçant T_A par $\widetilde{T}$, et vaut 0 si et seulement si f est un cobord.

Nous allons maintenant appliquer ces résultat à un exemple plus concret : le flot géodésique en courbure négative.

3.4 Flot géodésique en courbure négative

Après avoir travaillé avec des systèmes non inversibles, tels que les $\mathbb{Z}^d$ extensions de semi-flots Gibbs-Markov, nous traitons finalement quelques exemples de systèmes inversibles. Comme nous devrons nous ramener au cadre des applications Gibbs-Markov, le choix le plus simple est de remplacer les applications Gibbs-Markov avec des extensions naturelles d'applications Gibbs-Markov. Cette partie traite donc des $\mathbb{Z}^d$ extensions de semi-flots dont la base est l'extension naturelle d'une application Gibbs-Markov.

Nous ne chercherons pas à démontrer des théorèmes généraux sur des tels systèmes : en effet, les variations possibles des différentes hypothèses intervenant dans le problème sont nombreuses, ce qui nous contraindrait à avoir des listes de conditions encore plus longues, et à devoir moduler les énoncés en fonctions des systèmes auxquels on voudrait les appliquer. Nous avons choisi à la place de démontrer un théorème limite pour un exemple de système.

Dans cette partie, $(A_+, \pi, \lambda, \mathcal{B}_+, \mu_+, T_+)$ sera une application Gibbs-Markov et $(A, \mathcal{B}, \mu, T)$ sera son extension naturelle. Par abus de notation, on notera aussi $\mathcal{B}_+$ la ***tribu du futur*** sur A, c'est-à-dire le tiré en arrière de la tribu $\mathcal{B}_+$ sur A_+ par la projection canonique de A sur A_+. De même, la partition π sur A_+ est tirée en arrière en une partition de A, que l'on notera encore π. Le système dynamique (A, μ, T) est alors isomorphe à un sous-décalage de $\pi^{\mathbb{Z}}$.

La notion de temps de séparation permet de définir une distance sur l'extension naturelle : il suffit de prendre en compte le temps de séparation dans le futur comme dans le passé. Posons, pour tous x et y dans A :

$$\begin{aligned} s_+(x,y) &:= \inf\{n \geq 0 : \ T^n x \text{ et } T^n y \text{ n'appartiennent pas au même élément de } \pi\}, \\ s_-(x,y) &:= \inf\{n \geq 0 : \ T^{-n} x \text{ et } T^{-n} y \text{ n'appartiennent pas au même élément de } \pi\}. \end{aligned}$$

A partir de ces temps de séparation, on définit deux pseudo-distances $d_+ := \lambda^{-s_+}$ et $d_- := \lambda^{-s_-}$, ainsi qu'une distance $d := d_+ + d_-$ sur A. pour toute fonction f définie sur A, on peut définir des semi-normes lipschitziennes de la même façon qu'avec des application Gibbs-Markov.

Pour tout $a \in \pi$ dans A, on choisit arbitrairement un élément $x_a = (x_{a,n})_{n \in \mathbb{Z}}$ de a. Ensuite, on définit une fonction $p_+ \ : \ A \to A$ en posant, pour tout $a \in \pi$ et tout $x = (x_n)_{n \in \mathbb{Z}} \in a$,

$$p_+(x)_n := \begin{cases} x_{a,n} & \text{si} \quad n \leq 0 \\ x_n & \text{si} \quad n \geq 0 \end{cases} .$$

Autrement dit, on prend un élément $x \in a$, et on remplace son passé par celui de x_a. À a fixé, l'image de p_+ est l'ensemble des éléments de a dont le passé coïncide avec le passé de x_a. Pour tout $n \geq 0$, le diamètre de $T^{-n} p_+(a)$ est d'au plus λ^{-n} : l'image de p_+ est un morceau de variété instable. De plus, pour tout x dans $p_+(a)$, la pré-image de x par p_+ est un ensemble d'éléments qui ont tous le même futur. Pour tout $n \geq 0$, le diamètre de $T^n p_+^{-1}(\{a\})$ est d'au plus λ^{-n} : la pré-image de x est un morceau de variété stable. La fonction $p_+ \ : \ a \to a$ est une projection sur un morceau de variété instable par holonomie le long du feuilletage par les variétés stables.

Une fonction f est $\mathcal{B}_+$-mesurable si et seulement si $f \circ p_+ = f$; on dit qu'une telle fonction ***ne dépend que du futur***. Les fonctions ayant cette propriété sont d'un intérêt particulier, car elles se factorisent en une fonction sur le système Gibbs-Markov initial. Par conséquent, tout théorème limite pour les observables du système (A_+, μ_+, T_+) est un théorème limite pour les observables de (A, μ, T) qui ne dépendent que du futur.

La notion de $\mathbb{Z}^d$ extension d'un semi-flot Gibbs-Markov a été définie dans la partie 3.3. Si l'on se donne un entier $d \geq 0$, une fonction r sur A qui est mesurable et presque sûrement strictement positive, ainsi qu'une fonction F sur A à valeurs dans $\mathbb{Z}^d$ et $\sigma(\pi)$-mesurable, on peut définir de la même façon une $\mathbb{Z}^d$-extension de base (A, π, d, μ, T), de fonction de saut F et de temps de saut r. Le lemme suivant assure que, sous des hypothèses raisonnable concernant r, on peut supposer que r ne dépend que du futur. Bien que ce soit un résultat classique [15, lemme 1.6], l'expression précise du cobord qui apparaît dans la preuve sera utile par la suite. Nous choisissons donc d'en donner un démonstration.

Lemme 3.4.1.
Soit f une fonction définie sur A et à valeurs réelles. Supposons que $\|D(f)\|_\infty$ est fini. Alors il existe une fonction u qui est bornée par $\lambda(\lambda-1)^{-1}\|D(f)\|_\infty$, uniformément $1/2$-höldérienne, et telle que la fonction $f_+ := f + u \circ T - u$ est $\mathcal{B}_+$-mesurable.

Démonstration.
On pose :

$$u := \sum_{n=0}^{+\infty} f \circ T^n - f \circ T^n \circ p_+.$$

Soit $n \geq 0$. Si x et y sont dans le même élément de $\bigvee_{k=-n}^{+\infty} T^{-k}\pi$, alors $T^k x$ et $T^k y$ sont dans le même élément de π pour tout $k \geq -n$. Dans ce cas, $d(x,y) \leq \lambda^{-n}$ et $|r(x) - r(y)| \leq \lambda^{-n}\|D(f)\|_\infty$. Par conséquent, $|f \circ T^n - f \circ T^n \circ p_+| \leq \|D(f)\|_\infty \lambda^{-n}$. La fonction u est donc bien définie et est bornée par la quantité annoncée.

On calcule :

$$\begin{aligned} f + u \circ T - u &= f + \sum_{n=1}^{+\infty} \left(f \circ T^n - f \circ T^{n-1} \circ p_+ \circ T\right) - \sum_{n=0}^{+\infty} \left(f \circ T^n - f \circ T^n \circ p_+\right) \\ &= f \circ p_+ + \sum_{n=1}^{+\infty} f \circ T^n \circ p_+ - f \circ T^{n-1} \circ p_+ \circ T. \end{aligned}$$

La fonction $f_+ = f + u \circ T - u$ est donc bien $\mathcal{B}_+$-mesurable.

Il reste à montrer que u est $1/2$-höldérienne sur chaque élément de π, avec une semi-norme indépendante de l'élément choisi. Soit $a \in \pi$, et soient x, y dans a.

Si $d_+(x,y) = 0$, alors $p_+(x) = p_+(y)$ et $|f \circ T^n(x) - f \circ T^n(y)| \leq \|D(f)\|_\infty \lambda^{-n} d_-(x,y)$, d'où :

$$|u(x) - u(y)| \leq \frac{\lambda}{\lambda - 1}\|D(f)\|_\infty d_-(x,y).$$

Si $d_-(x,y) = 0$, on décompose la somme :

$$\begin{aligned} |u(x) - u(y)| \leq & \sum_{n=0}^{\lfloor s_+(x,y)/2 \rfloor} |f \circ T^n \circ p_+(x) - f \circ T^n \circ p_+(y)| + |f \circ T^n(x) - f \circ T^n(y)| \\ & + \sum_{n=\lfloor s_+(x,y)/2 \rfloor + 1}^{+\infty} |f \circ T^n(x) - f \circ T^n \circ p_+(x)| + |f \circ T^n(y) - f \circ T^n \circ p_+(y)| \\ \leq & \; 2\|D(f)\|_\infty \left(\frac{\lambda^{\lfloor s_+(x,y)/2 \rfloor - s_+(x,y) + 1}}{\lambda - 1} + \frac{\lambda^{-\lfloor s_+(x,y)/2 \rfloor - 1}}{\lambda - 1}\right) \\ \leq & \; \frac{4\lambda \|D(f)\|_\infty}{\lambda - 1} d_+(x,y)^{\frac{1}{2}}. \end{aligned}$$

Soient x et y des points génériques de a. Soit z l'élément de a tel que les coordonnées positives (c'est-à-dire le futur) de z sont celles de x, et les coordonnées négatives (c'est-à-dire le passé) de z sont celles de y. Alors $d_+(x,z) = d_-(y,z) = 0$, et de plus $d_-(x,z) = d_-(x,y)$ et $d_+(y,z) = d_+(x,y)$. Ainsi :

$$|u(x) - u(y)| \leq |u(x) - u(z)| + |u(z) - u(y)| \leq \frac{4\lambda \|D(f)\|_\infty}{\lambda - 1}\left(d_+(x,y)^{\frac{1}{2}} + d_-(x,y)^{\frac{1}{2}}\right). \qquad \square$$

Appliquons le lemme 3.4.1 au temps de saut r. Supposons $\|D(r)\|_\infty$ fini. On peut remplacer la constante de dilatation de l'application Gibbs-Markov λ par $\sqrt{\lambda}$; ainsi, sans perte de généralité, on peut supposer que les fonction u et r_+ du lemme 3.4.1 sont uniformément lipschitziennes. La fonction r_+ n'est *a priori* pas strictement positive, mais si r est bornée inférieurement par un réel strictement positif, on peut trouver un entier M tel que $\sum_{k=0}^{M-1} r_+ \circ T^k$ est strictement positif. On peut alors travailler avec la $\mathbb{Z}^d$ extension de base $(A, \pi_M, \lambda, \mu, T)$, de fonction de saut $\sum_{k=0}^{M-1} F \circ T^k$ et de temps de saut $\sum_{k=0}^{M-1} r_+ \circ T^k$. Ainsi, si r est bornée inférieurement et $\|D(r)\|_\infty$ est fini, alors on peut supposer sans perte de généralité que r ne dépend que du futur.

Nous sommes prêts à attaquer un premier exemple concret : le flot géodésique sur des variétés hyperboliques périodiques, et dont la période est compacte et de courbure sectionnelle strictement négative. Une partie de la preuve reste valable dans le cadre des gaz de Lorentz, mais des adaptations significatives sont malgré tout nécessaires. Les variétés hyperboliques négatives sont, sous plusieurs aspects, les exemples les plus simples sur lesquels nous pouvons travailler.

Soit M une variété riemannienne compacte, connexe et de courbure sectionnelle strictement négative. Soit $d \in \{1, 2\}$, et soit $p : N \to M$ un $\mathbb{Z}^d$ revêtement connexe de M. Notons μ_M et μ_N les mesures de Liouville sur T^1M et T^1N respectivement. Soit $(g_t)_{t\in\mathbb{R}}$ le flot géodésique sur le fibré tangent unitaire T^1N de N. Armés de ces notations, nous allons démontrer la proposition suivante :

Proposition 3.4.2.

Soit f une fonction définie sur T^1N, à valeurs réelles, à support compact et höldérienne. Supposons que $\int_{T^1N} f \, \mathrm{d}\mu_N = 0$.

Si $d = 1$, alors il existe un réel positif $K(f)$ tel que :

$$\lim_{t\to+\infty} \frac{1}{t^{\frac{1}{4}}} \int_0^t f \circ g_s(x, v) \, \mathrm{d}s = K(f)\sqrt{|\mathcal{N}'|}\mathcal{N} \text{ fortement en loi,}$$

où $\mathcal{N}$ sont $\mathcal{N}'$ indépendantes, où $\mathcal{N}$ suit un loi normale standard, et où $\mathcal{N}'$ suit une loi normale centrée de variance $\pi/2$ (de telle sorte que $\mathbb{E}(|\mathcal{N}'|) = 1$).

Si $d = 2$, alors il existe un réel positif $K(f)$ tel que :

$$\lim_{t\to+\infty} \frac{1}{\sqrt{\ln(t)}} \int_0^t f \circ g_s(x, v) \, \mathrm{d}s = K(f)X \text{ fortement en loi,}$$

où X suit une loi de Laplace centrée de paramètre 1.

Dans les deux cas, $K(f) = 0$ si et seulement si f est un cobord.

La preuve repose sur un isomorphisme entre le flot géodésique sur T^1M et le flot de suspension au-dessus de l'extension naturelle d'une application Gibbs-Markov avec un alphabet fini [14], [15, théorème 3.12]. Rappelons les principales propriétés de ces isomorphismes. On peut trouver un nombre fini de sous-variétés $A_1, \ldots, A_p$ transverses au flot géodésique et une fonction $r : \bigcup_i A_i \to \mathbb{R}_+^*$ telles que :

1. chaque A_i a une structure de boîte engendrée par la variété stable forte et la variété instable faible sur T^1M. En fait, chaque A_i peut être construit comme une union de morceaux de variété stable forte qui intersectent un morceau donné de variété instable forte ; la non-intégrabilité du feuilletage par les variétés stable et instable forte empêche ces boîtes d'être des unions de morceaux de variété instable forte ;
2. la fonction r est bornée inférieurement et supérieurement par des réels strictement positifs, et pour tout $x \in \bigcup_i A_i$, on a $r(x) = \inf\{t > 0 : g_t(x) \in \bigcup_i A_i\}$;
3. la fonction r est höldérienne sur chaque sous-boîte $\{x \in A_i : g_{r(x)}(x) \in A_j\}$;
4. le feuilletage par des variétés stables fortes étant invariant sous le flot, r est constant sur chaque morceau de variété stable forte dans chaque sous-ensemble $\{x \in A_i : g_{r(x)}(x) \in A_j\}$;
5. l'image d'un morceau maximal de variété instable faible inclus dans une boîte A_i par g_r est une union de morceaux maximaux de variété instable faible ;
6. l'image réciproque d'un morceau de variété stable forte inclus dans une boîte A_i par g_r est une union de morceaux maximaux de variété stable forte.

On pose $A := \bigcup_i A_i$, ainsi que $\pi := \{A_1, \ldots, A_p\}$ et $T(x) := g_{r(x)}(x)$ pour tout $x \in A$. Soit μ_A la mesure de Liouville sur A, renormalisée afin d'être une mesure de probabilité. Alors μ_A est invariante par T.

On raffine la partition π en une partition π', où un élément de π' est de la forme $\{x \in A_i : g_{r(x)}(x) \in A_j\}$ pour des boîtes A_i et A_j données. En particulier, r est höldérienne sur chaque élément de π'. Tout point $x \in A$ peut être codé par la suite de éléments de π' à travers lesquelles le flot $(g_t(x))_{t\in\mathbb{R}}$ passe, ou de façon équivalente par la suite des transversales dans π' auxquelles $T^n(x)$ appartient pour tout $n \in \mathbb{Z}$. D'après les propriétés (5) et (6), le système dynamique (A, μ_A, T) est un sous-décalage de type fini de $(\pi')^{\mathbb{Z}}$. On appelle d_A la distance sur A héritée du tenseur métrique sur T^1M. Alors le flot géodésique sur T^1M est mesurablement isomorphe au flot de suspension de base (A, π', d_A, μ_A, T) et de hauteur r. D'après la propriété (4), la fonction hauteur ne dépend que du futur.

Preuve de la proposition 3.4.2.
Encodage du flot géodésique sur T^1N

Nous partons de l'encodage symbolique du flot géodésique sur T^1M présenté ci-dessus, et allons démontrer ici que le flot géodésique sur T^1N est isomorphe à une $\mathbb{Z}^d$ extension d'un flot de suspension dont la base est l'extension naturelle d'une application Gibbs-Markov. Nous suivons la méthode utilisée entre autres dans [57].

Chaque transversale $a \in \pi$ a une structure de boîte, donc chaque pré-image $p^{-1}a$ est homéomorphe à $a \times \mathbb{Z}^d$. Pour chaque a, on choisit un des morceaux de $p^{-1}a$, et on le fait correspondre à $a \times \{0\} \subset \pi \times \mathbb{Z}^d$. On définit ainsi une origine sur $p^{-1}\pi$, ce qui distingue un isomorphisme particulier entre $p^{-1}\pi$ et $\pi \times \mathbb{Z}^d$. De là, on dispose d'un isomorphisme entre $p^{-1}\pi'$ et $\pi' \times \mathbb{Z}^d$.

Alors, le flot géodésique sur T^1N est mesurablement isomorphe à la $\mathbb{Z}^d$ extension de base (A, π', d_A, μ_A, T), de fonction de saut F et de temps de saut r. La fonction r ne dépend que du futur et est höldérienne sur chaque élément de π'. La fonction $x \mapsto g_{r(x)}(x)$ est continue de chaque élément de $p^{-1}\pi'$ sur son image, qui par connexité est contenue dans un seul morceau de $p^{-1}\pi$. Par conséquent, F est constante sur chaque élément de π'.

Le système sur lequel nous travaillons est mesurablement isomorphe à $A \times [0, r] \times \mathbb{Z}^d$. On étend la fonction r à cet espace tout entier en posant $r(x, t, q) := r(x)$. On choisit de une projection p_+ sur des feuilles instables de référence dans chaque A_i, que l'on étend en posant $p_+(x, t, q) = (p_+(x), t, q)$. Cette extension de p_+ est bien définie car r ne dépend que du futur. Finalement, pour toute fonction f intégrable sur $A \times [0, r] \times \mathbb{Z}^d$, on pose :

$$Y_f(x, q) := \int_0^{r(x)} f(x, t, q) \, \mathrm{d}t.$$

On a $Y_f(x, q) = X_{f,q}(x)$ pour tout $x \in A$ et $q \in \mathbb{Z}^d$. Il sera bientôt plus pratique de disposer d'une seule fonction définie sur $A \times \mathbb{Z}^d$, et non d'une famille de fonctions définies sur A et indexées par $\mathbb{Z}^d$. Finalement, on définit une application T' sur $A \times \mathbb{Z}^d$ par $T'(x, q) = (T(x), q + F(x))$.

Propriétés de la projection d'une fonction f sur $A \times \mathbb{Z}^d$

Étant donnée une observable f satisfaisant les hypothèses de la proposition, nous voulons démontrer que Y_f est höldérienne, et qu'à un cobord borné près, elle ne dépend que du futur. Commençons par étudier la régularité de Y_f.

Soit d_N la distance sur T^1N, et soit κ_- tel que la courbure sectionnelle de M est dominée par $-\kappa_- < 0$. Alors il existe une constante K strictement positive et ayant la propriété suivante. Soit $\alpha \in (0, 1]$, et soit f une fonction α-höldérienne et à support compact sur T^1N. Soient x et y deux points dans le même élément $a \in \pi$, et soit $q \in \mathbb{Z}^d$.

Si x et y sont dans le même morceau de variété stable forte, alors $d_N(g_t(x, q), g_t(y, q)) \leq K d_N(x, y) e^{-\kappa_- t}$ pour tout $t \geq 0$, d'où :

$$|f \circ g_t(x, q) - f \circ g_t(y, q)| \leq K |f|_{\mathrm{Hol}_\alpha} e^{-\alpha\kappa_- t} d_N(x, y)^\alpha,$$

et :

$$|Y_f(x, q) - Y_f(y, q)| \leq \frac{K |f|_{\mathrm{Hol}_\alpha}}{\alpha\kappa_-} d_N(x, y)^\alpha.$$

Si x et y sont dans le même morceau de variété instable forte, on peut répéter le même calcul avec des modifications mineures. Cependant, la transversale a est composée de morceaux de variété instable faible. Pour résoudre ce problème, choisissons un point x dans le morceau de variété instable forte inclus dans a, puis prenons des morceaux de variété instable forte qui intersectent le morceau de variété stable forte passant par x ; ce processus engendre une nouvelle transversale b. En choisissant des morceaux de variété stable forte de la bonne taille, on peut faire en sorte que b ait une structure de boîte, et soit l'image de a sous le flot géodésique arrêté à un temps non constant. Plus précisément, on peut faire en sorte qu'il existe une fonction $\delta : a \to \mathbb{R}$ qui est höldérienne, telle que $b = g_\delta(a)$ et $\delta = 0$ sur les morceaux de variété stable et instable fortes passant par x. On répète cette opération pour tout $a \in \pi$, de telle sorte que l'on obtient une fonction δ höldérienne sur la section de Poincaré π, et de là définie et höldérienne sur $p^{-1}\pi$. Soit α' l'exposant de Hölder de la fonction δ.

Soit $a' \in \pi'$. Si x et y sont dans le même morceau de variété instable faible dans a', alors $g_{r(x)}(x)$ et $g_{r(y)}(y)$ sont dans le même morceau de variété instable faible dans une certaine transversale $a \in \pi$. De là, $g_{r(x)+\delta(T'x)}(x)$ et $g_{r(y)+\delta(T'y)}(y)$ sont sur le même morceau de variété instable forte dans b. En inversant le flot, on obtient pour tout $t \geq 0$:

$$d_N(g_{r(x)+\delta(T'x)-t}(x), g_{r(y)+\delta(T'y)-t}(y)) \leq K d_N(g_{r(x)+\delta(T'x)}(x), g_{r(y)+\delta(T'y)}(y)) e^{-\kappa_- t},$$

d'où, en utilisant le fait que $L := r + \delta \circ T' - \delta$ est constant sur les morceaux de variété instable faible dans a' :

$$\begin{aligned}
|Y_f(x,q) & - Y_f(y,q)| \\
&= \left| \int_{\delta(T'x)}^{r(x)+\delta(T'x)} f(g_{r(x)+\delta(T'x)-t}(x,q,0)) \, \mathrm{d}t - \int_{\delta(T'y)}^{r(y)+\delta(T'y)} f(g_{r(y)+\delta(T'y)-t}(y,q,0)) \, \mathrm{d}t \right| \\
&\leq \int_{\max\{\delta(T'x),\delta(T'y)\}}^{L+\min\{\delta(x),\delta(y)\}} |f|_{\mathrm{Hol}_\alpha} d_N(g_{r(x)+\delta(T'x)-t}(x), g_{r(y)+\delta(T'y)-t}(y))^\alpha \, \mathrm{d}t \\
&\qquad + |\delta(x) - \delta(y)| \, \|f\|_\infty + |\delta(T'x) - \delta(T'y)| \, \|f\|_\infty \\
&\leq \frac{K^\alpha |f|_{\mathrm{Hol}_\alpha}}{\alpha \kappa_-} d_N(g_{r(x)+\delta(T'x)}(x), g_{r(y)+\delta(T'y)}(y))^\alpha \\
&\qquad + |\delta|_{\mathrm{Hol}_{\alpha'}(a)} \, \|f\|_\infty \, d_N(x,y)^{\alpha'} + |\delta|_{\mathrm{Hol}_{\alpha'}(Ta)} \, \|f\|_\infty \, d_N(T'x, T'y)^{\alpha'}.
\end{aligned}$$

Soit $\lambda > 1$ la constante de dilatation associée à l'application $x \mapsto g_{r(x)}(x)$ sur les feuilles instables de $T^1 M$. Il existe une constante $K' > 0$ telle que $K' d_N(x,y) \leq d(x,y) := \lambda^{-s_+(x,y)} + \lambda^{-s_-(x,y)}$, et le sous-décalage sur A_+ est Gibbs-Markov pour la distance λ^{-s_+}. Alors la fonction Y_f est bornée par $\|f\|_\infty r$ et est $\min\{\alpha, \alpha'\}$-höldérienne pour la distance d sur chaque sous-ensemble $A \times \{q\}$, uniformément en q.

Vers le théorème limite

Nous introduisons maintenant un cobord bien choisi et qui, une fois ajouté à f, produira un fonction qui ne dépend que du futur. Nous serons alors capables de factoriser le système dynamique sur l'application Gibbs-Markov sous-jacente, et utiliser les résultats de la sous-partie 3.3.3. Nous utilisons la même méthode que dans la preuve du lemme 3.4.1. Définissons des fonctions u et f_+, sur $A \times \mathbb{Z}^d$ et $T^1 N$ respectivement, par :

$$u(x,q) := Y_{f - f \circ p_+} + \sum_{n=1}^{+\infty} \left(Y_f \circ T'^n(x,q) - Y_f \circ T'^n \circ p_+(x,q) \right),$$

$$f_+(x,t,q) := f \circ p_+(x,t,q) + \frac{1}{r(x)} \sum_{n=1}^{+\infty} \left(Y_f \circ T'^n \circ p_+(x,q) - Y_f \circ T'^{n-1} \circ p_+ \circ T'(x,q) \right).$$

On étend u à $T^1 N$ en posant, pour tout $t < r(x)$:

$$u(x,t,q) := u(x,q) + \int_0^t (f_+ - f)(x,t,s) \, \mathrm{d}s.$$

La fonction r est höldérienne et strictement positive. Les fonctions u et f_+ sont bornées et $\min\{\alpha, \alpha'\}/2$-höldériennes, et $f_+ = f_+ \circ p_+$. La fonction $f_+ - f$ est un cobord borné pour le flot géodésique, de même que la fonction $Y_{f_+} - Y_f$ pour la transformation T et la fonction $X_{f_+} - X_f$ pour la transformation T_A.

Le temps de saut r et la nouvelle observable f_+ ne dépendent que du futur. La proposition 3.4.2 est une conséquence des corollaires 3.3.5 et 3.3.6 appliqués à f_+. Il reste à vérifier que f_+ remplit les conditions d'application de ces corollaires. Cependant, le critère "$X_{|f_+|} \in \mathbb{L}^p$ pour un certain $p > 2$ et $\int f_+ \, \mathrm{d}\mu = 0$" est ici trop restrictif, car il se comporte mal quand un cobord borné est ajouté à f. Par la remarque 2.3.4, cette condition peut être remplacée par :

- $\sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \, \mathrm{d}s \right| \in \mathbb{L}^p(A, \mu_A)$ pour un certain $p > 2$;
- $\int_A X_{f_+} \, \mathrm{d}\mu_A = 0$.

Quitte à changer de distance sur A, toutes les fonctions considérées sont uniformément lipschitziennes, et alors $\mathbb{E}(\sup_{q \in Z^d} D(X_{f_+,q}))$ est fini.

La section π est composée d'un nombre fini de boîtes, ce qui implique que le temps de saut r est borné. La famille de fonctions $X_{|f|,q}$ est alors uniformément bornée. De plus, $p^{-1}\pi$ est localement fini, donc il y a seulement un nombre fini d'éléments de $p^{-1}\pi$ qui sont à distance moins de $\|r\|_\infty$ du support de f, qui est supposé être compact. Par conséquent, $X_{|f|,q} \equiv 0$ pour tout q saut au plus un nombre fini d'entre eux, et $X_{|f|,q}$ est borné pour tout q. D'après le lemme 3.3.4, on sait que $X_{|f|} \in \mathbb{L}^p$ pour tout p fini. Alors :

$$\left\| \sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \, \mathrm{d}s \right| \right\|_{\mathbb{L}^p} \leq \left\| \sup_{0 \leq t \leq \varphi} \left| \int_0^t f \circ g_s \, \mathrm{d}s \right| \right\|_{\mathbb{L}^p} + 2 \|u\|_{\mathbb{L}^\infty} \leq \left\| X_{|f|} \right\|_{\mathbb{L}^p} + 2 \|u\|_{\mathbb{L}^\infty} ,$$

de telle sorte que $\sup_{0 \leq t \leq \varphi} \left| \int_0^t f_+ \circ g_s \, \mathrm{d}s \right|$ appartient aussi à $\mathbb{L}^p$ pour tout p fini. Finalement, $\int_A X_{f_+} \, \mathrm{d}\mu_A = \int_A X_f \, \mathrm{d}\mu_A = \int_{T^1 N} f \, \mathrm{d}\mu_N = 0$. □

Les billards périodiques, comme le gaz de Lorentz, sont plus difficiles à manipuler. Par exemple, la fonction de saut F et le temps de saut r sont non bornés. Par conséquent, l'argument utilisé pour montrer que $X_{|f|,q} = 0$ pour tout q sauf un nombre fini d'entre eux ne fonctionne pas ; cette propriété est fausse en général pour les billards. Nous n'avons actuellement pas les outils pour montrer un théorème qui s'applique, par exemple, aux fonctions à support compact, et l'hypothèse d'intégrabilité pour les observables est donc plus technique.

Bibliographie

[1] J. Aaronson, *An ergodic theorem with large normalising constants*, Israel Journal of Mathematics, **38** (1981), 182-188.

[2] J. Aaronson, *An introduction to infinite ergodic theory*, American Mathematical Society, 1997.

[3] J. Aaronson et M. Denker, *A local limit theorem for stationary processes in the domain of attraction of a normal distribution*, *Asymptotic methods in probability and statistics with applications (St. Petersburg, 1998)*, 215-223, Birkhäuser Boston, 2001.

[4] J. Aaronson et M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stochastics and Dynamics, **1** (2001), 193-237.

[5] J. Aaronson et O. Sarig, *Exponential chi-squared distributions in infinite ergodic theory*, preprint

[6] D.V. Anosov, *Ergodic properties of geodesic flows on closed Riemannian manifolds of negative curvature*, Doklady Akademii Nauk SSSR, **151** (1963), 1250-1252.

[7] A. Avila, D. Dolgopyat, E. Duriev et O. Sarig, *The visits to zero of a random walk driven by an irrational rotation*, preprint.

[8] M. Babillot et F. Ledrappier, *Geodesic paths and horocycle flows on abelian covers*, Lie groups and ergodic theory (Mumbai, 1996), 1-32.

[9] M. Babillot et F. Ledrappier, *Lalley's theorem on periodic orbits of hyperbolic flows*, Ergodic Theory and Dynamical Systems, **18** (1998), 17-39.

[10] P. Billingsley, *Probability and measure*, Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, 1979.

[11] N.H. Bingham, C.M. Goldie et J.L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, 1987.

[12] G.D. Birkhoff, *Proof of the ergodic theorem*, Proceedings of the National Academy of Sciences of the United States of America, **17** (1931), 656-660.

[13] G. Boole, *On the comparison of transcendents, with certain applications to the theory of definite integrals*, Philosophical Transactions of the Royal Society of London, **147** (1857), 745-803.

[14] R. Bowen, *Symbolic dynamics for hyperbolic flows*, American Journal of Mathematics, **95** (1973), 429-460.

[15] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, 1975.

[16] R. Bowen et C. Series, *Markov maps associated with Fuchsian groups*, Institut des Hautes Études Scientifiques, Publications Mathématiques, **50** (1979), 153-170.

[17] D.L. Burkholder, *Distribution function inequalities for martingales*, The Annals of Probability, **1** (1973), 19-42.

[18] F. Castell, N. Guillotin-Plantard, F. Pène et B. Schapira, *A local limit theorem for random walks in random scenery and on randomly oriented lattices*, The Annals of Probability, **39** (2011), 2079-2118.

[19] F. Castell, N. Guillotin-Plantard et F. Pène, *Limit theorems for one and two-dimensional random walks in random scenery*, preprint.

[20] F. Castell, N. Guillotin-Plantard, F. Pène et B. Schapira, *On the local time of random processes in random scenery*, preprint.

[21] J.-R.Chazottes et S. Gouëzel, *On almost-sure versions of classical limit theorems for dynamical systems*, Probability Theory and Related Fields, **138** (2007), 195-234.

[22] E. Csáki et A. Földes, *On asymptotic independence and partial sums*, Asymptotic methods in probability and statistics, A volume in honour of Miklós Csörgõ, Elsevier (2000), 373-381.

[23] E. Csáki et A. Földes, *Asymptotic independence and additive functionals*, Journal of Theoretical Probability, **13** (2000), 1123-1144.

[24] R.L. Dobrushin, *Two limit theorems for the simplest random walk on a line* (russe), Uspekhi Matematicheskikh Nauk, **10** (1955), 139-146.

[25] D. Dolgopyat, *Prevalence of rapid mixing in hyperbolic flows*, Ergodic Theory and Dynamical Systems, **18** (1998), 1097-1114.

[26] D. Dolgopyat, D. Szász et T. Varjú, *Recurrence properties of planar Lorentz process*, Duke Mathematical Journal, **142** (2008), 241-281.

[27] A. Erdélyi, W. Magnus, F. Oberhettinger et F.G. Tricomi, *Higher transcendental functions. Vol. III.*, McGraw-Hill Book Company, 1955.

[28] P. Gaspard et X.J. Wang, *Sporadicity : Between periodic and chaotic dynamical behaviors*, Proceedings of the National Academy of Science, **85** (1988), 4591-4595.

[29] M.I. Gordin, *The central limit theorem for stationary processes* (russe), Doklady Akademii Nauk SSSR, **188** (1969), 739-741.

[30] S. Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel Journal of Mathematics, **139** (2004), 29-65.

[31] S. Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probability Theory and Related Fields, **128** (2004), 82-122.

[32] S. Gouëzel, *Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes*, PhD thesis, 2008 version.

[33] S. Gouëzel, *Almost sure invariance principle for dynamical systems by spectral methods*, The Annals of Probability, **38** (2010), 1639-1671.

[34] P. Hall et C.C. Heyde, *Martingale limit theory and its application*, Probability and Mathematical Statistics. Academic Press, 1980.

[35] B. Hasselblatt et A. Katok, *Introduction to the modern theory of dynamical systems*, Cambridge University Press, 1996.

[36] B. Hasselblatt et A. Katok, *Handbook of dynamical systems, Volume 1A*, Elsevier, 2002.

[37] H. Hennion, *Sur un théorème spectral et son application aux noyaux lipschitziens*, Proceedings of the American Mathematical Society, **118** (1993), 627-634.

[38] M. Holland, *Slowly mixing systems and intermittency maps*, Ergodic Theory and Dynamical Systems, **25** (2005), 133-159.

[39] E. Hopf, *Ergodentheorie* (allemand), Springer, Berlin, 1937.

[40] I.A. Ibragimov, *Some limit theorems for stationary processes*, Theory of Probability and its Applications, **7** (1962), 349-382.

[41] I.A. Ibragimov et Y.V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971.

[42] C.T. Ionescu-Tulcea et G. Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Annals of Mathematics, **52** (1950), 140-147.

[43] N.C. Jain et W.E. Pruitt, *Maximal increments of local time of a random walk*, The Annals of Probablity, **15** (1987), 1461-1490.

[44] A. Jakubowski, *Tightness criteria for random measures with application to the principle of conditioning in Hilbert spaces*, Probability and Mathematical Statistics, **9** (1988), 95-114.

[45] M. Kac, *On the notion of recurrence in discrete stochastic processes*, Bulletin of the American Mathematical Society, **53** (1947), 1002-1010.

[46] T. Kato, *Perturbation theory for linear operators*, reprint of the 1980 edition, Classics in Mathematics, Springer-Verlag, Berlin (1995).

[47] A. Katsuda et T. Sunada, *Closed orbits in homology classes*, Institut des Hautes Études Scientifiques. Publications Mathématiques **71** (1990), 5-32.

[48] J. Korevaar, *Tauberian theory*, Grundlehren der Mathematischen Wissenschaften, **329**, Springer-Verlag, Berlin (2004).

[49] A. Lasota et J.A. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Transactions of the American Mathematical Society, **186** (1973), 481-488.

[50] F. Ledrappier et O. Sarig, *Unique ergodicity for non-uniquely ergodic horocycle flows*, Discrete and Continuous Dynamical Systems, **16** (2006), 411-433.

[51] F. Ledrappier et O. Sarig, *Invariant measures for the horocycle flow on periodic hyperbolic surfaces*, Israel Journal of Mathematics, **160** (2007), 281-315.

[52] F. Ledrappier et O. Sarig, *Fluctuations of ergodic sums for horocycle flows on $\mathbb{Z}^d$-covers of finite volume surfaces*, Discrete and Continuous Dynamical Systems, **22** (2008), 247-325.

[53] C. Liverani, B. Saussol et S. Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory and Dynamical Systems, **19** (1999), 671-685.

[54] M.B. Marcus et J. Rosen, *Laws of the iterated logarithm for the local times of recurrent random walks on $\mathbb{Z}^2$ and of Lévy processes and random walks in the domain of attraction of Cauchy random variables*, Annales de l'Institut Henri Poincaré (B). Probabilités et Statistiques, **30** (1994), 467-499.

[55] I. Melbourne et D. Terhesiu, *First and higher order uniform dual ergodic theorems for dynamical systems with infinite measure*, à paraître dans l'Israel Journal of Mathematics.

[56] G. Pisier, *Martingales in Hilbert spaces (in connection with type and cotype)*, URL : http ://people.math.jussieu.fr/~pisier/ihp-pisier.pdf (version du 9 février 2011).

[57] M. Pollicott et R. Sharp, *Orbit counting for some discrete groups acting on simply connected manifolds with negative curvature*, Inventiones Mathematicae, **117** (1994), 275-302.

[58] Y. Pomeau et P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Communications in Mathematical Physics, 74 (1980), 189-197.

[59] P. Révész, *Random walk in random and nonrandom environments*, World Scientific (1990).

[60] H.P. Rosenthal, *On the subspaces of $\mathbb{L}^p$ $(p > 2)$ spanned by sequences of independent random variables*, Israel Journal of Mathematics, **8** (1970), 273-303.

[61] F. Schweiger, *Numbertheoretical endomorphisms with σ-finite invariant measure*, Israel Journal of Mathematics, **21** (1975), 308-318.

[62] R.J. Serfling, *Moment inequalities for the maximum cumulative sum*, The Annals of Mathematical Statistics, **41** (1970), 1227-1234.

[63] R. Sharp, *Closed orbits in homology classes for Anosov flows*, Ergodic Theory and Dynamical Systems, **13** (1993), 387-408.

[64] D. Szász et T. Varjú, *Local limit theorem for the Lorentz process and its recurrence in the plane*, Ergodic Theory and Dynamical Systems, **24** (2004), 257-278.

[65] D. Szász et T. Varjú, *Markov towers and stochastic properties of billiards*, Modern dynamical systems and applications, Cambridge University Press, Cambridge (2004), 433-445.

[66] D. Szász et T. Varjú, *Limit laws and recurrence for the planar Lorentz process with infinite horizon*, Journal of Statistical Physics, **129** (2007), 59-80.

[67] M. Thaler, *Estimates of the invariant densities of endomorphisms with indifferent fixed points*, Israel Journal of Mathematics, **37** (1980), 303-314.

[68] M. Thaler, *A limit theorem for the Perron-Frobenius operator of transformations on* $[0,1]$ *with indifferent fixed points*, Israeli Journal of Mathematics, 91 (1995), 111–127.

[69] D. Thomine, *A generalized central limit theorem in infinite ergodic theory*, à paraître dans Probability Theory and Related Fields.

[70] D. Thomine, *Variations on a central limit theorem in infinite ergodic theory*, à paraître dans Ergodic Theory and Dynamical Systems.

[71] D. Thomine, *Local time and first return time for periodic semi-flows*, preprint.

[72] W. Woess, *Random walks on infinite graphs and groups*, Cambridge Tracts in Mathematics, Cambridge University Press, 2000.

[73] R. Zweimüller, *Ergodic properties of infinite measure preserving interval maps with indifferent fixed points*, Ergodic Theory and Dynamical Systems, **20** (2000), 1519-1549.

[74] R. Zweimüller, *Ergodic structure and invariant densities of non-markovian interval maps with indifferent fixed points*, Nonlinearity, **11** (1998), 1263-1276.

[75] R. Zweimüller, *Mixing limit theorems for ergodic transformations*, Journal of Theoretical Probability, **20** (2007), 1059-1071.

Résumé

Ce travail est consacré à certaines classes de systèmes dynamiques ergodiques, munis d'une mesure invariante infinie, telles que des applications de l'intervalle avec un point fixe neutre ou des marches aléatoires. Le comportement asymptotique des sommes de Birkhoff d'observables d'intégrale non nulle est assez bien connu, pour peu que le système ait une certaine forme d'hyperbolicité. Une situation particulièrement intéressante est celle des tours au-dessus d'une application Gibbs-Markov. Nous cherchons dans ce contexte à étudier le cas d'observables d'intégrale nulle. Nous obtenons ainsi une forme de théorème central limite pour des systèmes dynamiques munis d'une mesure infinie.

Après avoir introduit l'ensemble des notions nécessaires, nous adaptons des résultats de E. Csáki et A. Földes sur les marches aléatoires au cas des applications Gibbs-Markov. Les théorèmes d'indépendance asymptotique qui en découlent forment le cœur de cette thèse, et permettent de démontrer un théorème central limite généralisé. Quelques variations sur l'énoncé de ce théorème sont obtenues.

Ensuite, nous abordons les processus en temps continu, tels que des semi-flots et des flots. Un premier travail consiste à étudier les propriétés en temps grand du temps de premier retour et du temps local pour des $\mathbb{Z}^d$-extensions de systèmes dynamiques, ce qui se fait par des méthodes spectrales. Enfin, par réductions successives, nous pouvons obtenir une version du théorème central limite pour des flots $\mathbb{Z}^d$-périodiques, et en particulier le flot géodésique sur le fibré tangent unitaire de certaines variétés $\mathbb{Z}^d$-périodiques hyperboliques.

Abstract

This work is focused on some classes of ergodic dynamical systems endowed with an infinite invariant measure, such as transformations of the interval with a neutral fixed point or random walks. The asymptotic behavior of the Birkhoff sums of observables with a non-zero integral is well known, as long as the system shows some kind of hyperbolicity. The towers over a Gibbs-Markov map are especially interesting. In this context, we aim to study the case of observables whose integral is zero. We get the equivalent of a central limit theorem for some dynamical systems endowed with an infinite measure.

After we introduce the necessary definitions, we adapt some results by E. Csáki and A. Földes on random walks to the case of Gibbs-Markov maps. We derive a theorem on the asymptotic independence of Birhoff sums, which is the core of this thesis, and from this point we work out a generalised central limit theorem. We also prove a few variations on this generalised central limit theorem.

Then, we study dynamical systems in continuous time, such as semi-flows and flows. We first work on the asymptotic properties of the first return time and the local time for $\mathbb{Z}^d$-extensions of dynamical systems; this is done by spectral methods. Finally, step by step, we extend our generalised central limit theorem to cover some $\mathbb{Z}^d$-periodic flows, and in particular the geodesic flow on the unitary tangent bundle of some hyperbolic $\mathbb{Z}^d$-periodic manifolds.