Nº D'ORDRE : 7526

UNIVERSITÉ PARIS XI
UFR SCIENTIFIQUE D'ORSAY

THÈSE

présentée pour obtenir le

GRADE de DOCTEUR EN SCIENCES DE L'UNIVERSITÉ PARIS XI ORSAY

par

Sébastien GOUËZEL

Sujet :
Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes

Soutenue le 1er juin 2004 devant la Commission d'examen

Mme Viviane BALADI, directrice
M. Jean-Pierre CONZE, rapporteur
M. Giovanni FORNI
M. François LEDRAPPIER
M. Yves LE JAN
Mme Lai-Sang YOUNG

d'après le rapport de M. Gerhard KELLER

Remerciements

Je voudrais tout d'abord remercier ma directrice de thèse Viviane Baladi, pour ses encouragements constants et ses judicieux conseils. Elle a su me faire découvrir de nombreux domaines passionnants des mathématiques et me guider vers des problèmes à la fois intéressants et abordables. Sans elle, cette thèse n'aurait pas pu s'accomplir dans de si bonnes conditions.

Merci aussi aux rapporteurs, Jean-Pierre Conze et Gerhard Keller, qui ont accepté de se lancer dans un travail certainement ingrat de relecture attentive et critique. Je remercie également les membres du jury, Lai-Sang Young, Giovanni Forni, François Ledrappier et Yves Le Jan, qui m'ont fait l'honneur d'être présents lors de la soutenance.

Mon parcours mathématique doit assurément beaucoup à Yves Duval, qui a su me faire partager son goût pour la topologie et sans qui je serais peut-être en train de faire de la géométrie algébrique, ainsi qu'à Patrice Le Calvez qui m'a fait découvrir les systèmes dynamiques.

Je remercie également ma famille et mes amis, dont la présence au jour le jour m'a permis durant ces années de thèse de travailler dans un cadre idéal.

Ma reconnaissance va aussi évidemment à Agathe, qui a rédigé ces remerciements à ma place, m'enlevant là une douloureuse épine du pied. Les nombreuses personnes qu'elle a malencontreusement oublié de mentionner ici sont invitées à lui adresser directement leurs remontrances.

Enfin, l'expérience montre que je ne dois surtout pas oublier de souligner les excellentes ambiances dans les différentes équipes qui m'ont accueilli durant ma thèse, que ce soit à l'ENS ou à Orsay.

Version révisée du 11 avril 2008.

Résumé

Dans ce travail, nous nous intéressons aux propriétés de mélange des applications non uniformément dilatantes, par exemple les applications en dimension 1 ayant un point fixe neutre. Après un premier chapitre introductif rappelant des résultats bien connus sur les applications Gibbs-Markov, nous établissons au chapitre deux des résultats très précis sur la vitesse de décorrélation des applications localement Gibbs-Markov (bornes inférieures et supérieures), en appliquant des techniques d'algèbres de Banach aux opérateurs de transfert de premier retour. Dans le chapitre trois, nous montrons en utilisant des techniques de temps hyperboliques qu'une large classe d'applications admet un modèle localement Gibbs-Markov, ce qui permet de leur appliquer les résultats du deuxième chapitre.

Les deux derniers chapitres sont consacrés à des théorèmes de type probabiliste pour les sommes de Birkhoff de fonctions hölderiennes, dans le cadre des applications localement Gibbs-Markov. En étendant à un cadre perturbatif les techniques spectrales du chapitre deux, nous montrons d'une part la convergence des sommes de Birkhoff vers des lois normales ou des lois stables (suivant le système considéré), et d'autre part des théorèmes limites fins comme le théorème de Berry-Esseen (avec vitesse optimale) ou le théorème de la limite locale, sous des hypothèses légèrement plus fortes.

Abstract

In this thesis, we study the mixing properties of non uniformly expanding maps, for example maps in dimension 1 with a neutral fixed point. After an introductory chapter recalling well known facts on Gibbs-Markov maps, we establish in chapter 2 very precise results on the speed of decay of correlations of locally Gibbs-Markov maps (lower and upper bounds), by applying Banach algebra techniques to the first return transfer operators. In chapter 3, we show – using hyperbolic times techniques – that a large class of maps admit locally Gibbs-Markov models, which makes it possible to apply the results of the previous chapter to these maps.

The last two chapters are devoted to probabilistic limit theorems for the Birkhoff sums of Hölder functions, in the setting of locally Gibbs-Markov maps. We extend to a perturbative setting the spectral results of chapter 2, and prove on the one hand the convergence of the Birkhoff sums to normal laws or stable laws (depending on the system we are considering), and on the other hand fine limit theorems, the Berry-Esseen theorem (with optimal speed) and the local limit theorem, under slightly stronger assumptions.

Table des matières

Introduction **7**

1 Applications markoviennes **15**
1.1 Définition des applications markoviennes 16
1.2 Cobords pour les applications Gibbs-Markov 32
1.3 Applications induites et constructions de tours 40
1.4 Exemples en dimension 1 51
1.5 Un produit gauche en dimension 2 58

2 Décorrélation des tours localement Gibbs-Markov **77**
2.1 Opérateurs de transfert de premier retour 78
2.2 Algèbres de Banach et résultats spectraux 82
2.3 Vitesse de décorrélation 96
2.4 Raffinements sur la vitesse de décorrélation 104

3 Construction de structures markoviennes **113**
3.1 Résultats 114
3.2 Temps hyperboliques 119
3.3 Description de la partition 125
3.4 Tour de Young 136

4 Applications markoviennes et théorèmes limites **141**
4.1 Théorèmes limites des applications Gibbs-Markov 142
4.2 Tours localement Gibbs-Markov : approche élémentaire 153
4.3 Un résultat spectral perturbatif 168

4.4 Tours localement Gibbs-Markov : méthode spectrale 177

5 Théorèmes de la limite locale et de Berry-Esseen **189**

5.1 Un résultat spectral perturbatif quantitatif 190
5.2 L'estimée clé 197
5.3 Le théorème de la limite locale 203
5.4 Le théorème de Berry-Esseen 208
5.5 Applications 211

Dans cette thèse, tous les espaces métriques mesurables sont standards, c'est-à-dire polonais (autrement dit complets séparables) et munis de leur tribu borélienne. Cela permet d'utiliser le théorème de Lusin ([Aar97, théorème 1.0.0]), qui assure qu'une fonction mesurable entre un espace polonais muni d'une mesure de probabilité et un autre espace polonais est continue sur des compacts de mesure arbitrairement grande.

Quand on dira qu'il existe une unique mesure μ vérifiant certaines propriétés, l'unicité sera toujours à normalisation près, c'est-à-dire qu'une mesure μ' vérifiant les mêmes propriétés sera nécessairement multiple de μ. Souvent, quand les mesures seront de masse finie, on les supposera normalisées de masse 1.

On notera $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$ et $\overline{\mathbb{D}} = \{z \in \mathbb{C} \mid |z| \leqslant 1\}$. Pour $t \in \mathbb{R}$, la partie entière de t sera notée $\lfloor t \rfloor$.

Introduction

Décorrélations et théorèmes limites

L'objet d'étude principal en systèmes dynamiques discrets est le comportement asymptotique des itérés d'une application T d'un compact X dans lui-même. Lorsque T préserve une mesure μ, on s'intéresse en particulier à la suite $T^n(x)$ pour presque tout point x. Par exemple, si $f : X \to \mathbb{R}$ est continue, est-il possible de décrire le comportement asymptotique des sommes $S_n f(x) = \sum_{k=0}^{n-1} f(T^k(x))$?

Supposons que $\mu(X) = 1$. Alors les fonctions $Z_k = f \circ T^k$ sont des variables aléatoires sur l'espace probabilisé (X, μ), de même loi puisque μ est invariante. Ainsi, on peut espérer au moins formellement démontrer pour les sommes $S_n f(x)$ des résultats analogues aux théorèmes limites probabilistes, comme la loi forte des grands nombres ou le théorème central limite, le problème étant que les Z_k ne sont en général pas indépendants (sauf pour des applications T très particulières).

Le résultat le plus simple dans cette direction, analogue à la loi forte des grands nombres, est le théorème de Birkhoff : si T est ergodique (i.e. $A = T^{-1}(A) \Rightarrow \mu(A) = 0$ ou 1) et $f : X \to \mathbb{R}$ est intégrable, alors la moyenne $\frac{1}{n}\sum_{k=0}^{n-1} f(T^k(x))$ tend presque partout et dans L^1 vers $\int f$.

Pour obtenir des résultats plus fins, il faut faire des hypothèses plus fortes sur T. En particulier, on est souvent amené à quantifier le "défaut d'indépendance" entre f et $f \circ T^k$. On définit les *corrélations* de f et $f \circ T^k$ par

$$\mathrm{Cor}(f, f \circ T^k) := \int f \cdot f \circ T^k - \left(\int f\right) \cdot \left(\int f \circ T^k\right).$$

Lorsque f et $f \circ T^k$ sont indépendantes, ces corrélations sont nulles. Un outil technique essentiel pour étudier les corrélations est *l'opérateur de transfert* $\widehat{T}$ (aussi appelé opérateur de Ruelle-Perron-Frobenius), défini par dualité sur L^2 par

$$\int f \cdot g \circ T = \int \widehat{T} f \cdot g.$$

Si f est bornée et d'intégrale nulle, alors $\mathrm{Cor}(f, f \circ T^k) = \int f \cdot f \circ T^k = \int \widehat{T}^k f \cdot f$, donc $|\mathrm{Cor}(f, f \circ T^k)| \leqslant \left\|\widehat{T}^k f\right\|_1 \|f\|_\infty$. On peut alors énoncer le théorème suivant ([Liv96]), qui

affirme essentiellement que, si les $f \circ T^k$ sont suffisamment indépendants (plus précisément si $\sum |\operatorname{Cor}(f, f \circ T^k)| < \infty$), alors les sommes de Birkhoff $S_n f$ satisfont un théorème central limite :

Théorème (Gordin–Liverani)**.** *Soit $T : X \to X$ une application ergodique préservant une mesure de probabilité μ. Soit $f : X \to \mathbb{R}$ mesurable bornée et d'intégrale nulle telle que $\sum_{k=0}^{\infty} \left\| \widehat{T}^k f \right\|_1 < \infty$. Alors $\frac{1}{\sqrt{n}} S_n f$ tend en loi vers une gaussienne $\mathcal{N}(0, \sigma^2)$, où $\sigma^2 \geqslant 0$ est donné par*

$$\sigma^2 = \int f^2 + 2 \sum_{k=1}^{\infty} \int f \cdot f \circ T^k.$$

De plus, $\sigma^2 = 0$ si et seulement si il existe une fonction mesurable $\chi : X \to \mathbb{R}$ telle que $f = \chi - \chi \circ T$ presque partout.

La preuve de ce théorème est essentiellement probabiliste : les hypothèses permettent de se ramener à une martingale inverse dans L^2, pour laquelle le théorème central limite est bien connu (méthode de Gordin).

Il est donc important pour obtenir des théorèmes limites d'estimer la vitesse de décroissance des corrélations. Lorsque le spectre de $\widehat{T}$, agissant sur un certain espace de fonctions, est constitué d'une valeur propre simple en 1 et d'une partie d'un disque $B(0, 1-\delta)$ avec $\delta > 0$ (on dit que $\widehat{T}$ a un *trou spectral*), alors $\widehat{T}^k f$ tend exponentiellement vite vers la fonction constante égale à $\int f$, ce qui implique que $\operatorname{Cor}(f, f \circ T^k)$ décroît exponentiellement vite. Ainsi, le théorème de Gordin–Liverani s'applique, et on obtient un théorème central limite dans ce cadre.

Cette méthode est extrêmement féconde dans les systèmes qui ont une dynamique exponentiellement mélangeante, et en particulier pour les applications uniformément dilatantes, c'est-à-dire les applications lisses (mettons C^2) sur une variété riemannienne compacte M telles qu'il existe des constantes $C > 0$ et $\lambda > 1$ telles que, pour tous $x \in M$ et $v \in T_x M$, pour tout $n \in \mathbb{N}$,

$$\|DT^n(x).v\| \geqslant C\lambda^n \|v\|.$$

En effet, on vérifie alors que $\widehat{T}$ agissant sur l'espace de Banach des fonctions de classe C^1 admet un trou spectral.

Plus généralement, cette méthode s'applique aux applications uniformément hyperboliques, où l'espace tangent se décompose en chaque point en somme directe de deux sous-espaces supplémentaires, l'un correspondant à des vecteurs qui sont contractés par la différentielle de T et l'autre à des vecteurs dilatés, toutes les constantes étant uniformes sur la variété. En effet, on peut alors se ramener par un codage symbolique (une *partition de Markov* finie) à un modèle abstrait possédant des propriétés analogues aux applications uniformément dilatantes décrites ci-dessus ; en particulier, l'opérateur de transfert agissant sur un espace fonctionnel bien choisi a un trou spectral.

Applications non-uniformément dilatantes

L'objet de cette thèse est l'étude des applications *non-uniformément dilatantes* : pour presque tout point x, il existe des constantes $\lambda(x) > 1$ et $C(x) > 0$ telles que, pour tout $n \in \mathbb{N}$, pour tout $v \in T_xM$, $\|DT^n(x).v\| \geqslant C(x)\lambda(x)^n \|v\|$. Comme la dilatation est non-uniforme (i.e. les constantes dépendent de x), l'opérateur de transfert n'a en général plus de trou spectral. En particulier, les corrélations peuvent ne pas décroître exponentiellement vite.

Plus précisément, on étudie dans cette thèse des modèles abstraits d'applications, appelés *tours localement Gibbs-Markov* (définition 1.3.12). Ces tours sont analogues aux modèles des applications uniformément hyperboliques construits par l'intermédiaire d'une partition de Markov, la différence principale étant que, dans le cas non-uniforme, les partitions de Markov sont dénombrables et non plus finies. Comme on le montre dans le chapitre 3, les applications non-uniformément dilatantes peuvent, sous des hypothèses très générales, se ramener à des tours localement Gibbs-Markov (le problème est de construire une partition de Markov adéquate, et en particulier de contrôler la mesure des points dont le temps de retour est grand, voir le théorème 3.3.1). Cependant, pour la simplicité de cette introduction, on va se contenter dans la suite de décrire sur un exemple les résultats qu'on obtient dans des tours localement Gibbs-Markov générales.

Soit $\alpha \in]0,1[$, on définit une application T sur le segment $[0,1]$ par

$$T(x) = \begin{cases} x(1+2^\alpha x^\alpha) & \text{si } 0 \leqslant x \leqslant 1/2, \\ 2x-1 & \text{si } 1/2 < x \leqslant 1. \end{cases} \qquad (*)$$

Cette application, introduite dans [LSV99], présente un point fixe en 0 avec $T'(0) = 1$, ce qui introduit de la non-uniformité dans la dilatation. En particulier, plus α est grand, et plus les points mettent de temps à s'éloigner du point fixe, ce qui rend la dynamique de plus en plus pathologique. Il existe quand même une unique mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue, mais sa densité tend vers $+\infty$ en 0 (la condition $\alpha < 1$ sert à assurer que cette mesure invariante soit de masse finie).

Vitesse de décorrélation

Dans [You99], Young a montré, en utilisant des méthodes de couplage probabiliste, que les corrélations des fonctions hölderiennes pour l'application $(*)$ décroissaient au moins aussi vite que $\frac{1}{n^{1/\alpha-1}}$ (ce qui améliorait les résultats précédents de [LSV99]). En particulier, quand $\alpha < 1/2$, cette estimée est sommable, si bien que les sommes $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f \circ T^k(x)$ vérifient un théorème central limite d'après le théorème de Gordin–Liverani. En utilisant une méthode spectrale, Sarig a montré dans [Sar02] que cette estimée sur la vitesse de décorrélation est optimale lorsque $\alpha < 1/2$, pour une large classe de fonctions : si f est hölderienne et nulle sur $[0,1/2]$, alors $\mathrm{Cor}(f, f \circ T^n) = \frac{c\int f}{n^{1/\alpha-1}} + O(1/n^{\lfloor 1/\alpha \rfloor})$ pour une certaine constante c non nulle.

Dans le chapitre 2, je généralise et je renforce ces résultats, tant du point de vue des bornes inférieures que des bornes supérieures, pour des tours localement Gibbs-Markov générales.

Pour l'application (∗), on obtient par exemple le résultat suivant dans la proposition 2.3.9 et le théorème 2.4.14 :

Théorème. *Soient $0 < \alpha < 1$, T l'application (∗) et $f : [0,1] \to \mathbb{R}$ hölderienne.*

1. *On a toujours* $\mathrm{Cor}(f, f \circ T^n) = O\left(\frac{1}{n^{1/\alpha - 1}}\right)$.
2. *Il existe $c > 0$ tel que, si f est nulle au voisinage de 0, alors* $\mathrm{Cor}(f, f \circ T^n) = \frac{c \int f}{n^{1/\alpha - 1}} + O(E_\alpha(n))$, *où* $E_\alpha(n) = \frac{1}{n^{1/\alpha}}$ *si* $\alpha < 1/2$, $\frac{\log n}{n^2}$ *si* $\alpha = 1/2$, *et* $\frac{1}{n^{2/\alpha - 2}}$ *si* $\alpha > 1/2$.
3. *Si $f(0) = 0$ et $\int f = 0$, supposons aussi que $|f(x)| \leqslant Cx^\gamma$ pour un certain $\gamma > 0$. Alors* $\mathrm{Cor}(f, f \circ T^n) = O\left(\frac{1}{n^{\min(1/\alpha, (1+\gamma)/\alpha - 1)}}\right)$.

La première assertion avait déjà été démontrée par Young. Elle implique en particulier que, lorsque $\alpha < 1/2$, toute fonction hölderienne vérifie un théorème central limite.

La deuxième assertion montre que, pour une large classe de fonctions, la vitesse de décorrélation est exactement $\frac{1}{n^{1/\alpha - 1}}$, ce qui étend les résultats de Sarig au cas $1/2 \leqslant \alpha < 1$.

La troisième assertion montre de plus qu'il y a un gain dans la vitesse de décorrélation pour les fonctions qui s'annulent au point fixe neutre (ce n'est pas très surprenant : ce point fixe est la source de non-uniformité, et son influence diminue s'il n'est pas "vu" par la fonction f). En particulier, si f est lipschitzienne, on obtient que les corrélations $\mathrm{Cor}(f, f \circ T^n)$ décroissent en $\frac{1}{n^{1/\alpha}}$: il y a un gain de $1/n$ par rapport à une fonction générique. Cette estimée est toujours sommable, si bien que f vérifie encore un théorème central limite même dans le cas $\alpha \geqslant 1/2$.

Ces résultats sont encore valables pour d'autres applications, et d'autres vitesses de décroissance. Par exemple, en remplaçant T par une application du même type mais qui a un point fixe encore plus neutre en 0, on obtient qu'il existe des applications pour lesquelles les corrélations décroissent arbitrairement lentement. À l'opposé, on peut obtenir une vitesse de décorrélation exponentielle gauche (en e^{-cn^η} avec $0 < \eta < 1$) ou exponentielle. Par exemple, dans le chapitre 3, ces résultats sont appliqués entre autres à l'application de Viana (3.7), et impliquent que ses corrélations décroissent en $e^{-c\sqrt{n}}$.

La méthode de preuve est essentiellement spectrale, en considérant les *opérateurs de transfert de premier retour* R_n introduits par Sarig. Il s'agit alors de comprendre précisément la série $\sum z^n R_n$, pour $|z| \leqslant 1$, ainsi que son inverse. Pour cela, j'utilise des techniques d'algèbre de Banach (et plus précisément des généralisations du lemme de Wiener) dans certaines algèbres qui contiennent des informations sur la vitesse de décroissance des coefficients (alors que Sarig utilisait des méthodes plus élémentaires sur la régularité de certaines fonctions, qui ne s'appliquent plus dans le cas $1/2 \leqslant \alpha < 1$).

Théorèmes limites

Soit $T : X \to X$ une application préservant une mesure de masse 1 et $f : X \to \mathbb{R}$ une fonction continue. Pour étudier la convergence en loi de $\frac{1}{B_n}\sum_{k=0}^{n-1} f \circ T^k$ (pour une normalisation $B_n \to +\infty$ bien choisie), on peut utiliser la méthode de Gordin lorsque les corrélations $\mathrm{Cor}(f, f \circ T^n)$ sont sommables, pour obtenir un théorème central limite (correspondant à la normalisation $B_n = \sqrt{n}$), en se ramenant à une martingale. Cependant, cela ne traite pas tous les cas (par exemple l'application (∗) pour $1/2 < \alpha < 1$ et $f(0) \neq 0$, puisque $\mathrm{Cor}(f, f \circ T^n)$ n'est alors pas sommable, d'après ce qui précède), et cela ne permet pas d'accéder à des raffinements du théorème central limite qui ne sont pas vérifiés par des martingales.

Une méthode permettant d'obtenir des résultats plus fins (mais s'appliquant dans moins de situations) est la méthode spectrale de Nagaev et Guivarc'h ([GH88]), qui consiste à considérer des perturbations $\widehat{T}_t$ de l'opérateur de transfert $\widehat{T}$, pour obtenir une description précise de $\int e^{it\sum_{k=0}^{n-1} f\circ T^k}$. Cependant, cette méthode ne peut fonctionner que lorsque $\widehat{T}$ a un trou spectral, et ne s'applique donc pas dans les cas non uniformes (par exemple les tours localement Gibbs-Markov).

Dans le chapitre 4, on démontre des théorèmes limites pour les tours localement Gibbs-Markov en appliquant cette idée perturbative non pas à l'opérateur de transfert, mais aux opérateurs de transfert de premier retour. Moyennant un résultat de perturbation dans certaines algèbres de Banach, obtenu par des techniques d'analyse de Fourier, cela permet de démontrer non seulement des théorèmes de la limite centrale dans certains cas, mais également la convergence vers des lois stables (des lois apparaissant en théorie des probabilités comme limites de moyennes de variables aléatoires indépendantes identiquement distribuées) dans d'autres cas. Pour les applications (∗), on obtient par exemple dans le théorème 4.2.5 le résultat suivant (qui avait été conjecturé dans les années 80 par les physiciens qui avaient introduit des applications analogues à (∗) comme modèles pour des flots turbulents) :

Théorème. *Soient $0 < \alpha < 1$, T l'application (∗) et $f : [0,1] \to \mathbb{R}$ une fonction lipschitzienne d'intégrale nulle.*
- *Si $\alpha < 1/2$, alors il existe σ^2 tel que $\frac{1}{\sqrt{n}}\sum_{k=0}^{n-1} f \circ T^k$ converge en loi vers $\mathcal{N}(0,\sigma^2)$, i.e. cette suite satisfait un théorème central limite.*
- *Si $\alpha > 1/2$, il y a deux cas :*
  1. *Si $f(0) = 0$, alors f satisfait un théorème central limite, comme ci-dessus.*
  2. *Si $f(0) \neq 0$, alors $\frac{1}{n^\alpha}\sum_{k=0}^{n-1} f \circ T^k$ converge en loi vers une loi stable explicite.*

En utilisant le même type d'idées de perturbation des opérateurs de transfert de premier retour, et en poussant la précision des estimées techniques un cran plus loin, on peut aussi obtenir des raffinements du théorème central limite qui ne sont pas vérifiés par les martingales, et ne sont donc pas accessibles par la méthode de Gordin.

Le théorème de la limite locale décrit le comportement de $\mu\{x \in X \mid \sum_{k=0}^{n-1} f \circ T^k(x) \in [a,b]\}$ pour $[a,b]$ intervalle borné de $\mathbb{R}$ (alors que le théorème central limite décrit le comportement

de $\mu\{x \in X \mid \sum_{k=0}^{n-1} f \circ T^k(x) \in \sqrt{n}[a,b]\})$, et affirme que cette mesure se comporte comme $\frac{1}{\sqrt{n}} \frac{b-a}{\sigma\sqrt{2\pi}}$ (sous certaines hypothèses d'apériodicité : par exemple, si f est à valeurs entières et $[a,b] = [1/3, 2/3]$, le résultat ne peut être valable tel quel). Dans le chapitre 5, je démontre le théorème de la limite locale dans de nombreuses tours localement Gibbs-Markov. Cela s'applique par exemple à l'application (∗) lorsque $\alpha < 1/2$: le résultat suivant se déduit du théorème 5.5.6.

Théorème. *Soient $0 < \alpha < 1/2$, T l'application (∗) et $f : X \to \mathbb{R}$ hölderienne, qui ne s'écrit pas presque partout sous la forme $f = g - g \circ T + \rho + q$ où ρ est une constante, et q est à valeurs dans un sous-groupe discret de $\mathbb{R}$. Alors les sommes $\sum_{k=0}^{n-1} f \circ T^k$ vérifient un théorème de la limite locale.*

Notons qu'en pratique, il est difficile de vérifier que f ne s'écrit pas sous la forme $f = g - g \circ T + \rho + q$, car cette égalité n'aurait lieu que presque partout. Cependant, le théorème 1.3.16 assure alors que g est nécessairement régulière, et on peut parfois en déduire que l'égalité a lieu partout. Les informations sur les points périodiques du système peuvent alors permettre de conclure. Par exemple, le corollaire 1.4.4 assure que, pour T donné par (∗), la fonction $f = \log |T'|$ ne s'écrit pas presque partout sous la forme $g - g \circ T + \rho + q$.

Le même type de technique (et en fait les mêmes estimées) permet de démontrer le théorème de Berry-Esseen dans ce contexte : lorsque $\alpha < 1/3$, la vitesse de convergence dans le théorème central limite est en $O(1/\sqrt{n})$. Des résultats analogues (avec une vitesse un peu moins précise, en $O(1/n^{1/2-\varepsilon})$ pour tout $\varepsilon > 0$) avaient déjà été démontrés, par des techniques différentes et plus directes. En revanche, ces techniques ne permettaient pas d'arriver au théorème de la limite locale (puisque, même dans le cas probabiliste, la preuve de ce théorème requiert des techniques de Fourier élaborées).

Plan de la thèse

Cette thèse n'est pas directement une compilation d'articles, car il m'a semblé plus judicieux de tout formuler dans le cadre unifié des tours localement Gibbs-Markov. Cependant, les chapitres autres que le premier correspondent chacun plus ou moins à un article.

Dans le premier chapitre, on commence par rappeler des résultats classiques sur les applications Gibbs-Markov. Ces applications sont des analogues des décalages de type fini, mais sur un ensemble de symboles dénombrable. Les opérateurs de transfert correspondants ont un trou spectral sur un certain espace de fonctions hölderiennes, ce qui permet de décrire très précisément les propriétés ergodiques de ces applications. Ces résultats font plus ou moins partie du folklore, et se trouvent en partie dans [Aar97]. On définit ensuite les tours localement Gibbs-Markov, qui sont des analogues des tours de Young mais avec une combinatoire éventuellement plus complexe, avant de donner plusieurs exemples. On décrit en particulier un produit gauche en dimension 2, étudié dans [Gou07]. Dans ce chapitre, on démontre

également un résultat de régularité automatique des fonctions cobords (théorème 1.2.1), qui s'applique aux transformations non uniformément dilatantes.

Dans le chapitre 2, on expose les résultats de [Gou04b] (ainsi que quelques extensions) sur la vitesse de décorrélation des tours localement Gibbs-Markov. On obtient en particulier des bornes inférieures (théorème 2.3.1) et supérieures (théorème 2.4.13) très précises.

Dans le chapitre 3, correspondant à [Gou06], on montre que de nombreuses applications non-uniformément dilatantes admettent un modèle localement Gibbs-Markov, dont on peut contrôler explicitement les caractéristiques (théorème 3.3.1). En particulier, on peut alors appliquer les résultats du chapitre 2, et obtenir une majoration de la vitesse de décorrélation, même pour des décorrélations rapides (par exemple exponentielles gauches ou exponentielles).

Dans le chapitre 4, correspondant à [Gou04a], on commence par décrire les perturbations spectrales des opérateurs de transfert pour les applications Gibbs-Markov, et les théorèmes limites qui s'en déduisent (convergence vers des gaussiennes dans le théorème 4.1.4 ou des lois stables dans le théorème 4.1.7). On en déduit ensuite des théorèmes limites du même type pour des tours localement Gibbs-Markov générales, de deux manières différentes : on peut soit utiliser un argument élémentaire (partie 4.2, théorème 4.2.4), soit utiliser un argument spectral (partie 4.4, théorème 4.4.1), plus susceptible de généralisations.

Enfin, dans le chapitre 5, correspondant à [Gou05], on montre comment les arguments spectraux précédents permettent de démontrer le théorème de la limite locale et le théorème de Berry-Esseen dans les tours localement Gibbs-Markov, avant de l'appliquer aux transformations non-uniformément dilatantes (théorèmes 5.5.6 et 5.5.7).

Chapitre 1

Applications markoviennes

Sommaire

1.1 Définition des applications markoviennes 16
1.1.1 Définitions 17
1.1.2 Applications Gibbs-Markov 18
1.1.3 Propriétés ergodiques élémentaires 19
1.1.4 Rappels de théorie spectrale 21
1.1.5 Propriétés spectrales des applications Gibbs-Markov 24
1.2 Cobords pour les applications Gibbs-Markov 32
1.2.1 Preuve du théorème 1.2.1' 33
1.2.2 Le cas intégrable 39
1.3 Applications induites et constructions de tours 40
1.3.1 Induction 40
1.3.2 Le processus inverse : construction de tours 41
1.3.3 Applications induisant des applications Gibbs-Markov 43
1.3.4 Modélisation d'applications réelles 48
1.4 Exemples en dimension 1 51
1.4.1 L'application de Liverani-Saussol-Vaienti 51
1.4.2 Points fixes très neutres 54
1.5 Un produit gauche en dimension 2 58
1.5.1 Estimées préliminaires 59
1.5.2 La partition de Markov 61
1.5.3 Bornes sur la distorsion 63
1.5.4 Construction de la mesure invariante 65
1.5.5 Décorrélations multiples de l'application F, et bornitude L^p 67
1.5.6 Comportement asymptotique de X_n 70

L'objet de cette thèse est l'étude de propriétés fines des applications markoviennes. Dans ce chapitre, on s'intéresse à des propriétés ergodiques générales de plusieurs classes d'applications markoviennes, avant de donner plusieurs exemples.

Dans la première partie, on décrit le comportement des applications Gibbs-Markov (i.e. markoviennes et uniformément dilatantes en un certain sens). Les résultats de cette partie font plus ou moins partie du folklore (certains résultats sont démontrés par exemple dans [Aar97]), mais il est intéressant de les rappeler ici, d'une part parce qu'ils seront systématiquement utilisés dans le reste de cette thèse, et d'autre part parce qu'ils montrent l'utilité de certains résultats spectraux, rappelés dans le paragraphe 1.1.4. Dans la deuxième partie, on montre un résultat de régularité automatique des solutions de l'équation cohomologique pour les applications Gibbs-Markov. Dans la troisième partie, on s'intéresse aux applications qui induisent une application Gibbs-Markov sur une partie de l'espace, et on montre comment démontrer à leur sujet certains résultats à partir des résultats correspondants pour les applications Gibbs-Markov. Dans la quatrième partie, on décrit des exemples de telles applications en dimension 1, ayant un point fixe neutre. Enfin, dans la cinquième partie, qui reprend une partie de l'article [Gou07], on développe en détail un exemple analogue, mais en dimension 2, possédant toute une courbe neutre.

1.1 Définition des applications markoviennes

Une application markovienne est une transformation d'un espace qui préserve une partition finie ou dénombrable, i.e. l'image d'un élément de partition est une union d'éléments de la partition. Cette notion permet (sous certaines hypothèses de dilatation) de coder la trajectoire d'un point en regardant dans quels éléments de la partition ses itérés successifs se trouvent, et donc de se ramener à un système symbolique, souvent plus simple à étudier que le système initial.

Une légère variante de cette notion (séparant le cas des directions stables et instables) s'est révélée extrêmement féconde dans l'étude des systèmes uniformément hyperboliques, dans les années 1970 : un tel système admet toujours une partition de Markov finie, si bien qu'on peut le coder par un décalage de type fini. Cela permet de comprendre plus ou moins complètement les propriétés ergodiques de ces applications (existence et unicité des mesures d'équilibre pour des potentiels Hölder, existence de mesures SRB...)

Dans le cas non uniformément dilatant, ou plus généralement non uniformément hyperbolique, il n'existe en général pas de partition de Markov finie, et il faut donc utiliser des partitions dénombrables. Comme les décalages correspondants sont beaucoup moins bien compris que les décalages de type fini, les conclusions qu'on peut en tirer sont plus faibles, et il faut en général des hypothèses supplémentaires sur la forme de la partition de Markov pour obtenir des informations intéressantes.

1.1.1 Définitions

Définition 1.1.1. *Soient $(X, d, \mathcal{B})$ un espace métrique polonais borné muni de sa tribu borélienne, et m une mesure de probabilité sur X. Une application non-singulière $T : X \to X$ est dite markovienne s'il existe une partition α de X (modulo 0) par des ensembles de mesure non nulle telle que :*
- *pour tout $a \in \alpha$, $T(a)$ est une réunion (modulo 0) d'éléments de α et $T : a \to T(a)$ est un isomorphisme mesurable.*
- *le complété pour m de la tribu engendrée par $\bigvee_0^\infty T^{-n}(\alpha)$ est $\mathcal{B}$.*

On notera cette application markovienne par $(X, d, \mathcal{B}, m, T, \alpha)$, ou par un sous-ensemble de ce 6-uplet, en fonction du contexte.

On peut coder une application markovienne, en associant à un point x la suite des éléments $a_n(x)$ de α tels que $T^n(x) \in a_n(x)$ (qui est bien définie de façon unique pour presque tout x). L'application $\varphi : X \to \alpha^{\mathbb{N}}$ ainsi définie est alors un isomorphisme mesurable entre (X, m) et $(\alpha^{\mathbb{N}}, \varphi_* m)$ qui conjugue T et l'application décalage à gauche sur une partie de $\alpha^{\mathbb{N}}$. Dans la suite, on n'utilisera pas ce codage (car on voudra par exemple garder des informations sur la distance d), mais il est utile de le garder à l'esprit, par exemple pour comprendre les définitions suivantes. Parfois, on utilisera des applications markoviennes sur des espaces probabilisés abstraits, non munis d'une distance, et ce codage permettra alors de se ramener au cadre de la définition 1.1.1, avec un espace polonais (puisque $\alpha^{\mathbb{N}}$ est complet séparable pour la distance $d((u_n), (v_n)) = \sum \tau^n 1(u_n \neq v_n)$, lorsque $\tau < 1$).

Définition 1.1.2. *On dit qu'une application markovienne $(X, d, \mathcal{B}, m, T, \alpha)$ est* Markov-transitive *si, pour tous $a, b \in \alpha, \exists n \in \mathbb{N}$, $a \subset T^n(b) \mod 0$.*

Définition 1.1.3. *On dit qu'une application markovienne $(X, d, \mathcal{B}, m, T, \alpha)$ est* Markov-mélangeante *si, pour tous $a, b \in \alpha, \exists N \in \mathbb{N}, \forall n \geqslant N$, $a \subset T^n(b) \mod 0$.*

Pour $a_0, \dots, a_{n-1} \in \alpha$, on notera

$$[a_0, \dots, a_{n-1}] = \bigcap_0^{n-1} T^{-i}(a_i).$$

On dira que les ensembles $[a_0, \dots, a_{n-1}]$ de mesure non nulle sont des *cylindres* de l'application T. Ainsi, presque tout point de X est dans un unique cylindre de longueur n. Notons que, une fois qu'on a codé, on n'a plus besoin de ce "presque tout", mais qu'en général il faut faire attention aux points qui sont dans plusieurs éléments de la partition de Markov (les "bords" de la partition de Markov). Soit $\mathcal{F}_n$ la sous-tribu de X engendrée par les cylindres de longueur n, le théorème de convergence des martingales assure que, pour $f : X \to \mathbb{R}$ intégrable, alors $E(f|\mathcal{F}_n)(x) \to f(x)$ presque partout. En particulier, si K est un ensemble mesurable, presque tous ses points en sont des points de densité : pour presque tout $x \in K$, soit $a_n(x)$ l'élément de α contenant $T^n(x)$, alors $\frac{m(K \cap [a_0(x), \dots, a_{n-1}(x)])}{m[a_0(x), \dots, a_{n-1}(x)]} \to 1$.

Pour $x, y \in X$, on appellera *temps de séparation* de x et y par T l'entier

$$s(x,y) = \inf\{n \in \mathbb{N} \mid \nexists a \in \alpha, T^n(x) \in a, T^n(y) \in a\}.$$

Ainsi, $s(x,y)$ est la longueur du plus grand cylindre contenant x et y. Pour presque tous x et y, $s(x,y)$ est fini. Pour $0 < \tau < 1$, on peut alors définir une distance d_τ par

$$d_\tau(x,y) = \tau^{s(x,y)}. \tag{1.1}$$

(En fait, ce n'est une distance que sur une partie de X de mesure pleine). Toutes les distances d_τ induisent la même topologie, dont une base d'ouverts est donnée par les cylindres.

1.1.2 Applications Gibbs-Markov

Définition 1.1.4. *Une application markovienne $(X, d, \mathcal{B}, m, T, \alpha)$ est* Gibbs-Markov *si*

1. *T a la propriété de grande image :* $\inf_{a\in\alpha} m(Ta) > 0$.
2. *Dilatation : il existe $\lambda > 1$ tel que $\forall a \in \alpha$, pour presque tous $x, y \in a$, $d(Tx, Ty) \geqslant \lambda d(x,y)$.*
3. *Distorsion lipschitzienne : pour $a \in \alpha$, soit g l'inverse du jacobien de T sur a, i.e. $g(x) = \frac{\mathrm{d}m}{\mathrm{d}(m\circ T_{|a})}(x)$, défini et positif presque partout. Alors il existe C tel que, pour tout $a \in \alpha$, pour presque tous $x, y \in a$,*

$$\left|1 - \frac{g(x)}{g(y)}\right| \leqslant C d(Tx, Ty). \tag{1.2}$$

Remarque 1.1.5. Si T est une application markovienne, elle dilate toujours la distance d_τ de τ^{-1}, pour tout $0 < \tau < 1$.

Dans le cas d'une distance générale d, supposons T Gibbs-Markov. Pour presque tous x, y, la dilatation donne $d(T^{s(x,y)}x, T^{s(x,y)}y) \geqslant \lambda^{s(x,y)} d(x,y)$. Comme on a supposé X borné, disons par une constante K, on obtient $d(x,y) \leqslant K\lambda^{-s(x,y)}$, i.e. $d \leqslant K d_{\lambda^{-1}}$. Ainsi, T est également Gibbs-Markov pour la distance $d_{\lambda^{-1}}$.

Remarque 1.1.6. Il y a *a priori* deux interprétations possibles du point 2 de la définition :
– Il existe $U \subset a \times a$ de mesure pleine tel que, pour tous $(x,y) \in U$, $d(Tx, TY) \geqslant \lambda d(x,y)$.
– Il existe $V \subset a$ de mesure pleine tel que, pour tous $x \in V$, $y \in V$, $d(Tx, TY) \geqslant \lambda d(x,y)$.
Pour une propriété générale, la seconde exigence est en général plus forte que la première. Cependant, ici, les deux hypothèses sont équivalentes.

Supposons en effet la première exigence vérifiée. Par le théorème de Lusin, soit K_n une suite croissante avec $m(\bigcup K_n) = 1$ telle que T soit continue sur K_n. Soit L_n l'ensemble des points de densité de K_n, qui vérifie $m(K_n \backslash L_n) = 0$ par le théorème de convergence des martingales. Posons $V = \bigcup L_n$, alors $m(V) = 1$. Soient $x, y \in V$, on vérifie alors aisément

(en utilisant le fait que le diamètre des cylindres tend vers 0, et en approchant (x, y) par $(x', y') \in U \cap (K_n \times K_n)$) que $d(Tx, Ty) \geqslant \lambda d(x, y)$.

De la même façon, on peut interpréter le point 3 de la définition de deux manières, mais elles sont équivalentes (en considérant les points de continuité mesurable de g).

Dans la suite, toutes les propriétés portant sur des couples de points qu'on considérera seront du même type, et on pourra donc négliger cette subtilité.

Un intérêt des applications Gibbs-Markov est qu'elles se comportent presque comme des décalages de type fini, en particulier du point de vue spectral. Ainsi, on peut aisément obtenir des résultats d'existence de mesure invariante ou de mélange. Dans la suite de cette partie, on va démontrer un tel résultat, classique (cf. [Aar97, théorème 4.7.4]), essentiellement pour illustrer les méthodes spectrales qu'on va utiliser plus loin dans cette thèse, et pour mettre en place quelques outils.

1.1.3 Propriétés ergodiques élémentaires

Le lemme suivant permet de contrôler la distorsion des itérés d'une application Gibbs-Markov, et sera essentiel dans la suite.

Lemme 1.1.7 (Lemme de distorsion)**.** *Soit $g^{(n)}(x) = g(x) \cdots g(T^{n-1}x)$ (c'est l'inverse du jacobien de T^n). Il existe une constante $C > 0$ telle que pour presque tous x, y, on a $s(x, y) \geqslant n \Rightarrow \left|1 - \frac{g^{(n)}(x)}{g^{(n)}(y)}\right| \leqslant Cd(T^n x, T^n y)$.*

En particulier, il existe une constante $D > 0$ telle que, si $\underline{a} = [a_0, \ldots, a_{n-1}]$ est un cylindre de longueur n, pour presque tout $x \in \underline{a}$, $D^{-1} g^{(n)}(x) \leqslant \frac{m[\underline{a}]}{m[T^n \underline{a}]} \leqslant D g^{(n)}(x)$.

Démonstration. Soient x, y tels que $s(x, y) \geqslant n$, et dont tous les itérés vérifient (1.2) (cette dernière propriété étant vraie presque partout). Alors

$$\begin{aligned}\left|\log(g^{(n)}(x)) - \log(g^{(n)}(y))\right| &\leqslant \sum_{i=0}^{n-1} \left|\log(g(T^i x)) - \log(g(T^i y))\right| \\ &\leqslant \sum_{i=1}^{n} Cd(T^i x, T^i y) \leqslant \sum_{i=1}^{n} C\lambda^{-(n-i)} d(T^n x, T^n y) \\ &\leqslant \frac{C}{1 - \lambda^{-1}} d(T^n x, T^n y).\end{aligned}$$

En prenant l'exponentielle, on obtient la première partie du lemme.

En particulier, si $\underline{a}$ est un cylindre de longueur n et $x, y \in \underline{a}$, on a $\frac{g^{(n)}(x)}{g^{(n)}(y)} \leqslant D$ pour un certain D. Comme $1/g^{(n)}(y)$ est le jacobien de T^n, on obtient en intégrant sur $y \in \underline{a}$ que $g^{(n)}(x) m[T^n \underline{a}] \leqslant D m[\underline{a}]$. L'inégalité opposée est analogue. □

Ce lemme implique en particulier que les itérés d'applications Gibbs-Markov sont encore Gibbs-Markov.

Théorème 1.1.8. *Soit (X, T, m, α) une application Gibbs-Markov. Il existe alors un nombre fini (éventuellement nul) de parties deux à deux disjointes $X_1, \dots, X_k$ de X, chacune étant une réunion d'éléments de α, avec $T(X_i) = X_i$, et des mesures de probabilité $\mu_1, \dots, \mu_k$ vérifiant les propriétés suivantes :*

La mesure μ_i est invariante, ergodique, et équivalente à la restriction de m à X_i. De plus, toute mesure de probabilité invariante absolument continue par rapport à m est combinaison linéaire des μ_i.

Enfin, pour $1 \leqslant i \leqslant k$, le système dynamique $(X_i, T, m_{|X_i}, \alpha_{|X_i})$ est Gibbs-Markov, Markov-transitif et ergodique.

En particulier, pour étudier les propriétés des mesures invariantes des applications Gibbs-Markov, on peut sans perte de généralité se restreindre au cas transitif. Dans la proposition 1.1.17, on verra que $k \geqslant 1$, i.e., il existe effectivement des mesures de probabilité invariantes absolument continues.

Démonstration. Soit μ une mesure de probabilité invariante par T et absolument continue par rapport à m. Montrons que, si $a \in \alpha$ satisfait $\mu(a) > 0$, alors μ et m sont équivalentes sur a, i.e. la densité $f = \frac{\mathrm{d}\mu}{\mathrm{d}m}$ est presque partout positive sur a.

Soit $A = \{x \in a \mid f(x) > 0\}$: il vérifie $m(A) > 0$. Notons $\mathcal{F}_n$ la tribu engendrée par les cylindres de longueur n. Le théorème de convergence des martingales assure que, pour toute fonction $\varphi \in L^1(m)$, $E(\varphi \mid \mathcal{F}_n)$ tend presque sûrement et dans L^1 vers φ. Ainsi, pour presque tout x, $E(1_A \mid \mathcal{F}_n)(x) \to 1_A(x)$. De plus, le théorème de récurrence de Poincaré assure que μ-presque tout point de a reviendra infiniment souvent dans a. On choisit une fois pour toutes un point x de A satisfaisant ces deux convergences.

Notons $[a_0, \dots, a_n]$ le cylindre de longueur $n+1$ contenant x (avec $a_0 = a$). Comme x revient une infinité de fois dans a, on a $a_n = a$ pour une infinité d'indices n. De plus, le choix de x assure que $\frac{m\left([a_0,\dots,a_n]\cap A\right)}{m\left([a_0,\dots,a_n]\right)} \to 1$. En appliquant T^n et en utilisant la distorsion bornée donnée par le lemme précédent, on obtient que $T^n(A)$ occupe une proportion arbitrairement proche de 1 de a, puis

$$a \subset \bigcup_{n=0}^{\infty} T^n(A) \mod 0. \tag{1.3}$$

Par la formule du changement de variables, l'invariance de μ s'écrit ainsi : pour m-presque tout x,

$$f(x) = \sum_{T(y)=x} g(y) f(y).$$

De plus, $g > 0$ presque partout. Comme T est non-singulière, presque tout point x a donc toutes ses préimages y qui satisfont $g(y) > 0$. Soit x satisfaisant ces deux propriétés. Si $x \in T(A)$, alors il existe y dans A tel que $T(y) = x$. Alors $f(x) \geqslant g(y)f(y) > 0$. Cela montre que $T(A) \subset A \mod 0$. Ainsi, (1.3) assure que $a \subset A \mod 0$, ce qui démontre l'assertion initiale.

Soit maintenant μ une mesure de probabilité invariante ergodique absolument continue par rapport à m. Ce qui précède montre que l'ensemble $A(\mu) = \{\frac{\mathrm{d}\mu}{\mathrm{d}m} > 0\}$ est une réunion d'éléments de α, et est invariant par T. Le théorème de récurrence de Poincaré assure que T est Markov-transitive sur $A(\mu)$. Ainsi, le système $(A(\mu), T_{|A(\mu)}, m_{|A(\mu)}, \alpha_{|A(\mu)})$ est Gibbs-Markov transitif, et ergodique puisque m et μ sont équivalentes sur $A(\mu)$.

Montrons qu'il y a un nombre fini de telles mesures ergodiques. Soient $a_1, \dots, a_N$ un nombre fini d'éléments de α, avec $\sum m(a_j) > 1 - \eta$, où η est tel que l'image de tout élément de α par T soit de mesure au moins η. Ainsi, pour tout $a \in \alpha$, $T(a)$ contient l'un des a_j. Si $\mu_1, \dots, \mu_k$ sont invariantes, ergodiques et deux à deux distinctes, alors les ensembles $A(\mu_i)$ sont deux à deux disjoints, car deux mesures invariantes ergodiques sont mutuellement singulières. Mais chacun contient un élément de α et est invariant, donc chacun contient un des a_j. Cela montre donc que $k \leqslant N$.

Montrons finalement que toute mesure de probabilité ν invariante absolument continue par rapport à m est combinaison linéaire des μ_i. Si ν est ergodique, le résultat est clair. Sinon, soit A un ensemble invariant avec $0 < \mu(A) < 1$, il permet de décomposer ν en $\nu_1 + \nu_2$ où $\nu_1 = \nu_{|A}$ et $\nu_2 = \nu_{|X \setminus A}$. Si ces deux mesures sont ergodiques, on a terminé. Sinon, on continue la décomposition de la même manière. Si le processus se termine avant le rang N, on a encore terminé, et sinon on a réussi à écrire $\nu = \sum_{n=1}^{N+1} \nu_i$ où les ν_i sont invariantes et mutuellement singulières. C'est absurde, car les ensembles $A(\nu_i)$ doivent être deux à deux disjoints, et chacun contient un des ensembles a_j, avec $1 \leqslant j \leqslant N$. □

1.1.4 Rappels de théorie spectrale

Les résultats de ce paragraphe sont démontrés par exemple dans [DS58].

Soient $\mathcal{L}$ un espace de Banach complexe, et $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ l'ensemble des applications linéaires continues sur $\mathcal{L}$. Pour $Q \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$, on note $\mathrm{Sp}(Q)$ le spectre de Q, donné par

$$\mathrm{Sp}(Q) = \{\lambda \in \mathbb{C} \mid \lambda I - Q \text{ non inversible}\}.$$

C'est un compact de $\mathbb{C}$, toujours non vide. On note aussi $R(Q)$ le *rayon spectral* de Q, défini par

$$R(Q) = \sup\{|z| \mid z \in \mathrm{Sp}(Q)\}.$$

Il vérifie

$$R(Q) = \lim \|Q^n\|^{1/n} = \inf \|Q^n\|^{1/n}. \tag{1.4}$$

Si A est un ouvert fermé de $K := \mathrm{Sp}(Q)$, on peut définir le projecteur spectral $\Pi_A(Q)$ correspondant, de la façon suivante : soit γ un chemin dans $\mathbb{C}$ qui entoure A dans le sens direct, on pose

$$\Pi_A(Q) = \int_\gamma \frac{\mathrm{d}z}{zI - Q}.$$

Cet opérateur est indépendant du choix de γ. Le projecteur $\Pi_{K\backslash A}(Q)$ associé au reste du spectre vérifie alors $\mathrm{Id} = \Pi_A(Q) + \Pi_{K\backslash A}(Q)$, et $\Pi_A(Q)\Pi_{K\backslash A}(Q) = \Pi_{K\backslash A}(A)\Pi_A(Q) = 0$. De plus, ces projecteurs commutent avec Q. Ainsi, en posant $E = \mathrm{Im}\,\Pi_A(Q)$ et $F = \mathrm{Im}\,\Pi_{K\backslash A}(Q)$, on a $\mathcal{L} = E \oplus F$, et Q stabilise E et F. De plus, $\mathrm{Sp}(Q|E) = A$ et $\mathrm{Sp}(Q|F) = K\backslash A$. En ce sens, E et F sont les deux sous-espaces de $\mathcal{L}$ associés respectivement aux parties A et $K\backslash A$ du spectre de Q.

En particulier, si λ est un élément isolé du spectre de Q, on peut considérer le sous-espace spectral correspondant $E(\lambda) := \mathrm{Im}\,\Pi_{\{\lambda\}}(Q)$. On dit que λ est une *valeur propre isolée de multiplicité finie* si cet espace est de dimension finie, et on appelle *multiplicité de* λ sa dimension. Alors $E(\lambda)$ coïncide avec $\mathrm{Ker}(\lambda\,\mathrm{Id} - Q)^n$ pour tout n assez grand. En particulier, λ est une vraie valeur propre de Q.

Remarque 1.1.9. Si $E(\lambda)$ n'est pas de dimension finie, ces résultats ne sont plus valables. Par exemple, on peut avoir $\mathrm{Ker}(\lambda I - Q)^n = \{0\}$ pour tout n sans que $E(\lambda)$ ne soit réduit à zéro.

Le spectre ne dépend en général pas continûment de l'opérateur. Cependant, les valeurs propres isolées de multiplicité finie se perturbent bien :

Proposition 1.1.10. *Soit $Q \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$, et soit λ_0 une valeur propre simple isolée de Q. Il existe alors un voisinage U de Q dans $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ et un voisinage V de λ_0 dans $\mathbb{C}$ tels que tout élément $R \in U$ ait un unique élément $\lambda(R)$ de son spectre dans V, qui est en fait une valeur propre simple isolée (avec $\lambda(Q) = \lambda_0$)*

De plus, la fonction $R \mapsto \lambda(R)$ est lipschitzienne sur U.

Le même résultat est valide avec une valeur propre de multiplicité n de Q, mais elle peut donner naissance à plusieurs valeurs propres de R dans V (dont la somme des multiplicités fera toujours n). De plus, la valeur propre dépend régulièrement de l'opérateur : si on considère par exemple une perturbation holomorphe de Q, la valeur propre perturbée sera également holomorphe.

On note $\mathrm{R}_{\mathrm{ess}}(Q)$ le *rayon spectral essentiel* de Q, défini comme suit :

$$\mathrm{R}_{\mathrm{ess}}(Q) = \inf\{r > 0 \mid \forall \lambda \in \mathrm{Sp}(Q) \cap \{|z| > r\}, \\ \lambda \text{ est une valeur propre isolée de multiplicité finie de } Q\}.$$

Par exemple, le rayon spectral essentiel d'un opérateur compact est nul. Lorsque $\mathrm{R}_{\mathrm{ess}}(Q) < R(Q)$, on a donc une généralisation du cas compact, et les termes dominants dans Q^n viendront d'un nombre fini de valeurs propres (les éléments du spectre de Q de module maximal) : on dit que Q est *quasi-compact*, ou encore que Q a un *trou spectral*. Cette situation permet essentiellement de traiter Q^n comme une matrice. On a un analogue de la caractérisation dynamique (1.4) du rayon spectral pour le rayon spectral essentiel :

$$\mathrm{R}_{\mathrm{ess}}(Q) = \inf\{\|Q^n - K\|^{1/n} \mid n \in \mathbb{N}^*, K \text{ opérateur compact sur } \mathcal{L}\}. \tag{1.5}$$

Ce résultat est à rapprocher du théorème suivant de Hennion ([Hen93]) :

Théorème 1.1.11. *Soient $\mathcal{L}$ et $\mathcal{M}$ deux espaces de Banach, $i : \mathcal{L} \to \mathcal{M}$ une application compacte, et $Q \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$. On suppose qu'il existe $r \geqslant 0, C \geqslant 0$ et une suite $C_n \geqslant 0$ tels que, pour tout $n \in \mathbb{N}$, pour tout $x \in \mathcal{L}$,*

$$\|Q^n(x)\|_{\mathcal{L}} \leqslant Cr^n \|x\|_{\mathcal{L}} + C_n \|i(x)\|_{\mathcal{M}}. \tag{1.6}$$

Alors $\mathrm{R}_{\mathrm{ess}}(Q) \leqslant r$.

En effet, la partie bornée par $C_n \|i(x)\|_{\mathcal{M}}$ correspond moralement à une partie compacte de l'opérateur Q, et ne doit donc pas intervenir au niveau du rayon spectral essentiel. Les inégalités du type (1.6) sont appelées inégalités de *Doeblin-Fortet*, ou inégalités de *Lasota-Yorke*.

La proposition qui suit est due à [ITM50].

Proposition 1.1.12. *Soient $\mathcal{L} \subset \mathcal{M}$ deux espaces de Banach, avec $\mathcal{L}$ dense dans $\mathcal{M}$, et $Q \in \mathrm{Hom}(\mathcal{M}, \mathcal{M})$ qui induit un élément de $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ par restriction. On suppose que :*
- *il existe $M > 0$ tel que $\forall n \in \mathbb{N}, \|Q^n\|_{\mathcal{M}} \leqslant M$.*
- *il existe $C > 0$ et $r < 1$ tels que $\forall x \in \mathcal{L}, \forall n \in \mathbb{N}$,*

$$\|Q^n x\|_{\mathcal{L}} \leqslant Cr^n \|x\|_{\mathcal{L}} + C \|x\|_{\mathcal{M}}. \tag{1.7}$$

- *toute suite x_n bornée de $\mathcal{L}$ admet une sous-suite qui converge dans $\mathcal{M}$ vers un élément de $\mathcal{L}$.*

Alors, si $x \in \mathcal{M}$ vérifie $Q(x) = x$, on a $x \in \mathcal{L}$.

Démonstration. Soient $E = \{x \in \mathcal{L} \mid Q(x) = x\}$ et $F = \{y \in \mathcal{M} \mid Q(y) = y\}$. Montrons que E est dense dans F pour la norme de $\mathcal{M}$. Soit $y \in F$. Par densité de $\mathcal{L}$ dans $\mathcal{M}$, soit $x \in \mathcal{L}$ proche de y à ε près. Alors $\frac{1}{n}\sum_0^{n-1} Q^k(x)$ est borné dans $\mathcal{L}$ (par la deuxième hypothèse), donc on peut en extraire une sous-suite qui converge dans $\mathcal{M}$ vers $z \in \mathcal{L}$. De plus,

$$\left\| \frac{1}{n}\sum_0^{n-1} Q^k(x) - y \right\|_{\mathcal{M}} \leqslant \frac{1}{n}\sum_0^{n-1} \left\|Q^k(x-y)\right\|_{\mathcal{M}} \leqslant M \|x-y\|_{\mathcal{M}} \leqslant M\varepsilon.$$

En passant à la limite, on obtient $\|z-y\|_{\mathcal{M}} \leqslant M\varepsilon$. Comme $z \in E$, cela conclut la preuve de la densité de E dans F.

Soit maintenant $y \in F$, qu'on écrit comme limite d'une suite x_n de E. Alors x_n est une suite de Cauchy dans $(E, \| \|_{\mathcal{M}})$. Mais E est de dimension finie par le théorème 1.1.11, donc toutes les normes y sont équivalentes. En particulier, x_n est de Cauchy, donc convergente, pour $\| \|_{\mathcal{L}}$. Par conséquent, $y \in \mathcal{L}$. □

1.1.5 Propriétés spectrales des applications Gibbs-Markov

On va appliquer les méthodes spectrales du paragraphe précédent à l'étude des applications Gibbs-Markov. Les espaces considérés seront des espaces d'applications lipschitziennes sur chacun des éléments de partition, avec une condition globale de bornitude ou d'intégrabilité. En fait, le jeu de changements de distance décrit dans la remarque 1.1.5 montre que les résultats qui suivent s'appliqueront aussi à des espaces de fonctions hölderiennes.

Dans tout ce paragraphe, $(X, d, \mathcal{B}, m, T, \alpha)$ désignera une application Gibbs-Markov.

On définit l'opérateur de transfert $\widehat{T}$ associé à T par

$$\widehat{T}u(x) = \sum_{T(y)=x} g(y)u(y) \tag{1.8}$$

où g est l'inverse du jacobien de T, comme dans la définition 1.1.4. La formule du changement de variables assure que, si $u \in L^1$ et $v \in L^\infty$, on a

$$\int \widehat{T}u \cdot v \, \mathrm{d}m = \int u \cdot v \circ T \, \mathrm{d}m.$$

Ainsi, l'opérateur de transfert est le dual de l'opérateur de composition par T. Il agit sur L^1 par construction (et c'est une isométrie), mais on va voir qu'il agit sur des espaces de fonctions lipschitziennes dans le cas Gibbs-Markov. On vérifie aisément que

$$\widehat{T}^n u(x) = \sum_{T^n(y)=x} g^{(n)}(y)u(y), \tag{1.9}$$

où $g^{(n)}$ est défini dans le lemme 1.1.7.

Pour $u : X \to \mathbb{C}$, et $a \subset X$, on notera

$$Du(a) = \inf\{C > 0 \mid \text{pour presque tous } x, y \in a, |u(x) - u(y)| \leqslant Cd(x, y)\}. \tag{1.10}$$

C'est la meilleure constante de Lipschitz de u sur a. Soit

$$\mathcal{L}_{\alpha,\infty} = \left\{ u : X \to \mathbb{C} \mid \ \|u\|_\infty < \infty, \ \sup_{a \in \alpha} Du(a) < \infty \right\}.$$

C'est l'espace de Banach des fonctions lipschitziennes sur chaque élément de la partition, muni de la norme $\|u\|_{\mathcal{L}_{\alpha,\infty}} = \|u\|_\infty + \sup Du(a)$. Notons que cette norme vérifie

$$\|uv\|_{\mathcal{L}_{\alpha,\infty}} \leqslant \|u\|_{\mathcal{L}_{\alpha,\infty}} \|v\|_{\mathcal{L}_{\alpha,\infty}} . \tag{1.11}$$

On notera souvent simplement $\mathcal{L}$ au lieu de $\mathcal{L}_{\alpha,\infty}$. On utilisera également des variations de cet espace, en remplaçant ∞ par p (la condition sera alors $u \in L^p(m)$) et en changeant la partition α. Par exemple, l'espace $\mathcal{L}_{\alpha,1}$ est l'espace des fonctions lipschitziennes sur chaque élément de la partition α, et intégrables. De même, si β est la partition triviale $\beta = \{X\}$, alors $\mathcal{L}_{\beta,\infty}$ est l'espace des fonctions lipschitziennes sur tout X. En fait, les éléments de ces espaces ne sont pas des fonctions, mais des classes de fonctions (i.e. on identifie deux fonctions égales presque partout). On travaillera en général avec des représentants, sans détailler.

Dans la formule (1.9), presque tout point x a au plus un antécédent dans chaque cylindre $\underline{a} = [a_0, \dots, a_{n-1}]$ de longueur n, qu'on note $\underline{a}x$. Posons $M_{\underline{a}}u(x) = g^{(n)}(\underline{a}x)u(\underline{a}x)$ si ce point est défini, et 0 autrement. Ainsi, $\widehat{T}^n = \sum M_{\underline{a}}$ où la somme porte sur tous les cylindres de longueur n.

On notera α_* la partition engendrée par les images des éléments de α : c'est une partition, éventuellement moins fine que α, et tous les $T(a)$, pour $a \in \alpha$, sont α_*-mesurables. Rappelons que λ est le coefficient de dilatation de T.

Lemme 1.1.13. *Il existe une constante B telle que $\forall \underline{a} = [a_0, \dots, a_{n-1}]$, $\forall u : X \to \mathbb{C}$,*

$$\|M_{\underline{a}}u\|_{\mathcal{L}_{\alpha_*,\infty}} \leqslant B\frac{m[\underline{a}]}{m[T^n\underline{a}]}\left(\lambda^{-n}Du(a_0) + \frac{1}{m[\underline{a}]}\int_{[\underline{a}]} |u|\,\mathrm{d}m\right).$$

Démonstration. Soient x et y dans le même élément a_* de la partition α_*. Si cet élément n'est pas dans l'image $T(a_{n-1})$, $(M_{\underline{a}}u)_{|a_*} = 0$ et il n'y a rien à faire. Sinon, soient $x' = \underline{a}x$ et $y' = \underline{a}y$.

Pour presque tout $z' \in \underline{a}$, on a $|u(y') - u(z')| \leqslant Du(a_0)d(y', z') \leqslant C\,Du(a_0)\lambda^{-n}$. On obtient en intégrant sur z' que $|u(y')| \leqslant C\ Du(a_0)\lambda^{-n} + \frac{1}{m[\underline{a}]}\int_{\underline{a}} |u|\,\mathrm{d}m$. L'inégalité $g^{(n)}(y') \leqslant D\frac{m[\underline{a}]}{m[T^n\underline{a}]}$ (lemme 1.1.7) donne la borne annoncée sur $\|M_{\underline{a}}u\|_\infty$.

D'après le lemme 1.1.7, on a

$$\begin{aligned}|M_{\underline{a}}u(x) - M_{\underline{a}}u(y)| &\leqslant |g^{(n)}(x')||u(x') - u(y')| + |u(y')||g^{(n)}(y')|\left|1 - \frac{g^{(n)}(x')}{g^{(n)}(y')}\right| \\ &\leqslant D\frac{m[\underline{a}]}{m[T^n\underline{a}]}Du(a_0)\lambda^{-n}d(x,y) + |u(y')|D\frac{m[\underline{a}]}{m[T^n\underline{a}]}Cd(x,y),\end{aligned}$$

ce qui donne la conclusion en utilisant la borne précédente sur $|u(y')|$. □

Corollaire 1.1.14. *L'opérateur de transfert agit continûment sur $\mathcal{L}$. En fait, il envoie $\mathcal{L}_{\alpha,1}$ dans $\mathcal{L}_{\alpha_*,\infty}$. Finalement, on a une inégalité*

$$\left\|\widehat{T}^n u\right\|_{\mathcal{L}} \leqslant C\lambda^{-n}\|u\|_{\mathcal{L}} + C\|u\|_1 . \tag{1.12}$$

Démonstration. Comme T a la propriété de grande image, les mesures $m[T^n \underline{a}]$ (pour $\underline{a}$ cylindre de longueur n) sont uniformément minorées. En sommant les estimées données par le lemme 1.1.13, on trouve

$$\left\|\widehat{T}^n u\right\|_{\mathcal{L}_{\alpha_*,\infty}} \leqslant C\lambda^{-n} \sup_{a\in\alpha} Du(a) + C\left\|u\right\|_1 .$$

Comme $\|v\|_{\mathcal{L}_{\alpha,1}} \leqslant \|v\|_{\mathcal{L}} \leqslant \|v\|_{\mathcal{L}_{\alpha_*,\infty}}$ pour toute fonction v, cela donne la conclusion. □

On a même un résultat un peu plus précis :

Proposition 1.1.15. *Soit $u : X \to \mathbb{C}$ intégrable telle que $\sum_{a\in\alpha} m[a]Du(a) < \infty$. Alors $\widehat{T}u \in \mathcal{L}_{\alpha_*,\infty}$.*

Démonstration. Il suffit de sommer les estimées données par le lemme 1.1.13 sur tous les cylindres de longueur 1 : la condition $\sum m[a]Du(a) < \infty$ assure qu'on obtient une borne finie pour $\left\|\widehat{T}u\right\|_{\mathcal{L}_{\alpha_*,\infty}}$. □

L'équation (1.12) est du même type que les équations (1.6) et (1.7). Pour en déduire des résultats, il faut vérifier la compacité de l'injection de $\mathcal{L}$ dans L^1, qui est donnée par le lemme suivant (c'est une version du théorème classique d'Ascoli-Arzela) :

Proposition 1.1.16. *Soit f_p une suite bornée de $\mathcal{L}$. Elle admet alors une sous-suite qui converge dans $L^1(m)$ vers un élément de $\mathcal{L}$.*

Démonstration. Soit M une borne sur $\|f_p\|_{\mathcal{L}}$. Soit X' une partie de X de mesure pleine telle que, pour tout $n \in \mathbb{N}$, pour tous points x et y de X' dans un même élément de la partition α, on ait $|f_p(x) - f_p(y)| \leqslant Md(x, y)$ et $|f_p(x)| \leqslant M$ (par définition de $\mathcal{L}$).

Pour tout cylindre $\underline{a}$ de mesure positive, considérons un point $x_{\underline{a}}$ de ce cylindre, appartenant à X'. La suite $f_p(x_{\underline{a}})$, étant bornée, admet une sous-suite convergente. Par procédé diagonal, on peut supposer qu'on a convergence de toutes les suites $f_p(x_{\underline{a}})$ pour tous les cylindres $\underline{a}$ de mesure non nulle. Soient $x \in X'$, et $\underline{a} = [a_0, \dots, a_n]$ le cylindre de longueur $n+1$ le contenant. Alors

$$|f_p(x) - f_q(x)| \leqslant |f_p(x) - f_p(x_{\underline{a}})| + |f_p(x_{\underline{a}}) - f_q(x_{\underline{a}})| + |f_q(x_{\underline{a}}) - f_q(x)|. \tag{1.13}$$

Les premier et troisième termes sont majorés par $C\lambda^{-n}$, tandis que le second est arbitrairement petit si p et q sont assez grands. Ainsi, $(f_p(x))$ est une suite de Cauchy, qui converge donc. Autrement dit, f_p converge sur X' vers une fonction f. En passant à la limite simple dans les inégalités $|f_p(x) - f_p(y)| \leqslant Md(x, y)$, on obtient que f est lipschitzienne. De plus, f est bornée sur X' par M. Ainsi, $f \in \mathcal{L}$.

Montrons la convergence de f_p vers f dans L^1. Soient ε positif, et n tel que le diamètre de tout cylindre de longueur n soit majoré par ε. Pour $x \in \underline{a}$ de longueur n, (1.13) donne (en passant à la limite sur q)

$$|f_p(x) - f(x)| \leqslant 2M \operatorname{diam} \underline{a} + |f_p(x_{\underline{a}}) - f(x_{\underline{a}})| \leqslant 2M\varepsilon + |f_p(x_{\underline{a}}) - f(x_{\underline{a}})|.$$

Soient $\underline{a}_1, \dots, \underline{a}_k$ un nombre fini de cylindres de longueur n, remplissant au moins une mesure $1-\varepsilon$. Pour p assez grand, on a $|f_p(x_{\underline{a}_i}) - f(x_{\underline{a}_i})| \leqslant \varepsilon$ pour tout $1 \leqslant i \leqslant k$, et on obtient

$$\int |f_p - f| \leqslant \sum \int_{\underline{a}_i} |f_p - f| + 2Mm\left(X \setminus \bigcup \underline{a}_i\right) \leqslant \sum (2M+1)\varepsilon m[\underline{a}_i] + 2M\varepsilon \leqslant (4M+1)\varepsilon. \ \square$$

On peut alors décrire précisément le comportement spectral de $\widehat{T}$:

Proposition 1.1.17. *On reprend les hypothèses et les notations du théorème 1.1.8.*

L'opérateur de transfert $\widehat{T}$ agissant sur $\mathcal{L}$ a un trou spectral : il a un nombre fini de valeurs propres de modules 1, et le reste de son spectre est inclus dans un disque de rayon < 1.

De plus, 1 est une valeur propre de $\widehat{T}$, et une base du sous-espace propre correspondant est donnée par les densités $\frac{\mathrm{d}\mu_i}{\mathrm{d}m}$ des mesures de probabilité ergodiques absolument continues (qui sont donc en particulier dans $\mathcal{L}$, et même dans $\mathcal{L}_{\alpha_*,\infty}$).

Enfin, les autres valeurs propres de module 1 de $\widehat{T}$ sont des racines de l'unité. Si u est une fonction propre correspondante, alors $|u|$ est une fonction propre pour la valeur propre 1 et, sur chaque élément de partition a, il existe un complexe $e^{i\theta(a)}$ de module 1 tel que $u = e^{i\theta(a)}|u|$.

Démonstration. Comme le plongement de $\mathcal{L}$ dans $L^1(m)$ est compact d'après la proposition 1.1.16, l'inégalité (1.12) et le théorème 1.1.11 impliquent que $\widehat{T}$ a un trou spectral. De plus, la proposition 1.1.12 montre que les fonctions de L^1 satisfaisant $\widehat{T}u = u$ sont en fait dans $\mathcal{L}$. Ceci s'applique en particulier aux densités $f_i = \frac{\mathrm{d}\mu_i}{\mathrm{d}m}$. En fait, comme $\widehat{T}$ envoie $\mathcal{L}$ dans $\mathcal{L}_{\alpha_*,\infty}$ (corollaire 1.1.14), ces fonctions propres sont même dans $\mathcal{L}_{\alpha_*,\infty}$.

Montrons que ces densités forment une base du sous-espace propre associé à la valeur propre 1. Soit u une fonction propre, qu'on décompose comme $u = u_1 - u_2 + iu_3 - iu_4$ où les fonctions u_s sont dans $\mathcal{L}$ et à valeurs dans $\mathbb{R}_+$. Pour tout s, en utilisant la décomposition spectrale correspondant aux valeurs propres de module 1 et au reste du spectre, on obtient la convergence de $\frac{1}{n}\sum_0^{n-1} \widehat{T}^k(u_s)$ vers une fonction v_s, qui vérifie $\widehat{T}v_s = v_s$. La mesure $v_s\,\mathrm{d}m$ est donc invariante, si bien que le théorème 1.1.8 assure qu'il existe des coefficients λ_s^i tels que $v_s = \sum \lambda_s^i f_i$. Ainsi, il existe des coefficients λ_i tels que $v := v_1 - v_2 + iv_3 - iv_4$ s'écrive $\sum \lambda_i f_i$. Mais v est la limite de $\frac{1}{n}\sum_0^{n-1} \widehat{T}^k u$, par construction. Comme $\widehat{T}u = u$, on obtient $v = u$, ce qui conclut.

Soit maintenant u une fonction propre non nulle de $\widehat{T}$ pour une valeur propre λ de module 1 : on a presque partout

$$\lambda u(x) = \sum_{T(y)=x} g(y)u(y).$$

Comme $g \geqslant 0$, on obtient

$$|u(x)| \leqslant \sum_{T(y)=x} g(y)|u(y)|,$$

i.e. $|u| \leqslant \widehat{T}|u|$. Comme $\int |u| = \int \widehat{T}|u|$, on obtient $|u| = \widehat{T}|u|$ presque partout, i.e. $|u|$ est une fonction propre pour la valeur propre 1.

Notons θ l'argument de u lorsque $u \neq 0$, et montrons que θ est presque partout constant sur les éléments de la partition α. Tout d'abord, comme $|u|$ est combinaison linéaire des f_i, l'ensemble $\{u \neq 0\}$ est réunion d'éléments de la partition. Soit $a \in \alpha$ sur lequel $u \neq 0$. Le théorème de récurrence de Poincaré et le théorème de convergence des martingales assurent l'existence de $x \in a$ dont une infinité d'itérés appartiennent à a et tel que, en notant $[a_0, \dots, a_n]$ le cylindre de longueur $n+1$ contenant x, on ait pour tout $\varepsilon > 0$

$$\frac{m\{y \in [a_0, \dots, a_n] \mid |\theta(y) - \theta(x)| > \varepsilon\}}{m[a_0, \dots, a_n]} \to 0.$$

Par distorsion bornée, on obtient pour n assez grand

$$\frac{m\{y' \in a \mid \exists y \in [a_0, \dots, a_n], T^n(y) = y' \text{ et } |\theta(y) - \theta(x)| > \varepsilon\}}{m[a]} \leqslant \varepsilon.$$

Si y' n'est pas dans cet ensemble U_n, il admet un antécédent dans $[a_0, \dots, a_n]$ avec $|\theta(y) - \theta(x)| \leqslant \varepsilon$. Mais $\lambda^n u(y') = \sum_{T^n(z)=y'} g^{(n)}(z)u(z)$, et on a égalité dans l'inégalité de Cauchy-Schwarz, donc tous les points qui interviennent dans la somme ont même argument. En particulier, $\lambda^n e^{i\theta(y')} = e^{i\theta(y)}$. Comme $|\theta(y) - \theta(x)| \leqslant \varepsilon$, on en déduit que θ est constant à 2ε près sur $a \backslash U_n$, qui occupe une proportion $1 - \varepsilon$ de a. En passant à la limite, θ est constant presque partout sur a.

Montrons maintenant que λ est une racine de l'unité. Soit $a \in \alpha$ avec $|u| > 0$ sur a. Par le théorème de récurrence de Poincaré, il existe un entier n avec $T^n(a) \supset a$. Si $x \in a$, on a $\lambda^n u(x) = \sum_{T^n(y)=x} g^{(n)}(y)u(y)$, et tous les points du membre de droite de cette égalité ont même argument. En prenant pour y un antécédent de x dans a, on obtient $\lambda^n e^{i\theta(x)} = e^{i\theta(y)}$. Comme θ est constant sur a, cela montre que $\lambda^n = 1$.

Montrons finalement que 1 est effectivement une valeur propre de $\widehat{T}$ (i.e. qu'on n'a pas raisonné sur du vide). Si $\widehat{T}$ admet une valeur propre de module 1, de fonction propre correspondante u, alors $|u|$ est une fonction propre pour la valeur propre 1, et on a terminé. Sinon, par quasi-compacité, le spectre de $\widehat{T}$ est inclus dans un disque de rayon < 1, si bien que $\widehat{T}^n u$ tend exponentiellement vite vers 0 pour toute fonction $u \in \mathcal{L}$. Cependant, comme $\int \widehat{T}^n 1 = \int 1 = 1$, c'est absurde. □

Cette description implique les corollaires suivants :

Corollaire 1.1.18. *Soient (X, T, m) une application Gibbs-Markov, et $X_1, \dots, X_N$ les supports des mesures de probabilité invariantes $\mu_1, \dots, \mu_N$, donnés par le théorème 1.1.8. Alors*

$$\bigcup_{i=1}^{N} \bigcup_{n \geqslant 0} T^{-n}(X_i) = X \mod 0.$$

En particulier, toute mesure invariante absolument continue est combinaison linéaire des μ_i, et est donc de masse finie.

Démonstration. Pour toute fonction $u \in \mathcal{L}$, la proposition 1.1.17 implique que la suite $\frac{1}{n}\sum_{k=0}^{n-1} \widehat{T}^k(u)$ converge dans $\mathcal{L}$, et donc dans L^1. Par densité, on obtient la même convergence pour toute fonction de L^1.

Soit $A = X \setminus \bigcup_{i,n} T^{-n}(X_i)$. Alors $\frac{1}{n}\sum_{k=0}^{n-1} \widehat{T}^k(1_A)$ converge dans L^1 vers une fonction u, qui satisfait $\widehat{T}u = u$. Ainsi, par la proposition 1.1.17, u est combinaison linéaire des densités des mesures invariantes, si bien que u est supportée par $\bigcup X_i$. Mais, comme $T(A) \subset A$, les fonctions $\widehat{T}^k(1_A)$ sont toutes supportées par A, donc u est supportée par A. Comme $A \cap \bigcup X_i = \emptyset$, $u = 0$. Mais $\int u = m(A)$, donc $m(A) = 0$.

Soit maintenant μ une mesure invariante absolument continue, mais pas nécessairement de masse finie. Soit $X_i' = \bigcup T^{-n}(X_i)$. Alors $T^{-1}(X_i') = X_i'$, et la restriction de m à X_i' est ergodique : si $B \subset X_i'$ vérifie $T^{-1}(B) = B$ et $m(B) > 0$, alors $m(B \cap T^{-n}(X_i)) > 0$ pour un certain n, puis $m(B \cap X_i) > 0$ par invariance. Comme μ_i est ergodique, on obtient $X_i \subset B$ mod 0, puis $X_i' \subset B \mod 0$. Ainsi, μ_i est l'unique mesure invariante absolument continue sur X_i', par [Aar97, théorème 1.5.6]). En particulier, $\mu_{|X_i'}$ est proportionnelle à μ_i. En sommant (et comme $\bigcup X_i' = X \mod 0$), on obtient que μ est combinaison linéaire des μ_i. □

Corollaire 1.1.19. *Si (X, T, m) est une application Gibbs-Markov transitive, elle admet une unique mesure de probabilité μ invariante ergodique absolument continue par rapport à m. De plus, il existe un entier $s \geqslant 1$ et une décomposition $\mu = \nu_1 + \dots + \nu_s$ telle que $T_*\nu_i = \nu_{i+1}$ pour $i < s$, $T_*\nu_s = \nu_1$, et T^s est une application Gibbs-Markov mélangeante sur le support de ν_i, pour $1 \leqslant i \leqslant s$.*

En particulier, si tous les itérés de T sont Markov-transitifs, alors T est Markov-mélangeante.

Démonstration. Par le théorème de récurrence de Poincaré, T est transitive sur le support de chaque mesure μ_i donnée par la proposition 1.1.17. Comme T est transitive, il existe donc une seule mesure μ_i, qu'on note μ. Soient $\lambda_1, \dots, \lambda_k$ les autres valeurs propres de module 1 de $\widehat{T}$: ce sont des racines de l'unité, donc il existe N tel que $\lambda_i^N = 1$ pour tout i. L'application T^N est donc Gibbs-Markov, et son opérateur de transfert a 1 pour unique valeur propre de module 1. Ainsi, pour $k \geqslant 1$, toute fonction propre de $\widehat{T^N}^k$ pour la valeur propre 1 est déjà fonction propre de $\widehat{T^N}$ pour la même valeur propre.

Soient $\nu_1, \dots, \nu_l$ les mesures ergodiques invariantes de T^N, de densités respectives $f_1, \dots, f_l$. La mesure $T_*\nu_1$ est encore invariante par T^N et ergodique, c'est donc l'une des ν_k. On applique T_* à cette mesure, ce qui redonne une mesure $\nu_{k'}$. On continue à itérer ainsi, et on retombe sur une mesure déjà vue, ce qui donne un cycle. Quitte à renuméroter les mesures, on peut supposer que ce cycle est $(\nu_1, \dots, \nu_s)$. Alors la mesure $\nu_1 + \dots + \nu_s$ est invariante par T_*, et c'est donc μ par unicité.

Comme $T^s_*(\nu_i) = \nu_i$ pour tout i, T^s est une application Gibbs-Markov sur le support Y_i de ν_i, on va montrer qu'elle est mélangeante. Soit a un élément de α avec $\nu_i(a) > 0$, et montrons que $d := \mathrm{pgcd}\{n \mid a \subset T^{sn}(a)\} = 1$. Sinon, définissons des ensembles A_j, pour $0 \leqslant j \leqslant d-1$, par

$$A_j = \bigcup T^{s(nd+j)}(a).$$

Ces ensembles forment une partition de Y_i. Soit ν_i' la restriction de ν_i à A_0 : cette mesure est invariante par T^{sd}, et donc par T^{Nsd}. Mais toutes les mesures invariantes par T^{Nsd} sont invariantes par T^N, ce qui est absurde car le support de ν_i' est strictement inclus dans celui de ν_i. Cela montre que $d = 1$.

On en déduit que, pour tout n assez grand, $T^{sn}(a) \supset a$. La transitivité de T^s sur le support de ν_i (due au théorème de récurrence de Poincaré) implique alors aisément que T^s est mélangeante sur ce support. □

Corollaire 1.1.20. *Si T est Markov-transitive, le projecteur spectral associé à la valeur propre simple 1 de $\widehat{T}$ est donné par $Pu(x) = \left(\int u \,\mathrm{d}m\right) f(x)$, où f est la densité de l'unique mesure de probabilité invariante μ.*

Démonstration. La suite $u_n = \frac{1}{n}\sum_0^{n-1} \widehat{T}^k u$ converge vers Pu, qui est un multiple λf de f. Mais $\int u_n \,\mathrm{d}m = \int u \,\mathrm{d}m$, donc en passant à la limite, et en utilisant $\int f \,\mathrm{d}m = 1$, on obtient $\lambda = \int u \,\mathrm{d}m$. □

Corollaire 1.1.21. *Si T est une application Gibbs-Markov mélangeante, alors elle admet une unique mesure de probabilité invariante absolument continue et ergodique μ, qui est mélangeante. De plus, les corrélations décroissent exponentiellement vite pour cette mesure : il existe $\rho < 1$ et $C > 0$ tels que, pour toutes fonctions $u \in \mathcal{L}$ et $v \in L^\infty$,*

$$\left| \int u \cdot v \circ T^n \,\mathrm{d}\mu - \int u \,\mathrm{d}\mu \int v \,\mathrm{d}\mu \right| \leqslant C \left\|u\right\|_{\mathcal{L}} \left\|v\right\|_\infty \rho^n.$$

Démonstration. Comme T est mélangeante, la décomposition de la mesure invariante μ donnée par le corollaire 1.1.19 est nécessairement triviale. De plus, le même argument montre que 1 est l'unique valeur propre de module 1 de $\widehat{T}$, et qu'elle est simple. Soit f la fonction propre correspondante (c'est en fait la densité de μ), normalisée pour que μ soit une mesure de probabilité, et P le projecteur spectral associé à la valeur propre 1 de $\widehat{T}$.

Si $u \in \mathcal{L}$, alors $\widehat{T}^n(u)$ converge exponentiellement vite (à une vitesse ρ^n, avec $\rho < 1$) vers $P(u) = \left(\int u\,\mathrm{d}m\right) f$. En appliquant ce résultat à uf, on obtient

$$\widehat{T}^n(uf) = \left(\int uf\,\mathrm{d}m\right) f + O(\rho^n) = \left(\int u\,\mathrm{d}\mu\right) f + O(\rho^n).$$

Finalement,

$$\begin{aligned}\int u \cdot v \circ T^n\,\mathrm{d}\mu &= \int uf \cdot v \circ T^n\,\mathrm{d}m = \int \widehat{T}^n(uf) v\,\mathrm{d}m = \int \left(\int u\,\mathrm{d}\mu\right) fv\,\mathrm{d}m + O(\rho^n) \\ &= \left(\int u\,\mathrm{d}\mu\right)\left(\int v\,\mathrm{d}\mu\right) + O(\rho^n).\end{aligned}$$

□

Ces corollaires permettent de réduire l'étude des mesures invariantes des applications Gibbs-Markov d'abord au cas Markov-transitif, puis au cas Markov-mélangeant, dans lequel on a des propriétés de mélange très fortes, comme le montre le corollaire 1.1.21. Il serait également intéressant de pouvoir se restreindre au cas où m est la mesure invariante, mais ce n'est pas toujours possible (voir cependant la proposition 1.1.24) :

Remarque 1.1.22. Si T est une application Gibbs-Markov mélangeante pour une mesure m, elle n'est plus nécessairement Gibbs-Markov pour la mesure invariante.

Exemple 1.1.23. On considère la marche aléatoire sur $\mathbb{N}^*$, avec les probabilités de transition $P(n \to n+1) = \frac{1}{n^2}$ et $P(n \to 1) = 1 - \frac{1}{n^2}$, et la distribution initiale $P(X_0 = n) = \frac{6}{\pi^2}\frac{1}{n^2}$. Dynamiquement, cela correspond à considérer $X = \{(n_0, n_1, \ldots) \mid \forall i, n_{i+1} = n_i + 1 \text{ ou } n_{i+1} = 1\}$, muni de la mesure m définie sur les cylindres par

$$m[n_0, n_1, \ldots, n_k] = P(X_0 = n_0)P(n_0 \to n_1)\cdots P(n_{k-1} \to n_k).$$

Comme $m[n_0, \ldots, n_k] = \sum_n m[n_0, \ldots, n_k, n]$ par construction, le théorème de prolongement de Kolmogorov assure que m définit bien une mesure sur X. On considère enfin pour T le décalage à gauche. Le système (X, m, T) est alors markovien pour la partition constituée des $[n]$, pour $n \in \mathbb{N}^*$. La distorsion g est donnée par

$$\begin{aligned}g(n_0, n_1, \ldots) = \lim_{k\to\infty} \frac{m[n_0, \ldots, n_k]}{m[n_1, \ldots, n_k]} &= \frac{P(X_0 = n_0)P(n_0 \to n_1)}{P(X_0 = n_1)} \\ &= \begin{cases} \frac{(n_0+1)^2}{n_0^2}\frac{1}{n_0^2} & \text{si } n_1 = n_0 + 1, \\ \frac{1}{n_0^2}\left(1 - \frac{1}{n_0^2}\right) & \text{si } n_1 = 1. \end{cases}\end{aligned}$$

Ainsi, $g(x)/g(y) = 1$ si $s(x, y) \geqslant 2$, et $g(x)/g(y)$ est borné si $s(x, y) = 1$. Cela montre que T est Gibbs-Markov.

On vérifie aisément que la mesure stationnaire pour la chaîne de Markov est donnée par $P(X=n) = \frac{\left(\sum 1/(n!)^2\right)^{-1}}{\left((n-1)!\right)^2}$. Autrement dit, la mesure m' définie par

$$m'[n_0, \dots, n_k] = P(X = n_0)P(n_0 \to n_1) \cdots P(n_{k-1} \to n_k)$$

est invariante par T, et absolument continue par rapport à m, de densité $\frac{P(X=n_0)}{P(X_0=n_0)}$, qui tend vers 0 quand $n_0 \to \infty$. La distorsion g' de m' est donnée par

$$g'(n_0, n_1, \dots) = \begin{cases} 1 & \text{si } n_1 = n_0 + 1, \\ \left((n_0-1)!\right)^2 \left(1 - \frac{1}{n_0^2}\right) & \text{si } n_1 = 1. \end{cases}$$

et n'est donc pas lipschitzienne. Ainsi, (X, T, m') n'est pas Gibbs-Markov.

La mesure invariante se comporte mieux dans le cas où les éléments de la partition de Markov n'ont qu'un nombre fini d'images possibles ([Aar97, lemme 4.4.1]) :

Proposition 1.1.24. *Supposons que (X, T, α, m) est une application Gibbs-Markov mélangeante, et que $\mathrm{Card}(\alpha_*) < \infty$, où α_* est la partition engendrée par les images des éléments de α. Alors il existe une constante $C > 0$ telle que la densité f de la mesure de probabilité invariante μ satisfasse presque partout $\frac{1}{C} \leqslant f(x) \leqslant C$.*

En particulier, (X, T, α, μ) est une application Gibbs-Markov mélangeante, qui préserve la probabilité μ.

1.2 Cobords pour les applications Gibbs-Markov

Il sera plus tard important de savoir caractériser les fonctions qui sont des *cobords*, c'est-à-dire qui s'écrivent presque partout sous la forme $f = u - u \circ T$ pour une certaine fonction u. Sous des hypothèses de régularité sur f, on va voir que u est nécessairement également régulière. Ainsi, on pourra parfois montrer que l'égalité $f = u - u \circ T$ a en fait lieu partout, ce qui permettra d'utiliser le comportement de f le long des orbites périodiques. En particulier, si $T^n(x) = x$, alors le fait que la somme de Birkhoff $S_n f(x) = \sum_{k=0}^{n-1} f(T^k x)$ soit non nulle sera une obstruction au fait que $f = u - u \circ T$.

On aura dans la suite besoin de ce type de résultats pour des fonctions à valeurs dans $\mathbb{R}$ ou dans S^1. On énonce donc le théorème pour des fonctions à valeurs dans un groupe G abélien localement compact polonais, muni d'une distance invariante pour l'addition et l'inversion, qu'on notera $|x - y|$.

Théorème 1.2.1. *Soit (X, T, m, α) une application Gibbs-Markov qui préserve une mesure de probabilité m. Soient $f : X \to G$ satisfaisant $\sum_{a \in \alpha} m[a] Df(a) < +\infty$, et $u : X \to G$ une fonction mesurable telle que $f = u - u \circ T$ presque partout.*

Alors $\sup_{a_*\in\alpha_*} Du(a_*) < \infty$*, où* α_* *est la partition engendrée par les images des éléments de* α*.*

La preuve montrera en fait qu'il existe une constante C qui ne dépend que de T telle que $\sup_{a_*\in\alpha_*} Du(a_*) \leqslant C\sum m(a)Df(a)$. En particulier, quand $Df(a) = 0$ pour tout a, i.e. f est constante sur les éléments de la partition, on trouve que $Du(a_*) = 0$, i.e. u est constante sur les éléments de α_*. Dans le cas $G = S^1$, on retrouve [AD01b, théorème 3.1], avec une preuve totalement différente.

Le théorème serait beaucoup plus facile à démontrer sous l'hypothèse plus forte

$$\sup_{a\in\alpha} Df(a) < \infty. \tag{1.14}$$

Cependant, cette hypothèse est trop restrictive car elle n'est pas compatible avec le processus d'induction qu'on décrira dans la partie 1.3. Elle ne permettrait en particulier pas d'obtenir des extensions du théorème 1.2.1 à des cas où la dilatation est non uniforme, comme le théorème 1.3.16.

1.2.1 Preuve du théorème 1.2.1

D'après le théorème 1.1.8, l'espace X se décompose en un nombre fini de sous-systèmes sur lesquels T est Gibbs-Markov transitive. Il suffit de prouver le théorème sur chacun de ces sous-systèmes, et on peut donc supposer sans perte de généralité que T est transitive.

L'étape la plus importante de la preuve est le lemme suivant :

Lemme 1.2.2. *Il existe* $\alpha_1 \in \alpha$ *tel que* $Du(\alpha_1) < \infty$*.*

Démonstration. Soit $\lambda > 1$ le coefficient de dilatation de f. Ainsi, pour presque tous x, y dans le même élément de la partition α, on a $d(Tx, Ty) \geqslant \lambda d(x, y)$.

Soit Φ la fonction définie par $\Phi(x) = Df(a)$ lorsque $x \in a$: par hypothèse, cette fonction est intégrable. En particulier, sur un ensemble X_1 de mesure pleine, la somme de Birkhoff $S_n\Phi(x)$ est un $O(n)$.

Il existe également X_2 de mesure pleine tel que tous les itérés d'un point $x \in X_2$ vérifient : pour presque tout y dans le même élément de partition a que $T^n x$, $|f(y) - f(T^n x)| \leqslant Df(a)d(y, T^n x)$.

Enfin, le théorème de convergence des martingales assure que presque tout point est un point de continuité mesurable de u : il existe X_3 de mesure pleine tel que, si $x \in X_3$ et $a_0, a_1, \ldots$ désigne la suite des éléments de α qui contiennent $x, T(x), \ldots$, alors pour tout $\varepsilon > 0$

$$\frac{m\{y \in [a_0, \ldots, a_{n-1}] \mid |u(y) - u(x)| > \varepsilon\}}{m[a_0, \ldots, a_{n-1}]} \to 0.$$

Comme T est Gibbs-Markov, la distorsion de tous ses itérés est bornée d'après le lemme 1.1.7. Il existe donc une constante B telle que, pour tout ensemble mesurable Z et pour tout cylindre de longueur k,

$$\frac{m([a_0,\dots,a_{k-1}]\cap T^{-k}Z)}{m[a_0,\dots,a_{k-1}]} \leqslant B\frac{m(T(a_{k-1})\cap Z)}{m(Ta_{k-1})}.$$

Comme T a la propriété de grande image, on en déduit l'existence d'une constante B' telle que

$$\frac{m([a_0,\dots,a_{k-1}]\cap T^{-k}Z)}{m[a_0,\dots,a_{k-1}]} \leqslant B'm(Z). \tag{1.15}$$

Soient K un nombre assez grand pour que

$$K\log\lambda > 3, \tag{1.16}$$

puis $\alpha_1,\dots,\alpha_N$ un nombre fini d'éléments de la partition α dont le complémentaire est de mesure ε_0 assez petite pour que $K\log(1-B'\varepsilon_0)\geqslant -1/2$. Notons

$$Z_n = \{x \mid \forall n^3 \leqslant k < n^3 + \lfloor K\log n\rfloor, T^k(x)\in\alpha_1\cup\ldots\cup\alpha_N\}.$$

Soit enfin X_4 l'ensemble des points qui sont dans une infinité de Z_n.

Lemme 1.2.3. *L'ensemble X_4 est de mesure non nulle.*

Démonstration. On notera $A=\alpha_1\cup\ldots\cup\alpha_N$.

Minorons tout d'abord $m(Z_n)$. Pour tout cylindre $[a_0,\dots,a_{k-1}]$, on applique (1.15) au complémentaire de A, de mesure ε_0, et on obtient

$$m([a_0,\dots,a_{k-1}]\cap T^{-k}A)\geqslant(1-B'\varepsilon_0)m[a_0,\dots,a_{k-1}].$$

En additionnant ces inégalités pour a_{k-1} parcourant $\alpha_1,\dots,\alpha_N$, on obtient

$$m([a_0,\dots,a_{k-2}]\cap T^{-k+1}A\cap T^{-k}A)\geqslant(1-B'\varepsilon_0)m([a_0,\dots,a_{k-2}]\cap T^{-k+1}A).$$

Cette dernière expression majore $(1-B'\varepsilon_0)^2m[a_0,\dots,a_{k-2}]$, encore d'après (1.15). On démontre ainsi par récurrence que

$$m\big([a_0,\dots,a_l]\cap T^{-l-1}A\cap\cdots\cap T^{-k}A\big)\geqslant(1-B'\varepsilon_0)^{k-l}m[a_0,\dots,a_l].$$

En particulier, pour $l=-1$ et $k=\lfloor K\log n\rfloor-1$, on trouve en utilisant l'invariance de m que

$$m(Z_n)\geqslant(1-B'\varepsilon_0)^{K\log n}=n^{K\log(1-B'\varepsilon_0)}\geqslant\frac{1}{\sqrt{n}}$$

grâce au choix de ε_0.

Ainsi, $\sum m(Z_n) = \infty$. On va appliquer une version du lemme de Borel-Cantelli pour conclure. Comme les ensembles Z_n ne sont pas indépendants, on va utiliser la version suivante de ce lemme, due à Lamperti ([Spi64, proposition 6.26.3]) :

Si $\sum m(Z_n) = \infty$ *et*

$$\liminf_{n\to\infty} \frac{\sum_{j,k=1}^n m(Z_j \cap Z_k)}{\left(\sum_{k=1}^n m(Z_k)\right)^2} < \infty$$

alors l'ensemble des points appartenant à une infinité d'ensembles Z_n *est de mesure non nulle.*

Il faut donc estimer $m(Z_j \cap Z_k)$. On va utiliser l'opérateur de transfert $\widehat{T}$, qui agit sur $\mathcal{L}$ l'espace des fonctions lipschitziennes sur chaque élément de la partition, par le corollaire 1.1.14.

Soient χ la fonction caractéristique de A, et $\gamma_n = \prod_{0\leqslant k<\lfloor K\log n\rfloor} \chi\circ T^k$: ainsi, $m(Z_n) = \int \gamma_n \circ T^{n^3} = \int \gamma_n$, et $m(Z_n\cap Z_p) = \int \gamma_n\circ T^{n^3}\cdot\gamma_p\circ T^{p^3}$. La fonction χ est dans $\mathcal{L}$. Ainsi, pour $k > j$

$$m(Z_j\cap Z_k) = \int \gamma_j\circ T^{j^3}\cdot\gamma_k\circ T^{k^3} = \int \widehat{T}^{k^3-j^3}(\gamma_j)\cdot\gamma_k \leqslant \left\|\widehat{T}^{k^3-j^3}(\gamma_j)\right\|_{\mathcal{L}} m(Z_k). \tag{1.17}$$

Comme $\widehat{T}$ agit continûment sur $\mathcal{L}$, la fonction

$$\delta_j = \widehat{T}^{\lfloor K\log j\rfloor}(\gamma_j) = \widehat{T}(\chi\widehat{T}(\chi\cdots\widehat{T}(\chi))\cdots))$$

vérifie $\|\delta_j\|_{\mathcal{L}} \leqslant C^{K\log j}$. L'inégalité (1.12), appliquée à $k^3 - j^3 - \lfloor K\log j\rfloor$ et δ_j, donne

$$\left\|\widehat{T}^{k^3-j^3}\gamma_j\right\| \leqslant C\left(\lambda^{-(k^3-j^3-K\log j)}\|\delta_j\|_{\mathcal{L}} + \|\delta_j\|_1\right) \leqslant C\left(\lambda^{-(k^3-j^3-K\log j)}C^{K\log j} + m(Z_j)\right) \tag{1.18}$$

puisque $\|\delta_j\|_1 = \int\delta_j = \int\gamma_j$, toutes ces fonctions étant positives. Ainsi, (1.17) et (1.18) donnent

$$|m(Z_j\cap Z_k)| \leqslant C\lambda^{-(k^3-j^3)}(C\lambda)^{K\log j} + Cm(Z_j)m(Z_k).$$

Finalement,

$$\begin{aligned}\sum_{j<k\leqslant n} m(Z_j\cap Z_k) &\leqslant C\sum_{j<k} m(Z_j)m(Z_k) + C\sum_{j=1}^{\infty}\lambda^{j^3}(C\lambda)^{K\log j}\sum_{k=j+1}^{\infty}\lambda^{-k^3}\\ &\leqslant C\left(\sum_{k\leqslant n} m(Z_k)\right)^2 + C\sum_{j=1}^{\infty}\lambda^{j^3}(C\lambda)^{K\log j}\sum_{l=(j+1)^3}^{\infty}\lambda^{-l}.\end{aligned}$$

La dernière somme est majorée par

$$C\sum_{j=1}^{\infty}\lambda^{j^3}(C\lambda)^{K\log j}\frac{\lambda^{-(j+1)^3}}{1-\lambda^{-1}} < \infty,$$

ce qui montre qu'on peut appliquer le lemme de Borel-Cantelli énoncé ci-dessus. Cela conclut la preuve du lemme. □

On fixe maintenant $x_0 \in X_1 \cap X_2 \cap X_3 \cap X_4$, ce qui est possible puisque cet ensemble est de mesure non nulle. Soient $m_k \to \infty$ telle que $x \in Z_{m_k}$, et $n_k = m_k^3 + \lfloor K \log m_k \rfloor - 1$. Alors $T^{n_k}(x_0)$ est dans un ensemble parmi $\alpha_1, \dots, \alpha_N$. En particulier, un de ces ensembles est utilisé une infinité de fois. Quitte à prendre une sous-suite de m_k, on peut supposer que $T^{n_k}(x_0)$ est toujours dans le même ensemble α_i, mettons α_1. On va montrer que u est lipschitzienne sur α_1. On notera $a_0, a_1, \dots$ la suite des éléments de α qui contiennent les itérés successifs de x_0.

Soit $\varepsilon > 0$. Comme $x_0 \in X_3$,

$$\frac{m\{y \in [a_0, \dots, a_{n_k-1}] \mid |u(y) - u(x_0)| > \varepsilon\}}{m[a_0, \dots, a_{n_k-1}]} \to 0.$$

Quitte à prendre encore une sous-suite de m_k, on peut même supposer que

$$\sum \frac{m\{y \in [a_0, \dots, a_{n_k-1}] \mid |u(y) - u(x_0)| > \varepsilon\}}{m[a_0, \dots, a_{n_k-1}]} < \infty.$$

Pour tout $k \in \mathbb{N}$, le contrôle de la distorsion implique que

$$\frac{m\{y \in X \mid \exists y' \in [a_0, \dots, a_{n_k-1}], T^{n_k} y' = y, |u(y') - u(x_0)| > \varepsilon\}}{m[Ta_{n_k-1}]} \asymp \frac{m\{y' \in [a_0, \dots, a_{n_k-1}] \mid |u(y') - u(x_0)| > \varepsilon\}}{m[a_0, \dots, a_{n_k-1}]}.$$

Ainsi, $\sum_k m\{y \in X \mid \exists y' \in [a_0, \dots, a_{n_k-1}], T^{n_k} y' = y, |u(y') - u(x_0)| > \varepsilon\} < +\infty$. Ainsi, $U_\varepsilon := \{y \in X \mid \exists \kappa, \forall k \geqslant \kappa, \text{ si } y' \in [a_0, \dots, a_{n_k-1}] \text{ est tel que } T^{n_k} y' = y, \text{ alors } |u(y') - u(x_0)| \leqslant \varepsilon\}$ est de mesure pleine.

Soient y_1, y_2 dans $U_\varepsilon \cap \alpha_1$. Si k est assez grand, les préimages y_1' et y_2' de y_1 et y_2 dans $[a_0, \dots, a_{n_k-1}]$ satisfont $|u(y_i') - u(x_0)| \leqslant \varepsilon$, si bien que $|u(y_1') - u(y_2')| \leqslant 2\varepsilon$. Alors

$$\begin{aligned} |u(y_1) - u(y_2)| &= |u \circ T^{n_k}(y_1') - u \circ T^{n_k}(y_2')| \\ &\leqslant \sum_{i=0}^{n_k-1} |f \circ T^i(y_1') - f \circ T^i(y_2')| + |u(y_1') - u(y_2')|. \end{aligned} \tag{1.19}$$

Rappelons que $n_k = m_k^3 + \lfloor K \log m_k \rfloor - 1$, et que Φ est définie par $\Phi(x) = Df(a)$ lorsque $x \in a$. Alors

$$\begin{aligned} \sum_{i=0}^{m_k^3-1} |f \circ T^i(y_1') - f \circ T^i(y_2')| &\leqslant \sum_{i=0}^{m_k^3-1} \Phi(T^i(x_0)) d(T^i y_1', T^i y_2') \\ &\leqslant \sum_{i=0}^{m_k^3-1} \Phi(T^i(x_0)) \lambda^{i-n_k} d(T^{n_k} y_1', T^{n_k} y_2') \\ &\leqslant \lambda^{-K \log m_k + 2} S_{m_k^3} \Phi(x_0) d(y_1, y_2). \end{aligned} \tag{1.20}$$

Mais $x_0 \in X_1$, si bien qu'il existe C tel que $S_n\Phi(x_0) \leqslant Cn$ pour tout n. Comme $-K\log\lambda < -3$ par (1.16), on obtient que (1.20) tend vers 0.

Finalement, soit $D = \sup Df(\alpha_j)$ pour $1 \leqslant j \leqslant N$. Par définition de la suite m_k, on a $T^i(x_0) \in \bigcup_{j=1}^N \alpha_j$ pour tout $m_k^3 \leqslant i < n_k$, si bien que

$$\begin{aligned}\sum_{i=m_k^3}^{n_k-1} |f\circ T^i(y_1') - f\circ T^i(y_2')| &\leqslant \sum_{i=m_k^3}^{n_k-1} Dd(T^iy_1', T^iy_2') \leqslant D\sum_{i=m_k^3}^{n_k-1} \lambda^{i-n_k} d(y_1,y_2)\\ &\leqslant \frac{D}{\lambda-1} d(y_1,y_2).\end{aligned}$$

Ainsi, l'équation (1.19) donne

$$|u(y_1) - u(y_2)| \leqslant o(1) + \frac{D}{\lambda-1} d(y_1,y_2) + 2\varepsilon.$$

Finalement, sur $\alpha_1 \cap \bigcap U_\varepsilon$, de mesure pleine dans α_1, on a $|u(y_1)-u(y_2)| \leqslant \frac{D}{\lambda-1} d(y_1,y_2)$. □

Lemme 1.2.4. *On a $\sum_{a\in\alpha} m[a]Du(a) < \infty$.*

Démonstration. Soit $a \in \alpha$, montrons que $Du(a) < \infty$. Comme T est transitive, il existe n tel que $a \subset T^n(\alpha_1)$, et soit alors $[a_0,\dots,a_{n-1}]$ un cylindre inclus dans α_1 (avec $a_0 = \alpha_1$) avec $a \subset T(a_{n-1})$. Pour $y_1, y_2 \in a$, soient y_1' et y_2' leurs antécédents par T^n dans $[a_0,\dots,a_{n-1}]$. Alors

$$\begin{aligned}|u(y_1) - u(y_2)| &\leqslant \sum_{i=0}^{n-1} |f(T^iy_1') - f(T^iy_2')| + |u(y_1') - u(y_2')|\\ &\leqslant \sum_{i=0}^{n-1} Df(a_i)\lambda^{i-n} d(y_1,y_2) + \lambda^{-n} Du(\alpha_1) d(y_1,y_2),\end{aligned} \tag{1.21}$$

ce qui prouve que $Du(a) < \infty$.

Soit maintenant β une partie finie non vide de α. Soit $a \in \alpha\backslash\beta$, montrons que

$$m(a) = \sum_{a_0\in\beta, a_1,\dots,a_{n-1}\in\alpha\backslash\beta} m[a_0,a_1,\dots,a_{n-1},a]. \tag{1.22}$$

Soit $Y = \bigcup_{b\in\beta} b$. On notera $A_0 = a$, puis $A_{n+1} = T^{-1}(A_n)\backslash Y$ et $B_{n+1} = T^{-1}(A_n)\cap Y$. Ainsi, on a

$$A_n = \bigcup_{a_0,\dots,a_{n-1}\in\alpha\backslash\beta} [a_0,\dots,a_{n-1},a] \text{ et } B_n = \bigcup_{a_0\in\beta,a_1,\dots,a_{n-1}\in\alpha\backslash\beta} [a_0,\dots,a_{n-1},a].$$

On veut donc montrer que $m(a) = \sum_n m(B_n)$. On a $T^{-1}(A_n) = A_{n+1}\cup B_{n+1}$, donc $m(A_n) = m(A_{n+1}) + m(B_{n+1})$. Par récurrence, on obtient $m(a) = m(B_1) + \dots + m(B_n) + m(A_n)$. Pour

conclure, il faut donc voir que $m(A_n) \to 0$. Mais $A_n \subset C_n = \{x \mid \forall 0 \leqslant k \leqslant n, T^k(x) \notin Y\}$. On va montrer que $m(C_n) \to 0$. Il suffit pour cela de voir que $C = \bigcap C_n$ est de mesure nulle. On a $C \subset T^{-1}(C)$, donc par invariance de la mesure $C = T^{-1}(C) \mod 0$, donc $m(C) = 0$ ou 1 par ergodicité (car T est transitive). Mais C ne rencontre pas Y, de mesure positive, donc $m(C) = 0$. Cela démontre (1.22).

Si $[a_0, \dots, a_{n-1}, a]$ est un cylindre de mesure positive, on a d'après (1.21)

$$Du(a) \leqslant \sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) + \lambda^{-n} Du(a_0).$$

Ainsi, (1.22) implique que

$$m(a)Du(a) \leqslant \sum_{a_0 \in \beta, a_1, \dots, a_{n-1} \in \alpha \backslash \beta} m[a_0, \dots, a_{n-1}, a] \left(\sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) \right) + m(a) \sup_{b \in \beta} Du(b).$$

Comme $\sum m(a) \sup_{b\in\beta} Du(b) < \infty$, on va montrer que $\sum m(a)Du(a) < \infty$ en vérifiant que

$$\sum_{n=1}^{\infty} \sum_{a_0 \in \beta, a_1, \dots, a_{n-1} \in \alpha \backslash \beta} m[a_0, \dots, a_{n-1}] \left(\sum_{i=0}^{n-1} \lambda^{i-n} Df(a_i) \right) \tag{1.23}$$

est fini. Dans cette expression, pour $a \in \alpha \backslash \beta$, le facteur d'un terme $\lambda^{-k} Df(a)$ est

$$\sum_{\substack{a_0 \in \beta, a_1, \dots, a_{n-1} \in \alpha \backslash \beta \\ a_{n+1}, \dots, a_{n+k-1} \in \alpha \backslash \beta}} m[a_0, \dots, a_{n-1}, a, a_{n+1}, \dots, a_{n+k-1}] \leqslant \sum_{a_0 \in \beta, a_1, \dots, a_{n-1} \in \alpha \backslash \beta} m[a_0, \dots, a_{n-1}, a].$$

Mais, d'après (1.22), ce dernier terme est égal à $m(a)$. On vérifie de même que, dans (1.23), le facteur d'un terme $\lambda^{-k} Df(a)$ avec $a \in \beta$ est majoré par $m(a)$. Ainsi,

$$(1.23) \leqslant \sum_{a \in \alpha} \sum_{k=1}^{\infty} m(a) \lambda^{-k} Df(a),$$

qui est fini puisque $\sum m(a)Df(a) < \infty$ par hypothèse. □

Démonstration du théorème 1.2.1. Pour presque tout x, la conservation de la mesure implique que $\sum_{Ty=x} g(y) = 1$. Notons $x_0, x_1, \dots$ les antécédents de x, qui sont dans des éléments $a_0, a_1, \dots$ de la partition α. Par distorsion bornée, on a $g(x_n) \leqslant B \frac{m(a_n)}{m(Ta_n)}$ pour une certaine constante B. Comme T a la propriété de grande image, $m(Ta_n)$ est uniformément minoré, si bien qu'il existe $B' > 0$ tel que $g(x_n) \leqslant B' m(a_n)$. Comme $\sum g(x_n) = 1$, on obtient $B' \sum m(a_n) \geqslant 1$.

Soit a_* un élément de α_*, majorons $Du(a_*)$. Soient $x, y \in a_*$. Leurs antécédents par T, qu'on notera $x_0, x_1, \dots$ et $y_0, y_1, \dots$, sont dans les mêmes éléments $a_0, a_1, \dots$ de α, par définition de α_*. Comme $f = u - u \circ T$, on a pour tout n

$$\begin{aligned}|u(x) - u(y)| &\leqslant |f(x_n) - f(y_n)| + |u(x_n) - u(y_n)| \leqslant (Df(a_n) + Du(a_n))d(x_n, y_n) \\ &\leqslant (Df(a_n) + Du(a_n))\lambda^{-1} d(x, y).\end{aligned}$$

Ainsi,

$$|u(x) - u(y)| \leqslant B' \sum m(a_n)|u(x) - u(y)| \leqslant B' \sum m(a_n)(Df(a_n) + Du(a_n))\lambda^{-1} d(x, y).$$

Finalement, $Du(a_*) \leqslant \frac{B'}{\lambda} \sum_a m(a)(Df(a) + Du(a))$, qui est fini par le lemme 1.2.4. □

1.2.2 Le cas intégrable

Le critère suivant peut également être utile pour assurer qu'une fonction n'est pas un cobord : il affirme qu'un cobord intégrable est nécessairement borné (sous des hypothèses de régularité suffisantes).

Proposition 1.2.5. *Sous les hypothèses du théorème 1.2.1, supposons que f est intégrable. Alors u est essentiellement bornée.*

Démonstration. Montrons tout d'abord que u est intégrable. Soient $a_1, \dots, a_N$ des éléments de α dont la mesure dépasse $1 - \eta$, si bien que l'image de tout élément de α contient l'un des a_i. Comme u est lipschitzienne sur les a_i, elle y est en particulier bornée par une constante C.

Soit $x \in a \in \alpha$. Il existe $y \in a$ dont l'image est dans l'un des a_i. Alors $u(y) = f(y) + u(Ty)$ donc $|u(y)| \leqslant |f(y)| + C$, puis

$$|u(x)| \leqslant Du(a) + |u(y)| \leqslant Du(a) + C + |f(y)| \leqslant Du(a) + C + Df(a) + |f(x)|.$$

Le membre de droite est intégrable, d'après le lemme 1.2.4, donc u est également intégrable.

Montrons que u est en fait bornée. Soit $x \in X$, dont les antécédents $x_0, x_1, \dots$ sont dans des éléments $a_0, a_1, \dots$ de α, avec $B' \sum m(a_n) \geqslant 1$ d'après la preuve du théorème 1.2.1. Alors, pour tout n, $u(x) = u(x_n) - f(x_n)$. De plus, pour $y \in a_n$, $|u(y) - u(x_n)| \leqslant Du(a_n)$, donc en intégrant $|u(x_n)| \leqslant Du(a_n) + \frac{1}{m(a_n)} \int_{a_n} |u|$. La fonction f satisfait la même équation. Ainsi,

$$\begin{aligned}|u(x)| &\leqslant B' \sum m(a_n)|u(x)| \leqslant B' \sum m(a_n)(|u(x_n)| + |f(x_n)|) \\ &\leqslant B' \left(\sum m(a_n)(Du(a_n) + Df(a_n)) + \int_{a_n} (|u| + |f|) \right).\end{aligned}$$

Ce dernier terme est majoré par $B' \sum m(a)(Du(a) + Df(a)) + B' \int_X (|u| + |f|)$, qui est fini. □

Remarque (post-soutenance) : l'hypothèse d'intégrabilité de f est en fait inutile, la fonction u est toujours essentiellement bornée. En effet, d'après [Sar03], une application Gibbs-Markov préservant une mesure de probabilité a automatiquement la *big preimage property* : il existe un nombre fini d'éléments de partition $a_1, \dots, a_n$ tels que, pour tout $a \in \alpha$, il existe i avec $a \subset T(a_i)$ (modulo 0). Si $f = u - u \circ T$, les fonctions f et u sont hölderiennes, donc bornées, sur chacun des a_i, et cela implique que u est bornée sur tout l'espace.

Pour démontrer le résultat de Sarig, on note que la distorsion bornée implique l'existence d'une constante $C > 0$ telle que, pour tout cylindre $[b, a]$ non vide,

$$C^{-1} m[b] m[a] \leqslant m[b, a] \leqslant C m[b] m[a].$$

Soit $a \in \alpha$, notons $b_1, \dots$ les éléments de α tels que $a \subset T(b_j)$. Alors $T^{-1}[a] = \bigcup [b_j, a]$, puis

$$m[a] = m(T^{-1}[a]) = \sum m[b_j, a] \leqslant C \sum m[b_j] m[a].$$

On obtient donc $\sum m[b_j] \geqslant C^{-1}$. Pour conclure, il suffit donc de prendre des éléments $a_1, \dots, a_n$ de α tels que $\sum m[a_j] > 1 - C^{-1}$. En effet, l'un des b_j sera alors nécessairement égal à l'un des a_i (sans quoi on aurait $m(X) > 1 - C^{-1} + C^{-1} = 1$, ce qui est absurde).

1.3 Applications induites et constructions de tours

Dans la suite, on va s'intéresser à des applications qui ne seront pas Gibbs-Markov, à cause de singularités en lesquelles on n'aura pas de dilatation uniforme. Pour pouvoir malgré tout utiliser les outils disponibles pour les applications Gibbs-Markov, l'idée (classique) sera de travailler sur une partie de l'espace qui ne verra pas les singularités, et d'utiliser les premiers retours des points dans cette partie : c'est le processus *d'induction*.

1.3.1 Induction

Les preuves des résultats de cette partie se trouvent dans [Aar97, partie 1.5].

Définition 1.3.1. *Soient $T : X \to X$ une application conservative sur un espace X muni d'une mesure m, et Y une partie de X de mesure non nulle. Pour $y \in Y$, soit $\varphi_Y(y) = \inf\{n \geqslant 1 \mid T^n(y) \in Y\}$ le* temps de retour *de y dans Y. On définit une application T_Y sur l'ensemble $\{y \in Y \mid \varphi_Y(y) < \infty\}$ par*

$$T_Y(y) = T^{\varphi_Y(y)}(y). \tag{1.24}$$

C'est l'application induite par T sur Y.

L'hypothèse de conservativité ([Aar97, partie 1.1]) assure que presque tout point de Y revient une infinité de fois dans Y, ce qui va permettre d'itérer T_Y. Elle est satisfaite en particulier dès que T préserve la mesure m et que celle-ci est de masse finie.

On a une correspondance entre les mesures invariantes pour T et pour T_Y :

Proposition 1.3.2. *Si T préserve une mesure de probabilité m, alors T_Y préserve la restriction m_Y de cette mesure à Y, et $m = \sum_{n=0}^{\infty} T_*^n(m_Y|\{\varphi_Y > n\})$.*

Réciproquement, si T_Y préserve une mesure de probabilité m_Y, alors T préserve la mesure

$$m = \sum_{n=0}^{\infty} T_*^n(m_Y|\{\varphi_Y > n\}).$$

Il faut noter que la mesure m donnée par cette dernière formule n'est pas toujours une mesure de masse finie : il est donc important d'étudier la proportion des points qui mettent longtemps à revenir, i.e. la *taille des queues* $m_Y\{\varphi_Y > n\}$.

On a un lien entre l'ergodicité de T et de T_Y :

Proposition 1.3.3. *Supposons que m est une mesure de probabilité invariante et ergodique pour T. Alors $m_{|Y}$ est ergodique pour T_Y.*

Réciproquement, si m_Y est invariante et ergodique pour T_Y, soit $m = \sum_{n=0}^{\infty} T_*^n(m_{|\{\varphi_Y > n\}})$. Alors m est ergodique pour T.

De plus, on peut estimer le nombre moyen d'itérés nécessaires pour revenir à Y : si T préserve une mesure de probabilité ergodique m, alors un point passe en moyenne une proportion du temps $1/m(Y)$ dans Y, par le théorème de Birkhoff, donc le temps de retour doit en moyenne être égal à $1/m(Y)$. On a effectivement la proposition suivante (*formule de Kac*) :

Proposition 1.3.4. *Si T préserve une mesure de probabilité m, et Y est une partie de X de mesure non nulle, alors*

$$\int_Y \varphi_Y \, \mathrm{d}m = 1.$$

1.3.2 Le processus inverse : construction de tours

On va décrire le processus inverse de celui du paragraphe précédent, qui va consister à partir d'une transformation S sur un espace Y et d'une fonction à valeurs entières qui représentera le temps de retour, et à construire une extension (dont l'application induite sur Y sera exactement S).

Plus précisément, soient S une transformation d'un espace probabilisé (Y, μ) et $\varphi : Y \to \mathbb{N}^*$ une fonction mesurable. On définit un espace $X = \{(y, i) \mid y \in Y, i < \varphi(y)\}$, et une application T sur X par $T(y, i) = (y, i + 1)$ si $i < \varphi(y) - 1$, et $T(y, \varphi(y) - 1) = (S(y), 0)$. On définit

également une mesure m sur X, en partant de μ dans la base et en la poussant vers le haut, i.e. $m|Y \times \{i\} = \mu$.

Dans X, on identifiera $Y \times \{0\}$ avec Y. La proposition qui suit est alors une conséquence directe des définitions.

Proposition 1.3.5. *L'application T_Y induite par T sur Y est S.*

Un intérêt de l'application T ainsi définie est que deux points distincts de la base Y gardent des images distinctes avant le premier retour à la base.

Supposons maintenant que l'on parte d'une application T conservative sur un espace probabilisé (X, m). Soit $Y \subset X$ une partie de mesure non nulle. Soit $Y_1, Y_2, \ldots$ une partition finie ou dénombrable de Y (à un ensemble de mesure nulle près) et $\varphi : Y \to \mathbb{N}^*$ une fonction constante sur chacun des ensembles Y_n, telle que $T^{\varphi(Y_n)}(Y_n) \subset Y$. On peut alors faire la même construction que ci-dessus, avec l'application $S = T^\varphi$, le temps de retour φ, et la mesure $\mu = m_{|Y}$. On obtient ainsi un espace X', une application $T' : X' \to X'$ et une mesure m' sur X'.

Par exemple, on peut prendre pour φ le temps de premier retour à Y. Même en ce cas, l'application T' n'est pas toujours conjuguée à T : il se peut que deux points différents de Y soient envoyés par un itéré de T sur le même point de X avant le premier retour à Y, tandis que ce ne sera pas le cas pour T' : par construction, deux points distincts de Y ont des itérés par T' qui restent distincts avant le premier retour à Y. Ainsi, (X', T') peut être vu comme une extension de (X, T), qui sépare les orbites avant le premier retour à Y.

Pour préciser ceci, on introduit une application $\pi : X' \to X$ en posant $\pi(y, i) = T^i(y)$.

Proposition 1.3.6. *On a un diagramme commutatif*

$$\begin{array}{ccc} X' & \xrightarrow{T'} & X' \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ X & \xrightarrow{T} & X \end{array}$$

Démonstration. Soit $(y, i) \in X'$. Si $i < \varphi(y) - 1$, alors $T'(y, i) = (y, i + 1)$, si bien que $\pi \circ T'(y, i) = \pi(y, i + 1) = T^{i+1}(y) = T \circ \pi(y, i)$. Supposons maintenant que $i = \varphi(y) - 1$, si bien que $T'(y, i) = (T^{\varphi(y)}(y), 0)$. Ainsi,

$$\pi \circ T'(y, \varphi(y) - 1) = T^{\varphi(y)}(y) = T(T^{\varphi(y)-1}(y)) = T \circ \pi(y, \varphi(y) - 1). \qquad \square$$

Ce diagramme implique en particulier que, si T' préserve une mesure ν', alors T préserve la mesure $\nu = \pi_*(\nu')$. On a une réciproque partielle dans le cas où φ est le temps de premier retour.

Proposition 1.3.7. *Supposons que T préserve la mesure de probabilité ergodique m, et que φ est le temps de premier retour à Y. Alors $\pi_*(m') = m$, et T' préserve la mesure m'.*

Démonstration. Si T préserve m, alors $T^\varphi = T_Y$ préserve la mesure $\mu = m_Y$, d'après la première implication de la proposition 1.3.2. Mais $T'_Y = T_Y$, donc T'_Y préserve μ. Ainsi, T' préserve la mesure $\sum_0^\infty (T')^n_*(\mu|\{\varphi > n\})$, c'est-à-dire m'.

Montrons maintenant que $\pi_*(m') = m$. Soit $K \subset X$, et montrons que $m'(\pi^{-1}K) = m(K)$.

Soient $A_0 = K\backslash Y$ et $B_0 = K \cap Y$. On définit par récurrence $A_k = T^{-1}(A_{k-1})\backslash Y$ et $B_k = T^{-1}(A_{k-1}) \cap Y$: les points de A_k n'ont pas encore été traités au temps k, alors que les points de B_k sont traités exactement au temps k. On a alors $\pi^{-1}(K) = \{(x, i) \mid i \in \mathbb{N}, x \in B_i\}$. Ainsi, $m'(\pi^{-1}K) = \sum m(B_k)$.

Comme $m(K) = m(A_0) + m(B_0)$ et, par construction, $m(A_{k-1}) = m(T^{-1}A_{k-1}) = m(A_k) + m(B_k)$, on obtient pour tout n que $m(K) = m(B_0) + \ldots + m(B_n) + m(A_n)$. Pour conclure, il suffit de montrer que $m(A_n) \to 0$ quand $n \to \infty$.

Soit $C_n = \{x \mid \forall i \leqslant n, T^i x \notin Y\}$. Alors $A_n \subset C_n$. De plus, C_n décroît. Soit $C = \bigcap C_n$. Il suffit de montrer que $m(C) = 0$, car cela impliquera que $m(C_n) \to 0$. Comme $C \subset T^{-1}C$ et T préserve la mesure, on a $C = T^{-1}C \mod 0$. Par ergodicité de T, $m(C) = 0$ ou 1. Mais $C \cap Y = \emptyset$ et $m(Y) > 0$, donc $m(C) = 0$. □

Ceci implique que, si on part d'une application ergodique T qui préserve une mesure m mais qui peut avoir un comportement mal contrôlé entre deux retours successifs à une partie Y de X, alors on peut la comprendre en étudiant une application plus simple qui aura les mêmes propriétés ergodiques que T mais sera en plus gentille entre deux retours à Y. Par exemple, si on s'intéresse à des théorèmes limites pour une fonction $f : X \to \mathbb{R}$ sous l'itération de T, il suffira de démontrer les mêmes théorèmes pour $f' = f \circ \pi : X' \to \mathbb{R}$ sous l'action de T', ce qui se révélera plus facile.

1.3.3 Applications induisant des applications Gibbs-Markov

On va appliquer les techniques qui précèdent à des applications qui induisent des applications Gibbs-Markov sur une partie de l'espace.

Soient (X, T, m, α, d) une application markovienne, et Y une réunion d'éléments de α. Montrons que l'application induite T_Y est automatiquement markovienne. Soit $\gamma \subset \alpha$ tel que $Y = \bigcup_{a\in\gamma} a$. Pour $a \in \gamma$ et $\xi_1, \ldots, \xi_{n-1} \in \alpha\backslash\gamma$, notons

$$[a, \xi_1, \ldots, \xi_{n-1}, Y] = \{y \in a \mid \forall 1 \leqslant i \leqslant n-1, T^i(y) \in \xi_i, \text{ et } T^n y \in Y\}.$$

C'est une réunion de cylindres de T. Notons

$$\delta = \{[a, \xi_1, \ldots, \xi_{n-1}, Y] \mid a \in \gamma, \xi_1, \ldots, \xi_{n-1} \in \alpha\backslash\gamma, m[a, \xi_1, \ldots, \xi_{n-1}, Y] > 0\}.$$

C'est une partition (modulo un ensemble de mesure nulle) de l'ensemble des points de Y qui reviennent dans Y. En particulier, lorsque T est conservative, cet ensemble est de mesure

pleine. L'application T_Y est alors une application de Y dans lui-même définie presque partout, markovienne pour la partition δ.

On notera les cylindres de la partition δ par $[d_0, \dots, d_{n-1}]_Y$, avec $d_0, \dots, d_{n-1} \in \delta$. Si $d = [a, \xi_1, \dots, \xi_{n-1}, Y] \in \delta$, son image par T_Y est $T_Y(d) = T(\xi_{n-1}) \cap Y$, et est donc α-mesurable. En particulier, si γ est fini, alors T_Y a automatiquement la propriété de grande image.

Définition 1.3.8. *Soit (X, T, m, α, d) une application markovienne conservative. On dit qu'elle est localement Gibbs-Markov si elle satisfait les deux conditions suivantes :*
- *Il existe une partie Y de X, réunion d'éléments de α, avec $\bigcup T^{-n}(Y) = X \mod 0$, telle que T_Y soit Gibbs-Markov pour la partition δ correspondante sur Y.*
- *Il existe $C > 0$ vérifiant la condition qui suit. Soient $x, y \in X$, notons $x_0 = x, x_1 = T(x), \dots, x_n = T^n(x) \in Y$ et $y_0 = y, y_1 = T(y), \dots, y_m = T^m(y) \in Y$ les premiers itérés de ces points, jusqu'à la première entrée dans Y. On suppose que $m = n$, et que pour $0 \leqslant i < n$, x_i et y_i sont dans le même élément de α. Alors, pour presque tous x, y vérifiant cette propriété, $d(x, y) \leqslant C d(x_n, y_n)$.*

La deuxième condition affirme que la distance ne se comporte pas trop mal vis-à-vis de l'induction, et permet d'assurer que des fonctions lipschitziennes sur tout l'espace induisent des fonctions assez régulières sur Y.

Remarque 1.3.9. Quand l'espace X n'est pas muni d'une distance au départ, ou quand cette distance ne vérifie pas la deuxième condition, on peut définir une distance d sur X de la façon suivante : soient $s(x, y)$ le temps de séparation de $x, y \in Y$ sous l'action de T_Y, et $\tau < 1$. On pose alors, pour $x, y \in X$, $d(x, y) = 1$ si x et y ne restent pas dans le même élément de la partition α avant leur première entrée dans Y, et $d_\tau(x, y) = \tau^{s(x_n, y_n)}$ sinon, où x_n et y_n sont les premières entrées de x et y dans Y. Cette distance est une extension à X de la distance définie en (1.1). Si T_Y est Gibbs-Markov pour une certaine distance, alors T est localement Gibbs-Markov pour la distance d_τ si τ est assez proche de 1, d'après la remarque 1.1.5.

On peut aisément relier les propriétés de transitivité de T et T_Y (mais il n'y a pas d'analogue pour les propriétés de mélange) :

Lemme 1.3.10. *Si T est Markov-transitive, alors T_Y est Markov-transitive.*

On peut alors déduire des résultats sur les mesures invariantes des applications localement Gibbs-Markov, en utilisant les résultats correspondants pour les applications Gibbs-Markov.

Proposition 1.3.11. *Soit (X, T, m) une application localement Gibbs-Markov. Il existe alors un nombre fini non nul de mesures $m_1, \dots, m_N$ invariantes, absolument continues par rapport à m, ergodiques. Elles satisfont $m_i(Y) < \infty$. Il existe également des ensembles $X_1, \dots, X_N$ tels que m_i est équivalente à m sur X_i, avec $X = \bigcup_{i=1}^N \bigcup_{n \geqslant 0} T^{-n}(X_i)$. De plus, toute mesure μ invariante absolument continue est combinaison linéaire des m_i.*

Enfin, si T est Markov-transitive, alors $N = 1$, i.e. il existe une unique mesure invariante absolument continue μ, et cette mesure est ergodique et satisfait $\mu(Y) < \infty$.

Démonstration. Si T est localement Gibbs-Markov pour une partie Y de X, alors T_Y préserve un nombre fini non nul de mesures de probabilité $\mu_1, \dots, \mu_N$ invariantes absolument continues par rapport à m, par la proposition 1.1.17 et le théorème 1.1.8. De plus, si T est transitive, alors T_Y est transitive par le lemme 1.3.10, donc $N = 1$ par le théorème 1.1.8.

Soient $m_1, \dots, m_N$ obtenues à partir de $\mu_1, \dots, \mu_N$ par le processus d'induction décrit dans la proposition 1.3.2. Comme les μ_i sont ergodiques, les mesures m_i sont également ergodiques d'après la proposition 1.3.3.

Soit Y_i tel que μ_i est équivalente à m sur Y_i. Alors m_i est équivalente à m sur $X_i = \bigcup T^n(Y_i)$. D'après le corollaire 1.1.18, $\bigcup_{i=1}^N \bigcup_{n\geqslant 0} T_Y^{-n}(Y_i) = Y \mod 0$. Comme $\bigcup T^{-n}(Y) = X$, on en déduit que $\bigcup_{i=1}^N \bigcup_{n\geqslant 0} T^{-n}(X_i) = X \mod 0$.

Soit finalement μ une mesure invariante non nulle, absolument continue par rapport à m. Alors la restriction de μ à Y est invariante par T_Y et absolument continue par rapport à m. Par le corollaire 1.1.18, $\mu_{|Y}$ est combinaison linéaire des μ_i. En induisant (et en utilisant $X = \bigcup T^{-n}(Y)$), on en déduit que μ est combinaison linéaire des m_i. □

Cependant, les mesures m_i obtenues de cette façon ne sont pas nécessairement de masse finie : pour l'assurer, il faut de plus imposer une condition de sommabilité.

On peut également définir les tours localement Gibbs-Markov, qui sont des applications localement Gibbs-Markov avec une combinatoire simple entre deux retours successifs :

Définition 1.3.12. *On dit qu'une application markovienne conservative (X, T, m, α, d) est une* tour localement Gibbs-Markov *s'il existe une partie α_0 de α et une application $\varphi : \alpha_0 \to \mathbb{N}^*$ (le* temps de retour*) telles que*

1. *$X = \{(x, i) \mid x \in a \in \alpha_0, i \in \{0, \dots, \varphi(a) - 1\}\}$, et la partition α est donnée par $\{a \times \{i\} \mid a \in \alpha_0, i \leqslant \varphi(a) - 1\}$.*
2. *Pour $a \in \alpha_0$ et $i < \varphi(a) - 1$, T envoie $a \times \{i\}$ dans $a \times \{i + 1\}$ en agissant comme l'identité sur la première coordonnée (i.e. $T(x, i) = (x, i+1)$) et $T_* m_{|a\times\{i\}}$ est équivalente à $m_{|a\times\{i+1\}}$.*
3. *Pour $a \in \alpha_0$, $T^{\varphi(a)}$ envoie $a \times \{0\}$ sur une réunion d'éléments de α_0, avec une image de mesure uniformément minorée et une dilatation d'au moins $\lambda > 1$. De plus, il existe $C > 0$ tel que la distorsion $g(x) = \frac{\mathrm{d}m}{\mathrm{d}\left(m \circ T_{|a}^{\varphi(a)}\right)}$ vérifie : pour presque tous $x, y \in a \times \{0\}$,*

$$\left|1 - \frac{g(x)}{g(y)}\right| \leqslant C d(T^{\varphi(a)}x, T^{\varphi(a)}y).$$

4. *Il existe $C > 0$ tel que, pour tout $a \in \alpha_0$, pour tout $i < \varphi(a)$, pour presque tous $x, y \in a \times \{0\}$,*
$$d(T^i x, T^i y) \leqslant C d(T^{\varphi(a)} x, T^{\varphi(a)} y).$$

En particulier, une tour localement Gibbs-Markov est une application localement Gibbs-Markov, par définition, mais les retours à la base vérifient également une combinatoire très simple : deux points distincts de la base ne peuvent pas être envoyés sur un même point avant le premier retour à la base. Ainsi, il sera souvent plus facile de démontrer des théorèmes pour des tours localement Gibbs-Markov que pour des applications localement Gibbs-Markov générales. On pourra alors tirer profit du résultat suivant :

Proposition 1.3.13. *Soit (X, T, m, α, d) une application localement Gibbs-Markov sur une base $Y \subset X$, préservant la mesure de probabilité m. Il existe alors une tour localement Gibbs-Markov $(X', T', m', \alpha', d')$ préservant une mesure de probabilité m', et une projection $\pi : X' \to X$, telles que $\pi_*(m') = m$, $\pi \circ T' = T \circ \pi$, et la restriction de la projection π à un élément a' de α' est une isométrie, dont l'image est incluse dans un élément de α. Finalement, π est un isomorphisme entre (Y', α') et (Y, δ).*

Démonstration. Si T est localement Gibbs-Markov, on peut lui appliquer la construction décrite dans le paragraphe 1.3.2, en prenant pour φ le temps de premier retour. On définit une partition α' sur X', en poussant δ dans chaque étage de la tour. Plus précisément, α' est donnée par $\{d \times \{k\} \mid d \in \delta, k < \varphi(d)\}$. Notons $\Delta_k = \{(y, k) \mid y \in Y, \varphi(y) > k\}$ le k-ème étage de la tour. On définit aussi une distance d' sur X', de la façon suivante : pour (x, k) et (y, k) dans le même étage Δ_k de la tour, on pose $d'((x, k), (y, k)) = d(T^k x, T^k y)$. De plus, pour $k \neq l$, on pose $d'((x, k), (y, l)) = \sup_{X \times X} d$. Par construction, l'application T' est alors markovienne pour la partition α' et la distance d', et l'application induite sur la base de la tour est T_Y. Finalement, T' vérifie clairement toutes les propriétés annoncées. □

Ainsi, pour étudier les applications localement Gibbs-Markov, on peut se restreindre au cas transitif, puis aux tours localement Gibbs-Markov transitives. On poursuit la réduction dans la proposition qui suit :

Proposition 1.3.14. *Soit (X, T, m) une tour localement Gibbs-Markov transitive. Il existe alors un entier d et une partition $X_0, \dots, X_{d-1}$ de X (moins fine que α) telle que T envoie X_i dans X_{i+1} pour $0 \leqslant i < d-1$, et X_{d-1} dans X_0. De plus, pour $0 \leqslant i \leqslant d-1$, $(X_i, T^d, m_{|X_i})$ est une tour localement Gibbs-Markov mélangeante.*

Démonstration. Pour $a \in \alpha$, soit $d = \mathrm{pgcd}\{n \mid a \subset T^n(a)\}$. Par transitivité, on vérifie qu'il est indépendant de a. Fixons alors $a \in \alpha$, et posons, pour $0 \leqslant i \leqslant d-1$, $X_i = \{T^n(a) \mid n \equiv i \mod d\}$. Ces ensembles forment une partition de X, et T^d est Markov-mélangeante sur chacun d'eux (voir [Aar97, proposition 4.2.2] pour plus de détails). Finalement, on vérifie aisément que (X_i, T^d) est une tour localement Gibbs-Markov, de base $X_i \cap \bigcup_0^{d-1} \Delta_j$, où Δ_j désigne les points de la tour d'altitude j. □

On a l'analogue suivant du corollaire 1.1.21 :

Proposition 1.3.15. *Soit (X, T, m) une application localement Gibbs-Markov mélangeante. Elle est alors exacte : si $A \subset X$ s'écrit sous la forme $T^{-n}(A_n) \mod 0$ pour tout n, alors $m(A) = 0$ ou 1.*

En particulier, si T préserve une mesure de probabilité absolument continue par rapport à m, alors cette mesure est mélangeante.

Démonstration. La preuve du théorème 4.4.7 de [Aar97] s'applique directement à ce cas. □

Pour montrer l'utilité de l'extension donnée par la proposition 1.3.13, on va démontrer l'analogue suivant du théorème 1.2.1. Rappelons que $Df(a)$ est défini dans (1.10).

Théorème 1.3.16. *Soient (X, T, m, α) une application localement Gibbs-Markov qui préserve une mesure de probabilité m, et G un groupe abélien localement compact polonais. Soit $f : X \to G$ satisfaisant $\sum_{a\in\alpha} m[a]Df(a) < +\infty$. Soit $u : X \to G$ une application mesurable telle que $f = u - u \circ T$ presque partout.*

Alors il existe $C > 0$ tel que, pour tout $d \in \delta$ inclus dans Y, pour presque tous $x, y \in d$, $|u(x) - u(y)| \leqslant Cd(T^{\varphi(d)}x, T^{\varphi(d)}y)$.

Autrement dit, une solution de l'équation cohomologique est nécessairement régulière sur Y. On peut ensuite utiliser l'équation $f = u - u \circ T$ pour en déduire que u est régulière sur tout l'espace X.

En particulier, ce résultat s'applique aux fonctions lipschitziennes.

Démonstration. Montrons tout d'abord qu'il suffit de démontrer le résultat dans une tour pour conclure.

Soient X' l'extension de X en forme de tour décrite dans la proposition 1.3.13, $\pi : X' \to X$ la projection canonique et $f' = f \circ \pi$. Un élément a' de α' se projette sur une partie d'un élément a de α, et on a alors $Df'(a') \leqslant Df(a)$. En particulier, pour $a \in \alpha$,

$$\sum_{\substack{a'\in\alpha' \\ \pi(a')\subset a}} m'(a')Df'(a') \leqslant \sum_{\substack{a'\in\alpha' \\ \pi(a')\subset a}} m'(a')Df(a) = m'(\pi^{-1}(a))Df(a) = m(a)Df(a).$$

Ainsi, $\sum m'(a')Df'(a') \leqslant \sum m(a)Df(a) < \infty$. Les hypothèses du théorème sont donc vérifiées pour X', f' et $u' = u \circ \pi$. Il suffit alors de montrer le théorème dans ce cadre pour conclure.

On peut donc supposer que X est une tour localement Gibbs-Markov, c'est-à-dire que $X = \{(y, i) \mid y \in Y, i < \varphi(y)\}$. Notons f_Y la fonction définie sur Y par $f_Y(x) = \sum_0^{\varphi(x)-1} f(T^k x)$. Elle vérifie $f_Y = u - u \circ T_Y$, par construction, mais n'est en général pas lipschitzienne. On définit une distance d' sur Y en posant $d'(x, y) = d(T_Y x, T_Y y)$ si x et y sont dans le même

élément de la partition α_0 de Y, et $d'(x,y) = \lambda \sup_{Y\times Y} d$ sinon. Alors les propriétés de dilatation de T_Y assurent que $d' \geqslant d$. Comme T_Y est Gibbs-Markov pour la distance d, on en déduit qu'elle est également Gibbs-Markov pour la distance d' et la partition α_0.

Pour $b \in \alpha_0$, et $x, y \in b$, on a

$$|f_Y(x) - f_Y(y)| \leqslant \sum_0^{\varphi(b)-1} |f(x,k) - f(y,k)| \leqslant \sum_0^{\varphi(b)-1} Df(b,k) d((x,k),(y,k))$$
$$\leqslant C \sum_0^{\varphi(b)-1} Df(b,k) d(T_Y x, T_Y y) = C \sum_0^{\varphi(b)-1} Df(b,k) d'(x,y).$$

Ainsi, $Df_Y(b) \leqslant \sum_{k<\varphi(b)} Df(b,k)$, puis

$$\sum_{b\in\beta} m(b) Df_Y(b) \leqslant \sum_{b\in\beta, k<\varphi(b)} m(b) Df(b,k) = \sum_{b\in\beta,k<\varphi(b)} m(b,k) Df(b,k)$$
$$= \sum_{a\in\alpha} m(a) Df(a) < \infty.$$

Ainsi, le théorème 1.2.1 s'applique, et donne le résultat. □

Ainsi, certains résultats sur les applications Gibbs-Markov se généralisent bien aux applications localement Gibbs-Markov, mais ce n'est pas le cas de tous. Par exemple, le corollaire 1.1.21 sur la vitesse exponentielle de mélange ne s'étend pas, et on verra en fait dans le chapitre 2 de cette thèse que la vitesse de mélange est reliée à la mesure des points qui reviennent en un temps plus grand que n à la base Y. En particulier, la décroissance peut être polynomiale. Un des objectifs de cette thèse sera également d'étendre à certaines applications localement Gibbs-Markov des théorèmes limites bien connus pour les applications Gibbs-Markov, et qui peuvent se démontrer dans ce cas par des méthodes spectrales (qui ne fonctionnent plus dans le cas localement Gibbs-Markov).

1.3.4 Modélisation d'applications réelles

De nombreuses applications réelles admettent des extensions en forme de tour localement Gibbs-Markov, ce qui permet de leur appliquer les théorèmes du paragraphe précédent. Dans ce paragraphe, on décrit plus précisément sous quelles hypothèses on peut construire ce type d'extension.

Définition 1.3.17. *Soient (X,m,d) un espace métrique mesuré et $T : X \to X$ une application mesurable non-singulière. On dit que T* admet un modèle localement Gibbs-Markov *s'il existe un nombre fini de parties bornées deux à deux disjointes $U_1,\dots,U_N$ de X, une partition (modulo 0) finie ou dénombrable $(W_j)_{j\in J}$ de $\bigcup U_i$ dont chaque élément est inclus dans l'un des U_i, et des temps $\varphi_j \in \mathbb{N}^*$ tels que*

– *Pour tout $j \in J$, T^{φ_j} est un isomorphisme mesurable entre W_j et l'un des U_i, et sa distorsion $g(x) = \frac{\mathrm{d}m}{\mathrm{d}(m\circ T^{\varphi_j})}(x)$ vérifie : pour presque tous $x, y \in W_j$,*

$$\left|1 - \frac{g(x)}{g(y)}\right| \leqslant Cd(T^{\varphi_j}x, T^{\varphi_j}y)$$

où C est une constante indépendante de j.

– *Il existe $C > 0$ tel que, pour tout $j \in J$, pour presque tous $x, y \in W_j$, pour tout $k < \varphi_j$,*

$$d(T^k x, T^k y) \leqslant Cd(T^{\varphi_j}x, T^{\varphi_j}y).$$

– *Il existe $\lambda > 1$ tel que, pour tout $j \in J$, pour presque tous $x, y \in W_j$,*

$$d(T^{\varphi_j}x, T^{\varphi_j}y) \geqslant \lambda d(x, y).$$

– $\sum \varphi_j m(W_j) < \infty$.

Ces propriétés signifient que, après une attente éventuellement longue (le temps de retour φ_j), on a dilatation et contrôle de la distorsion. Dans des cas pratiques, on peut construire explicitement $U_1, \dots, U_N$ et la partition W_j pour montrer qu'une application admet un modèle localement Gibbs-Markov (voir par exemple le paragraphe 1.4.1), mais il est aussi possible d'utiliser des arguments abstraits de temps hyperboliques pour les construire (voir le chapitre 3).

Lorsque $N = 1$, i.e. il n'y a qu'un seul ensemble U_i, on dit qu'on a affaire à une *tour de Young* ([You99]).

Lorsque T admet un modèle localement Gibbs-Markov, on peut construire une tour localement Gibbs-Markov $X' = \{(x, k) \mid x \in W_j, k < \varphi_j\}$ et une application T' définie par $T'(x, k) = (x, k + 1)$ si $k < \varphi_j - 1$, et $T'(x, \varphi_j - 1) = (T^{\varphi_j}x, 0)$, comme dans le paragraphe 1.3.2. On obtient une projection $\pi : X' \to X$ telle que $\pi \circ T' = T \circ \pi$, et les résultats sur la tour localement Gibbs-Markov abstraite (X', T') permettent d'obtenir les résultats correspondants suivants pour (X, T) :

Proposition 1.3.18. *Soit $T : X \to X$ une application admettant un modèle localement Gibbs-Markov, de partition $\{U_1, \dots, U_N\}$. Soit $\Lambda = \bigcup_{n\geqslant 0} \bigcup_{i=1}^N T^n(U_i)$. Il existe alors un nombre fini de mesures de probabilité $m_1, \dots, m_k$ invariantes ergodiques absolument continues par rapport à m supportées par Λ. De plus, toute mesure μ invariante absolument continue supportée par Λ est combinaison linéaire des m_j, et presque tout point de Λ est dans le bassin d'une des mesures m_j.*

De plus, chaque mesure m_j a un support qui contient l'un des U_i, et sa densité est alors minorée sur U_i. Finalement, m_j se décompose comme $\mu_{j,0} + \dots + \mu_{j,d-1}$ (où d peut dépendre de j), avec $T_*(\mu_{j,s}) = \mu_{j,s+1}$ pour $s \in \mathbb{Z}/d\mathbb{Z}$, et $\mu_{j,s}$ mélangeante pour T^d.

Démonstration. Soient X' la tour localement Gibbs-Markov décrite ci-dessus, et $\pi : X' \to X$ la projection canonique. Alors X' admet un nombre fini de mesures invariantes $m'_1, \dots, m'_t$, par la proposition 1.3.11. Par la proposition 1.1.24, leurs densités sont majorées et minorées sur la base. La condition $\sum \varphi_j m(W_j) < \infty$ assure de plus que ces mesures sont de masse finie, et on peut donc supposer que ce sont des mesures de probabilité. La proposition 1.3.11 garantit alors que la réunion de leurs bassins recouvre presque tout X'.

Les mesures $\pi_*(m'_j)$ ne sont pas nécessairement toutes différentes. Soient $m_1, \dots, m_k$ ces mesures, sans répétition. Leurs bassins recouvrent alors presque tout Λ, donc il n'y a pas d'autre mesure invariante.

La seule assertion qu'il reste à vérifier est la décomposition en composantes mélangeantes. Soit $m_j = \pi_*(m'_j)$ l'une des mesures. Il existe une décomposition $m'_j = \nu_0 + \ldots + \nu_{d_1-1}$ où les mesure ν_s sont invariantes par $(T')^{d_1}$ et mélangeantes, d'après les propositions 1.3.14 et 1.3.15. Les mesures $\pi_*(\nu_s)$ sont invariantes et ergodiques par T^{d_1}, donc elles sont deux à deux égales ou singulières. On en déduit qu'il existe d (qui divise d_1) tel que $\pi_*(\nu_s) = \pi_*(\nu_t)$ si et seulement si $s \equiv t \mod d$. On obtient la conclusion en prenant $\mu_{j,s} = \pi_*(\nu_s)$ pour $s \leqslant d-1$. □

Notons que, dans le cas d'une tour de Young, il existe un unique ensemble U_i, et donc une unique mesure m_j.

On obtient le résultat suivant sur la régularité des cobords :

Théorème 1.3.19. *: Soient $T : X \to X$ une application admettant un modèle localement Gibbs-Markov de partition $\{U_1, \dots, U_N\}$ et G un groupe abélien localement compact polonais. Soient $f : X \to G$ une fonction lipschitzienne (ou plus généralement θ-hölderienne) et $u : X \to G$ telle que $f = u - u \circ T$ presque partout. Alors f est lipschitzienne (respectivement θ-hölderienne) sur l'un des U_j.*

Démonstration. Quitte à remplacer d par d^θ, il suffit de traiter le cas où f est lipschitzienne.

On travaille dans l'extension X' de X. Soit μ une des mesures de probabilité invariantes ergodiques données par la proposition 1.3.11. Elle est équivalente à m' sur un certain nombre des U_i, mettons $U_1, \dots, U_k$. Notons $Z = \bigcup_1^k U_i$.

Le théorème 1.3.16 donne une constante C telle que, pour presque tous x, y dans le même élément de la partition α et inclus dans Z, $|u(x) - u(y)| \leqslant C d(T^\varphi x, T^\varphi y)$.

Soit $a \subset Z$ un élément de α, de temps de retour n. Alors $T^n a$ est l'un des U_j (avec $j \leqslant k$), si bien que presque tous $x, y \in U_j$ ont des antécédents $x', y' \in a$. Alors, comme $f_Y = u - u \circ T_Y$ (avec les notations de la preuve du théorème 1.3.16), et $|f_Y(x) - f_Y(y)| \leqslant C' d(T^n x, T^n y)$, on a

$$|u(x') - u(y')| \leqslant |u(x) - u(y)| + |f_Y(x) - f_Y(y)| \leqslant C d(x', y') + C' d(x', y'). \qquad \square$$

Il est aussi possible d'utiliser les tours localement Gibbs-Markov abstraites pour modéliser des applications non-uniformément hyperboliques, en suivant la construction de [You98] :

dans ce cas, on obtient une partition de Markov au sens hyperbolique du terme, c'est-à-dire qu'il y a, en plus de la direction dilatante, une direction contractante. En identifiant les points d'une même feuille contractante, on obtient un espace quotient, qui est une tour de Young au sens précédent, et dans lequel on peut donc appliquer la proposition 1.3.18 ou le théorème 1.3.19.

Montrons en particulier comment utiliser le théorème 1.3.19 dans ce cadre. Supposons que $f = u - u \circ T$ sur un espace X dans lequel on peut construire une tour de Young hyperbolique X' de base $Y \subset X$. La fonction $f' = f \circ \pi$ sur X' s'écrit alors $f' = u' - u' \circ T'$, où $u' = u \circ \pi$. De plus, il existe une fonction f'', constante le long des feuilles stables et hölderienne, et une fonction v hölderienne, telles que $f' = v - v \circ T' + f''$. On a alors $f'' = (u' - v) - (u' - v) \circ T'$. Comme f'' est constante le long des feuilles stables, on en déduit par un argument de points de continuité mesurable que c'est aussi le cas de $u' - v$. On quotiente par les feuilles stables et on applique le théorème 1.3.19 dans le quotient, ce qui donne que $u' - v$ est hölderienne sur Y. Comme v est hölderienne par construction, u' est donc aussi hölderienne.

1.4 Exemples en dimension 1

Dans ce paragraphe, on donne des exemples d'applications en dimension 1 auxquelles on peut appliquer les résultats qui précèdent. Ces applications ne seront pas uniformément dilatantes (sans quoi on pourrait appliquer des méthodes spectrales directes, voir par exemple [Bro96]). La non uniformité dans la dilatation proviendra de la présence d'un point fixe où la dérivée est égale à 1. Ainsi, les points proches du point fixe vont mettre longtemps à s'en éloigner, et c'est le comportement précis au voisinage du point fixe qui dictera le comportement asymptotique de la dynamique.

1.4.1 L'application de Liverani-Saussol-Vaienti

Soit $\alpha \in]0, 1[$. On considère l'application T définie sur le segment $[0, 1]$ par

$$T(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{si } 0 \leqslant x \leqslant 1/2, \\ 2x - 1 & \text{si } 1/2 < x \leqslant 1. \end{cases} \tag{1.25}$$

Cette application (LSV) envoie chacun des deux intervalles $[0, 1/2]$ et $]1/2, 1]$ sur $[0, 1]$ (en négligeant les bords des intervalles, qui ne joueront aucun rôle dans la suite puisqu'on s'intéressera au comportement presque partout par rapport à la mesure de Lebesgue). Sur la branche de droite, elle est affine, et en particulier uniformément dilatante, tandis qu'elle a un point fixe en 0 qui est topologiquement répulsif mais où la dérivée vaut 1. De plus, le comportement du point fixe dépend du coefficient α : plus il est élevé, et plus le point fixe est neutre, i.e. plus les autres points mettent de temps à s'éloigner du point fixe. Notons que la dérivée seconde de T tend vers $+\infty$ en 0 : le comportement précis du point fixe est

en $x + cx^{1+\alpha}$ pour une certaine constante c sans importance (ajustée ici de telle sorte que $T(1/2) = 1$).

Cette application a été introduite par Liverani, Saussol et Vaienti dans [LSV99]. C'est un modèle simple pour d'autres applications, dont l'étude a débuté avec Pomeau et Manneville. Citons par exemple l'application donnée sur $[0, 1]$ par $x \mapsto \{x(1 + x^\alpha)\}$, où $\{x\}$ est la partie fractionnaire de x : elle a exactement le même comportement que l'application LSV, avec deux branches dont l'image est tout l'intervalle $[0, 1]$, celle de droite étant uniformément dilatante, et celle de gauche admettant un point fixe neutre en $x + x^{1+\alpha}$. On peut aussi considérer les applications données par $\{x/(1 - x)^\alpha\}$ ou $\{x(1 - x^\alpha)^{-1/\alpha}\}$, qui ont un point fixe neutre du même type et une infinité de branches uniformément dilatantes (qui ne changent rien ni aux problèmes qui se posent, ni aux résultats).

Tous les résultats qu'on va décrire dans la suite restent valables pour ces transformations, avec les mêmes preuves. Par souci de simplicité, on ne les formulera cependant que dans le cadre des applications LSV.

On ne donnera pas de preuves détaillées dans ce paragraphe, puisqu'on traitera en détail un cas plus général dans le paragraphe 1.4.2

Théorème 1.4.1. *Soient $\alpha \in]0, 1[$, et T l'application LSV correspondante. Alors l'application induite T_Y sur $Y =]1/2, 1]$ est Gibbs-Markov. Par conséquent, T admet une mesure m invariante absolument continue par rapport à la mesure de Lebesgue. De plus, cette mesure est de masse finie, et sa densité h est lipschitzienne strictement positive sur tout intervalle $[\varepsilon, 1]$.*

Finalement, soit φ_Y le temps de premier retour de Y à lui-même. Alors $m(\varphi_Y > n) \sim \frac{h(1/2)}{4(\alpha n)^{1/\alpha}}$.

Ce théorème se trouve dans [LSV99]. Pour construire une partition markovienne pour T_Y, on pose $x_0 = 1$, puis on définit par récurrence $x_{n+1} = T^{-1}(x_n) \cap [0, 1/2]$. On vérifie alors que $x_n \sim \frac{1}{2}(\alpha n)^{-1/\alpha}$. Soit ensuite $y_n = T^{-1}(x_{n-1}) \cap]1/2, 1]$. Ainsi, $J_n :=]y_{n+1}, y_n]$ est l'ensemble des points de Y qui reviennent à Y en exactement n itérations. On vérifie alors en contrôlant explicitement la distorsion que T_Y est Gibbs-Markov pour la partition $\{J_n\}$. De plus, en utilisant l'estimée sur x_n, on vérifie que

$$\sum nm(J_n) \leqslant C \sum n \operatorname{Leb}(J_n) \leqslant C \sum n \frac{1}{n^{1/\alpha+1}} < \infty$$

car $\alpha < 1$. Ainsi, l'application T admet un modèle localement Gibbs-Markov (définition 1.3.17). Par la proposition 1.3.18, elle admet donc une mesure de probabilité invariante m, dont la densité h est lipschitzienne sur Y.

Il reste à vérifier que h est localement lipschitzienne partout, et en particulier continue en $1/2$, ce qui n'est *a priori* pas évident. [LSV99] le prouve en construisant h comme un élément d'un cône de fonctions continues. Dans la partie 1.5, on verra sur un autre exemple comment on peut démontrer ce type de continuité, sans hypothèse *a priori* sur la forme de la densité.

Remarquons qu'on s'est restreint au cas $\alpha < 1$ puisque, lorsque $\alpha \geqslant 1$, la mesure invariante absolument continue obtenue de cette façon est de masse infinie. On montre en fait que, pour Lebesgue presque tout x, la mesure $\frac{1}{n}\sum_0^{n-1} \delta_{T^k(x)}$ converge faiblement vers δ_0 la masse de Dirac au point fixe neutre. Il ne peut donc exister de mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue.

On peut appliquer dans ce cadre le théorème 1.3.19 sur les cobords :

Proposition 1.4.2. *Soient T l'application LSV de paramètre $\alpha \in]0,1[$ et $G = S^1$ ou $\mathbb{R}$. Soit $f : [0,1] \to G$ une application hölderienne d'exposant $\gamma > 0$ sur les deux intervalles $[0,1/2]$ et $]1/2,1]$. Si $u : [0,1] \to G$ est mesurable et vérifie $f = u - u\circ T$ Lebesgue presque partout, alors il existe une fonction $\widetilde{u}$ égale à u sur un ensemble de mesure de Lebesgue pleine, hölderienne d'exposant γ, et telle que $f = \widetilde{u} - \widetilde{u}\circ T$ partout.*

Démonstration. Le théorème 1.3.19 donne que, si f est hölderienne sur $[0,1/2]$ et $]1/2,1]$ et s'écrit $f = u - u\circ T$, alors u est presque partout hölderienne sur $]1/2,1]$. Comme $T :]1/2,1] \to]0,1]$ est lipschitzienne et d'inverse lipschitzienne, l'équation de cobord donne donc que u est presque partout hölderienne sur $[0,1]$, i.e. il existe un ensemble V de mesure de Lebesgue pleine et une constante C tels que, pour tous $x, y \in V$, $|u(x) - u(y)| \leqslant C|x-y|^\gamma$. On peut aussi supposer que $f = u - u\circ T$ partout sur V.

En particulier, l'application u est uniformément continue sur V, si bien qu'elle se prolonge de façon unique en une application continue (et même γ-hölderienne) $\widetilde{u}$ sur $\overline{V} = [0,1]$. Sur $V \cap T^{-1}(V)$, de mesure de Lebesgue pleine et par conséquent dense, on a $f(x) = \widetilde{u}(x) - \widetilde{u}\circ T(x)$. Comme les deux membres de cette égalité sont continus sur les deux intervalles $[0,1/2]$ et $]1/2,1]$, cette égalité s'étend à $[0,1]$. □

En particulier, si f est hölderienne sur $[0,1/2]$ et $]1/2,1]$ et est un cobord mesurable, on a pour tout x périodique de période n que $\sum_0^{n-1} f(T^k x) = \widetilde{u}(T^n x) - \widetilde{u}(x) = 0$.

Corollaire 1.4.3. *Si $f : [0,1] \to \mathbb{R}$ est hölderienne sur $[0,1/2]$ et $]1/2,1]$ et vérifie $f(0) \neq 0$, alors f n'est pas un cobord mesurable pour l'application LSV.*

Cela résout une conjecture de [FHV03] : dans cet article, les auteurs ont besoin d'admettre que $f = \log|T'| - \int \log|T'|$ n'est pas un cobord, pour obtenir une variance non nulle dans le théorème central limite. Comme f est α-hölderienne sur $[0,1/2]$ et $]1/2,1]$ et $f(0) = -\int \log|T'| < 0$, le corollaire garantit que ce n'est effectivement pas le cas.

En utilisant le résultat pour $G = S^1$, on peut même prouver un résultat plus fort :

Corollaire 1.4.4. *Soit $f(x) = \log|T'| - \int \log|T'|$. Il n'existe pas de fonction mesurable $u : [0,1] \to \mathbb{R}$, de fonction mesurable $q : X \to \mathbb{Z}$ et de réels ρ, λ tels que $f = u - u\circ T + \lambda q + \rho$ presque partout.*

Démonstration. Supposons que $f = u - u\circ T + \lambda q + \rho$. Dans $G = \mathbb{R}/\lambda\mathbb{Z}$, on a $\overline{f} = \overline{u} - \overline{u}\circ T + \overline{\rho}$. Ainsi, la proposition 1.4.2 assure que $\overline{u}$ est hölderienne presque partout. On la prolonge en une fonction hölderienne, puis on la relève sur $[0,1]$ simplement connexe. Cette nouvelle fonction $\widetilde{u}$ coïncide presque partout avec u, à un élément de $\lambda\mathbb{Z}$ près. Quitte à remplacer q par une fonction $\widetilde{q}$, on peut supposer que $u = \widetilde{u}$ presque partout.

Ainsi, $\widetilde{u} : [0,1] \to \mathbb{R}$ est continue et $f = \widetilde{u} - \widetilde{u} \circ T + \lambda\widetilde{q} + \rho$ presque partout, $\widetilde{q}$ étant à valeurs entières. Cette équation donne que $\widetilde{q}$ admet une version continue, et donc constante, sur $[0,1/2]$ et $]1/2,1]$. En identifiant les valeurs en 0 et en 1, on obtient finalement que

$$\begin{aligned} f &= \widetilde{u} - \widetilde{u} \circ T + f(0) \text{ sur } [0,1/2], \\ f &= \widetilde{u} - \widetilde{u} \circ T + f(1) \text{ sur }]1/2,1]. \end{aligned} \tag{1.26}$$

Dans l'intervalle $]1/2,1]$, l'application T^2 a deux points fixes x_0 et 1. Alors (1.26) donne que $f(x_0) + f(Tx_0) = f(0) + f(1)$. Mais $f(x_0) = f(1)$ et $f(0) < f(Tx_0)$, c'est donc absurde. □

Ce corollaire est un résultat fort d'apériodicité de la fonction f. On verra dans la partie 5.5 qu'il implique un théorème limite fort pour f, le théorème de la limite locale.

1.4.2 Points fixes très neutres

Dans ce paragraphe, on étudie une classe d'exemples introduite dans [Hol05]. Ce sont des analogues des applications LSV, mais où le point fixe peut être encore plus neutre. Cela donnera des vitesses de décorrélation arbitrairement lentes.

Définition 1.4.5. *On dit qu'une fonction $l : \mathbb{R}_+^* \to \mathbb{R}_+^*$ est à variation lente en $+\infty$ si, pour tout $\lambda > 0$, $\frac{l(\lambda x)}{l(x)} \to 1$ quand $x \to \infty$.*

On dit que $\rho : \mathbb{R}_+^* \to \mathbb{R}_+^*$ est à variation lente en 0 si $l(x) = \rho(1/x)$ est à variation lente en $+\infty$.

On dira qu'une application $T : [0,1] \to [0,1]$ est une *application de Holland* si elle s'écrit sous la forme

$$T(x) = \begin{cases} x(1 + x^\alpha \rho(x)) & \text{si } 0 \leqslant x \leqslant 1/2, \\ 2x - 1 & \text{si } 1/2 < x \leqslant 1, \end{cases}$$

où $\alpha > 0$ et $\rho : \mathbb{R}_+^* \to \mathbb{R}_+^*$ est une fonction à variation lente en 0, de classe C^2, avec $x\rho'(x) = o(\rho(x))$ et $x^2\rho''(x) = o(\rho(x))$, normalisée de telle sorte que $T(1/2) = 1$. On suppose de plus que, pour tout $x \neq 0$, $T'(x) > 1$.

Remarque 1.4.6. Les conditions $x\rho'(x) = o(\rho(x))$ et $x^2\rho''(x) = o(\rho(x))$ interviennent dans le contrôle de la distorsion. En général, pour toute fonction $\widetilde{\rho}$ à variation lente, il existe une fonction ρ qui lui est équivalente et vérifie ces conditions, par [BGT87, théorème 1.8.2]. Ainsi, on ne restreint pas les comportements asymptotiques possibles en faisant cette hypothèse.

Par exemple, les applications LSV sont des applications de Holland, pour $\rho(x) = 2^\alpha$.

On définit une suite x_n par $x_0 = 1$ et $x_{n+1} = T^{-1}(x_n) \cap [0, 1/2]$.

Proposition 1.4.7. *On a*

$$\frac{1}{x_n^\alpha \rho(x_n)} \sim \alpha n. \tag{1.27}$$

Démonstration. La suite x_n, étant strictement positive et décroissante, converge vers le point fixe de la branche gauche de T, c'est-à-dire 0.

Comme $x\rho'(x) = o(\rho(x))$, la fonction $\rho(x)$ est une fonction à variation lente normalisée, c'est-à-dire qu'on peut l'écrire sous la forme $\rho(x) = \exp\left(\int_x^1 \frac{\varepsilon(u)}{u}\,\mathrm{d}u\right)$, où ε tend vers 0 quand x tend vers 0 (on peut prendre $\varepsilon(x) = -\frac{x\rho'(x)}{\rho(x)}$ sur un voisinage de 0). Pour n assez grand, on a alors

$$\left|\int_{x_{n+1}}^{x_n} \frac{\varepsilon(u)}{u}\,\mathrm{d}u\right| \leqslant \varepsilon \ln\left(\frac{x_n}{x_{n+1}}\right) = \varepsilon \ln(1 + x_{n+1}^\alpha \rho(x_{n+1})) \sim \varepsilon x_{n+1}^\alpha \rho(x_{n+1}),$$

puisque $x^\alpha\rho(x)$ tend vers 0 quand x tend vers 0, car $\alpha > 0$ et ρ est à variation lente. Ainsi, $\int_{x_{n+1}}^{x_n} \frac{\varepsilon(u)}{u} = o(x_{n+1}^\alpha \rho(x_{n+1}))$. En prenant l'exponentielle, on trouve que

$$\frac{\rho(x_{n+1})}{\rho(x_n)} = 1 + o(x_{n+1}^\alpha \rho(x_{n+1})).$$

On calcule alors

$$\begin{aligned}
\frac{1}{x_n^\alpha \rho(x_n)} &= \frac{\rho(x_{n+1})}{\rho(x_n)} \frac{1}{x_{n+1}^\alpha \rho(x_{n+1})} \frac{1}{(1 + x_{n+1}^\alpha \rho(x_{n+1}))^\alpha} \\
&= \frac{1}{x_{n+1}^\alpha \rho(x_{n+1})} \left(1 + o(x_{n+1}^\alpha \rho(x_{n+1}))\right) \left(1 - \alpha x_{n+1}^\alpha \rho(x_{n+1}) + o(x_{n+1}^\alpha \rho(x_{n+1}))\right) \\
&= \frac{1}{x_{n+1}^\alpha \rho(x_{n+1})} - \alpha + o(1).
\end{aligned}$$

On obtient la conclusion en sommant les équivalents. □

Remarque 1.4.8. Dans la preuve, on a utilisé le fait que ρ est normalisée, et donc que $x\rho'(x) = o(\rho(x))$. La proposition reste en fait vraie sans cette hypothèse : on peut encadrer ρ par deux fonctions ρ_1 et ρ_2 normalisées qui lui sont équivalentes ([BGT87, théorème 1.8.2]), appliquer le résultat pour les deux applications T_1 et T_2 correspondantes, et obtenir ainsi deux suites y_n et z_n qui encadrent x_n, avec $\frac{1}{y_n^\alpha \rho_1(y_n)} \sim n\alpha$ et $\frac{1}{z_n^\alpha \rho_2(z_n)} \sim n\alpha$. Comme $\rho_1 \sim \rho_2$, on en déduit que $y_n \sim z_n$. Comme $y_n \leqslant x_n \leqslant z_n$, cela implique donc le comportement asymptotique désiré pour x_n.

Proposition 1.4.9. *Il existe $C > 0$ tel que, pour tout $n \in \mathbb{N}$, pour tous $x, y \in [x_{n+1}, x_n]$,*

$$\left|1 - \frac{(T^n)'(x)}{(T^n)'(y)}\right| \leqslant C|T^n(x) - T^n(y)|.$$

Démonstration. Les hypothèses sur ρ garantissent que $\frac{T''(x)}{T'(x)} \sim (\alpha+1)\alpha x^{\alpha-1}\rho(x)$ quand $x \to 0$. De plus, comme $x_{n+1} \sim x_n$ et ρ est à variation lente, on en déduit que, pour tout point x dans $[x_{n+1}, x_n]$, $\frac{T''(x)}{T'(x)} \sim (\alpha+1)\alpha x_{n+1}^{\alpha-1}\rho(x_{n+1})$.

Fixons maintenant $x < y \in [x_{n+1}, x_n]$. L'égalité des accroissements finis donne, pour tout $0 \leqslant k \leqslant n-1$, un point $\xi_k \in [T^k(x), T^k(y)] \subset [x_{n+1-k}, x_{n-k}]$ tel que $\frac{\log T'(T^k x)}{\log T'(T^k y)} = \frac{T''(\xi_k)}{T'(\xi_k)}(T^k x - T^k y)$. Ainsi,

$$\begin{aligned}\left|\log\left(\frac{(T^n)'(x)}{(T^n)'(y)}\right)\right| &\leqslant \sum_{k=0}^{n-1}\left|\log\left(\frac{T'(T^k x)}{T'(T^k y)}\right)\right| = \sum_{k=0}^{n-1}\frac{T''(\xi_k)}{T'(\xi_k)}|T^k x - T^k y| \\ &\leqslant \sum_{k=0}^{n-1}\frac{T''(\xi_k)}{T'(\xi_k)}|x_{n-k} - x_{n-k+1}|.\end{aligned}$$

Mais $\frac{T''(\xi_k)}{T'(\xi_k)} \leqslant C x_{n-k+1}^{\alpha-1}\rho(x_{n-k+1})$, et $x_{n-k} - x_{n-k+1} = x_{n-k+1}^{1+\alpha}\rho(x_{n-k+1})$, donc

$$\sum_{k=0}^{n-1}\frac{T''(\xi_k)}{T'(\xi_k)}|x_{n-k} - x_{n-k+1}| \leqslant C\sum_{k=0}^{n-1}\left(x_{n-k+1}^{\alpha}\rho(x_{n-k+1})\right)^2 \leqslant C'\sum_{k=0}^{n-1}\frac{1}{(n-k+1)^2} \leqslant C''$$

par (1.27).

On a montré que la distorsion de $T^n : [x_{n+1}, x_n] \to [x_1, x_0]$ est bornée uniformément, par une constante indépendante de n. En appliquant ce résultat à $T^k(x), T^k(y) \in [x_{n-k+1}, x_{n-k}]$ et T^{n-k}, on obtient $\frac{|T^k(x)-T^k(y)|}{|x_{n-k}-x_{n-k+1}|} \leqslant C|T^n x - T^n y|$. Cela permet de reprendre les calculs précédents, en remplaçant la majoration de $|T^k x - T^k y|$ en $|x_{n-k} - x_{n-k+1}|$ par une majoration en $C|T^n x - T^n y||x_{n-k} - x_{n-k+1}|$. On obtient donc $\left|\log\left(\frac{(T^n)'(x)}{(T^n)'(y)}\right)\right| \leqslant C''C|T^n x - T^n y|$. □

Théorème 1.4.10. *Soit T une application de Holland. Alors elle admet une mesure m invariante absolument continue par rapport à la mesure de Lebesgue, qui est unique à multiplication par un scalaire près. De plus, la densité h de cette mesure est lipschitzienne et strictement positive sur tout intervalle $[\varepsilon, 1]$ avec $\varepsilon > 0$.*

De plus, si φ_Y désigne le premier temps de retour à $Y =]1/2, 1]$, on a $m(\varphi_Y > n) \sim h(1/2)\frac{x_n}{2}$. En particulier, m est de masse finie si et seulement si $\sum x_n < \infty$. Dans ce cas, m est mélangeante.

Démonstration. Notons $y_n = T^{-1}(x_{n-1}) \cap]1/2, 1]$: ainsi, $J_n :=]y_{n+1}, y_n]$ est l'ensemble des points qui reviennent en temps n dans $Y =]1/2, 1]$. De plus, la proposition 1.4.9 permet de contrôler la distorsion de T^n sur cet intervalle, et T^n y est dilatante d'au moins $\lambda := \min(2, \inf_{x_2 \leqslant x \leqslant x_1} T'(x)) > 1$. Ainsi, les propositions 1.1.17 et 1.1.24 assurent que T_Y admet une mesure invariante absolument continue, dont la densité h est lipschitzienne sur $]1/2, 1]$, majorée et minorée. En induisant, on obtient la mesure invariante absolument continue de

T. De plus, elle est ergodique d'après la proposition 1.3.3. Cela implique donc qu'elle est unique, par [Aar97, théorème 1.5.6]. On verra dans un cadre plus compliqué (la preuve du théorème 1.5.11) une technique permettant de vérifier que h est en fait continue, et même lipschitzienne sur tout intervalle $[\varepsilon, 1]$ lorsque $\varepsilon > 0$.

De plus, $m(\varphi_Y > n) = m(]1/2, y_{n+1}]) = \int_{1/2}^{y_{n+1}} h \, \mathrm{dLeb}$. Comme h est continue et $y_{n+1} = \frac{1}{2} + \frac{x_{n+1}}{2}$, avec $x_{n+1} \sim x_n$, on trouve donc que

$$m(\varphi_Y > n) \sim h\left(\frac{1}{2}\right)\frac{x_n}{2}.$$

De plus, m est de masse finie si et seulement si $\sum nm(\varphi_Y = n) < \infty$, i.e. $\sum m(\varphi_Y > n) < \infty$. En ce cas, la proposition 1.3.18 montre que m est mélangeante. □

Pour estimer x_n, on peut utiliser la notion de conjuguée de de Bruijn :

Définition 1.4.11. *Soit l une fonction à variation lente en $+\infty$. Il existe une unique (à équivalence asymptotique près) fonction $l^{\#}$ à variation lente en $+\infty$, appelée* conjuguée de de Bruijn *de l, satisfaisant $l(x)l^{\#}(xl(x)) \to 1$ et $l^{\#}(x)l(xl^{\#}(x)) \to 1$.*

On a en particulier $l^{\#\#} \sim l$. Cette notion est utile pour définir un inverse asymptotique : si $f(x) = xl(x)$ avec l à variation lente, alors la fonction $g(x) = xl^{\#}(x)$ vérifie $f(g(x)) \sim x$ et $g(f(x)) \sim x$.

Proposition 1.4.12. *On a*

$$x_n \sim \frac{1}{(\alpha n)^{1/\alpha} l^{\#}(n^{1/\alpha})}, \tag{1.28}$$

où $l(x) = \frac{1}{\rho(1/x)^{1/\alpha}}$.

De plus, $\sum x_n < \infty$ si et seulement si $\alpha < 1$, ou $\alpha = 1$ et $\int_0^1 \frac{1}{x\rho(x)} < \infty$.

Remarque 1.4.13. On en déduit que, pour toute fonction $\widetilde{l}$ à variation lente en $+\infty$, il existe un choix de ρ pour lequel $x_n \sim \frac{1}{n^{1/\alpha}\widetilde{l}(n)}$: il suffit de choisir $\rho(x) \sim \left[\left[\alpha^{-1/\alpha}\widetilde{l}(x^{-\alpha})\right]^{\#}\right]^{-\alpha}$.

Démonstration. La suite x_n satisfait $\frac{1}{x_n^{\alpha}\rho(x_n)} \sim \alpha n$. Alors, pour $u_n = 1/x_n$, on a $u_n l(u_n) \sim (\alpha n)^{1/\alpha}$, donc $u_n \sim (\alpha n)^{1/\alpha} l^{\#}(n^{1/\alpha})$, ce qui donne (1.28).

Lorsque $\alpha < 1$, la suite x_n est toujours sommable. Lorsque $\alpha > 1$, ce n'est jamais le cas. On peut donc se concentrer sur le cas $\alpha = 1$ pour étudier la sommabilité de x_n. On a alors $\sum x_n < \infty$ si et seulement si $\sum \frac{1}{nl^{\#}(n)} < \infty$, ce qui équivaut encore à $\int_1^{\infty} \frac{1}{xl^{\#}(x)} \, \mathrm{d}x < \infty$. Le lemme suivant et un changement de variables en $1/x$ donnent donc la conclusion de la proposition. □

Lemme 1.4.14. *Soient l une fonction à variation lente en $+\infty$ et $l^\#$ sa conjuguée de de Bruijn. Alors $\int_1^\infty \frac{l(x)}{x}\,\mathrm{d}x < \infty$ si et seulement si $\int_1^\infty \frac{1}{xl^\#(x)}\,\mathrm{d}x < \infty$.*

Démonstration. La finitude des intégrales ne dépend que de la classe asymptotique de l (et $l^\#$). On peut donc sans perte de généralité supposer que $l \in SR_0$ ([BGT87, paragraphe 1.8]). En ce cas, $f(x) = xl(x)$ admet un véritable inverse $g(x) = xl^\#(x)$ au voisinage de $+\infty$. On calcule alors $\int_1^\infty \frac{\mathrm{d}x}{g(x)}$ par le changement de variables $x = f(y)$. Notons que $yf'(y) \sim f(y)$ car $f \in SR_1$. Ainsi,

$$\int_1^\infty \frac{\mathrm{d}x}{xl^\#(x)} = \int_1^\infty \frac{\mathrm{d}x}{g(x)} = \int_{g(1)}^\infty \frac{f'(y)}{g(f(y))}\,\mathrm{d}y \sim \int_{g(1)}^\infty \frac{f(y)}{y^2}\,\mathrm{d}y = \int_{g(1)}^\infty \frac{l(y)}{y}\,\mathrm{d}y.$$

Par conséquent, l'une des intégrales est finie si et seulement si l'autre l'est. □

Remarque 1.4.15. En général, la conjuguée de de Bruijn peut être très compliquée à calculer, mais, dans les cas pratiques, on a souvent $l^\#(x) \sim 1/l(x)$. C'est par exemple le cas quand $l(x) = \prod_1^s (\log_k x)^{\alpha_k}$, où $\log_k$ désigne le log itéré k fois ([BGT87, appendice 5]). Ainsi, pour $\rho(x) = \prod_1^s (\log_k(1/x))^{\alpha_k}$, on trouve que

$$x_n \sim \frac{1}{(\alpha n)^{1/\alpha} \prod_1^s \log_k(n^{1/\alpha})^{\alpha_k/\alpha}}.$$

1.5 Un produit gauche en dimension 2

Dans cette partie, on étudie un exemple en dimension 2 qui présente le même type de propriétés que les applications LSV en dimension 1. Contrairement à [Hu01] ou [PY01], qui étudient le cas des points fixes isolés, il y aura dans notre exemple toute une courbe neutre. On va voir que les résultats valables en dimension 1 restent quand même essentiellement vrais.

Plus précisément, soit T_α l'application LSV de paramètre α, donnée par (1.25). Pour mélanger les comportements de différents T_α, on va considérer un produit gauche, analogue à l'application d'Alves-Viana ([Via97]) mais où les applications unimodales sont remplacées par des T_α. Soit $\alpha : S^1 \to]0,1[$ une application de minimum $\alpha_{\min}$ et de maximum $\alpha_{\max}$. On suppose que

1. α est C^2.
2. $0 < \alpha_{\min} < \alpha_{\max} < 1$.
3. α prend la valeur $\alpha_{\min}$ en un unique point x_0, avec $\alpha''(x_0) > 0$.
4. $\alpha_{\max} < \frac{3}{2}\,\alpha_{\min}$ (ce qui implique $\alpha_{\max} < \alpha_{\min} + 1/2$).

Ces conditions sont par exemple satisfaites par $\alpha(\omega) = \alpha_{\min} + \varepsilon(1 + \sin(2\pi\omega))$ où $\alpha_{\min} \in]0,1[$ et ε est assez petit.

On définit une application T sur $S^1 \times [0,1]$ par

$$T(\omega, x) = (F(\omega), T_{\alpha(\omega)}(x)) \tag{1.29}$$

où $F(\omega) = 4\omega$.

Dans la suite, on va généraliser à ce produit gauche les résultats en dimension 1 sur les applications T_α. On va montrer qu'il existe une unique probabilité m invariante absolument continue par rapport à la mesure de Lebesgue, dont la densité h sera en fait Lipschitz sur tout compact de $S^1 \times]0,1]$ (théorème 1.5.11). Pour cela, on induira sur $Y = S^1 \times]1/2, 1]$, et on verra que l'application induite est Gibbs-Markov. On étudiera ensuite précisément le comportement asymptotique de $m(\varphi_Y > n)$, où φ_Y est le temps de retour à Y. Cela permettra, dans le chapitre 4 de cette thèse, d'obtenir des théorèmes limites, plus précisément la convergence des sommes de Birkhoff de fonctions hölderiennes vers une loi normale ou une loi stable, suivant la valeur de $\alpha_{\min}$ (théorème 4.2.6). L'étude précise de $m(\varphi_Y > n)$ utilisera des propriétés très fortes de mélange de l'application F sur S^1, et en particulier des propriétés de décorrélation multiple, dont l'importance a été notée dans [Pèn02].

Remarque 1.5.1. La forme particulière de F n'a pas d'importance, tous les résultats restent vrais lorsque F est C^2 avec $|F'| \geqslant 4$ (par exemple $F(\omega) = d\omega$, pour $d \geqslant 4$). De la même façon, la seule propriété importante des applications T_α est leur comportement près de 0 et le fait qu'elles soient markoviennes. Finalement, l'hypothèse $\alpha''(x_0)$ (qui n'est utile que pour l'estimation précise de $m(\varphi_Y > n)$) peut être remplacée par : il existe k tel que $\alpha^{(k)}(x_0) \neq 0$ (mais les facteurs de normalisation doivent alors être modifiés en conséquence). Par souci de simplicité, on se limitera dans la suite au cas évoqué plus haut.

1.5.1 Estimées préliminaires

Pour appliquer la proposition 1.3.11, on va construire une partition de Markov pour l'application T_Y induite par T sur $Y = S^1 \times]1/2, 1]$, et on va contrôler la distorsion des différentes branches de T_Y.

On écrira $T_\omega^n = T_{\alpha(F^{n-1}\omega)} \circ \cdots \circ T_{\alpha(\omega)}$, si bien que $T^n(\omega, x) = (F^n\omega, T_\omega^n(x))$. On pose aussi $d((\omega_1, x_1), (\omega_2, x_2)) = |\omega_1 - \omega_2| + |x_1 - x_2|$. Un point de $S^1 \times [0,1]$ sera noté $\mathbf{x} = (\omega, x)$. Finalement, soit $\mathrm{d}_{\mathrm{vert}}((\omega_1, x_1), (\omega_2, x_2)) = |x_2 - x_1|$.

On définit $X_0(\omega) = 1$, $X_1(\omega) = 1/2$ et, pour $n \geqslant 2$, soit $X_n(\omega)$ la préimage dans $[0, 1/2]$ de $X_{n-1}(F\omega)$ par $T_{\alpha(\omega)}$. Ces X_n serviront à construire une partition de Markov pour T dans le paragraphe 1.5.2.

Proposition 1.5.2. *Il existe $C > 0$ tel que $\forall n \in \mathbb{N}^*, \forall \omega \in S^1$,*

$$\frac{1}{Cn^{1/\alpha_{min}}} \leqslant X_n(\omega) \leqslant \frac{C}{n^{1/\alpha_{max}}}.$$

Démonstration. Soit $V(x) = x(1 + 2^{\alpha_{\max}} x^{\alpha_{\min}})$. On définit une suite Z_n par $Z_1 = 1/2$ et $V(Z_{n+1}) = Z_n$ pour $n \geqslant 1$. Comme $V(x) \geqslant T_{\alpha(\omega)}(x)$ pour tout ω, on vérifie par récurrence que $Z_n \leqslant X_n(\omega)$ pour tout ω. Ainsi, il suffit d'estimer Z_n pour obtenir la minoration. Comme $V(x) \geqslant x$, la suite Z_n, décroissante et positive, tend vers le point fixe 0 de V.

On a

$$\begin{aligned}\frac{1}{Z_n^{\alpha_{\min}}} &= \frac{1}{Z_{n+1}^{\alpha_{\min}}} \left(1 + 2^{\alpha_{\max}} Z_{n+1}^{\alpha_{\min}}\right)^{-\alpha_{\min}} = \frac{1}{Z_{n+1}^{\alpha_{\min}}} \left(1 - \alpha_{\min} 2^{\alpha_{\max}} Z_{n+1}^{\alpha_{\min}} + o(Z_{n+1}^{\alpha_{\min}})\right) \\ &= \frac{1}{Z_{n+1}^{\alpha_{\min}}} - \alpha_{\min} 2^{\alpha_{\max}} + o(1).\end{aligned}$$

En sommant, on trouve $\frac{1}{Z_m^{\alpha_{\min}}} \sim m\, \alpha_{\min} 2^{\alpha_{\max}}$, donc $Z_m \sim C/m^{1/\alpha_{\min}}$. Cela montre la minoration.

La majoration est analogue, en utilisant une suite Z'_n avec $Z'_n \geqslant X_n(\omega)$. □

On fixe une fois pour toutes une constante D assez grande. La définition qui suit est analogue à une définition de Viana ([Via97]).

Définition 1.5.3. *Soit $\psi : K \to [0,1]$, où K est un intervalle de S^1. On dit que le graphe de ψ est une* courbe admissible *si ψ est C^1 avec $|\psi'| \leqslant D$.*

Proposition 1.5.4. *Soit ψ une courbe admissible, définie sur K avec $|K| < 1/4$, et incluse dans $K \times [0, 1/2]$ ou $K \times]1/2, 1]$. Alors l'image du graphe de ψ par T est encore une courbe admissible.*

Démonstration. Soit (u, v) un vecteur tangent à (ω, x), avec $|v| \leqslant D|u|$. On doit vérifier que son image (u', v') par $DT(\omega, x)$ vérifie encore $|v'| \leqslant D|u'|$.

Supposons tout d'abord que $x \leqslant 1/2$, si bien que $u' = 4u$ et $v' = (1 + (2x)^{\alpha(\omega)}(\alpha(\omega) + 1))v + x \ln(2x)\alpha'(\omega)(2x)^{\alpha(\omega)}u$. Comme $\alpha(\omega) \leqslant \alpha_{\max} \leqslant 1$, on obtient $|v'| \leqslant 3|v| + C|u|$ pour une certaine constante C qui ne dépend que de $\|\alpha'\|_\infty$. Ainsi,

$$\frac{|v'|}{|u'|} \leqslant \frac{3}{4}\frac{|v|}{|u|} + \frac{C}{4}.$$

Cela donne $|v'|/|u'| \leqslant D$ si $\frac{3}{4}D + \frac{C}{4} \leqslant D$, ce qui est vrai si D est assez grand.

Supposons maintenant que $x > 1/2$. Alors $u' = 4u$ et $v' = 2v$, et il n'y a rien à faire. □

Corollaire 1.5.5. *Soient (ω_1, x_1) et (ω_2, x_2) deux points de $S^1 \times [0, 1/2]$ avec $|x_1 - x_2| \leqslant D|\omega_1 - \omega_2|$ et $|\omega_1 - \omega_2| \leqslant \frac{1}{8}$. Alors leurs images satisfont $|x'_1 - x'_2| \leqslant D|\omega'_1 - \omega'_2|$.*

Démonstration. On utilise un segment entre les deux points : c'est une courbe admissible, donc son image est encore admissible. □

1.5.2 La partition de Markov

Soit $Y = S^1 \times]1/2, 1]$. Pour $\mathbf{x} \in Y$, soit $\varphi_Y(\mathbf{x}) = \inf\{n > 0 \mid T^n(\mathbf{x}) \in Y\}$: c'est le temps de premier retour à Y, partout fini. L'application $T_Y(\mathbf{x}) := T^{\varphi_Y(\mathbf{x})}(\mathbf{x})$ est l'application induite par T sur Y. On va montrer que T_Y est Gibbs-Markov, en construisant une partition appropriée.

Si I est un intervalle de S^1, on notera abusivement $I \times [X_{n+1}, X_n]$ pour $\{(\omega, x) \mid \omega \in I, x \in [X_{n+1}(\omega), X_n(\omega)]\}$.

Soit $I_n(\omega) = [X_{n+1}(\omega), X_n(\omega)]$ (ou $\{\omega\} \times [X_{n+1}(\omega), X_n(\omega)]$, en fonction du contexte). Par définition de X_n, T envoie $\{\omega\} \times I_n(\omega)$ bijectivement sur $\{F\omega\} \times I_{n-1}(F\omega)$. Ainsi, l'intervalle $I_n(\omega)$ retourne à $[1/2, 1]$ en exactement n itérations.

Soit $Y_n(\omega)$ la préimage dans $[1/2, 1]$ de $X_{n-1}(F\omega)$ par $T_{\alpha(\omega)}$. Ainsi, l'intervalle $J_n(\omega) = [Y_{n+1}(\omega), Y_n(\omega)]$ retourne à $[1/2, 1]$ en n étapes.

On fixe une fois pour toutes $0 < \varepsilon_0 < \frac{1}{8}$, assez petit pour que $D\varepsilon_0$ minore la longueur de tout intervalle $I_1(\omega)$. (Cette condition sera utile dans les estimées de distorsion).

Soit q assez grand pour que $\frac{1}{4^q} < \varepsilon_0$, et considérons $A_{s,n} = \left[\frac{s}{4^{q+n}}, \frac{s+1}{4^{q+n}}\right] \times J_n$, pour $n \in \mathbb{N}^*$ et $0 \leqslant s \leqslant 4^{q+n} - 1$: cet ensemble est envoyé par T^n sur $\left[\frac{s}{4^q}, \frac{s+1}{4^q}\right] \times [1/2, 1]$. Soient $K_0, \dots, K_{4^q-1}$ les ensembles $\left[\frac{i}{4^q}, \frac{i+1}{4^q}\right] \times [1/2, 1]$. Alors T_Y est un isomorphisme entre chaque $A_{s,n}$ et l'un des K_i. Par conséquent, l'application T_Y est markovienne pour la partition $\{A_{s,n}\}$, et elle a la propriété de grande image.

Pour obtenir une mesure invariante pour T_Y (et donc pour T), on va vérifier que T_Y est Gibbs-Markov. On aura donc besoin de dilatation et de contrôle de la distorsion. La dilatation est donnée par la proposition qui suit, tandis que la distorsion est estimée dans le prochain paragraphe.

Sur les intervalles $[X_3(\omega), X_1(\omega)]$, la dérivée de $T_{\alpha(\omega)}$ est strictement plus grande que 1. Elle est donc minorée par une constante $\lambda \in]1, 2[$, indépendante de ω.

Pour (ω_1, x_1) et $(\omega_2, x_2) \in S^1 \times [0, 1]$, posons

$$d'((\omega_1, x_1), (\omega_2, x_2)) = a|x_1 - x_2| + |\omega_1 - \omega_2| \tag{1.30}$$

où $a = \frac{1-\lambda/4}{D}$.

Proposition 1.5.6. *Sur chaque $A_{s,n}$, l'application T^n dilate la distance d' d'au moins λ.*

Démonstration. Pour $n = 1$ (les points reviennent directement dans Y), tout est linéaire et le résultat est clair. Supposons $n \geqslant 2$.

Soient (ω_1, x_1) et $(\omega_2, x_2) \in A_{s,n}$, avec par exemple $x_2 \geqslant x_1$. Les points (ω_1, x_1) et (ω_2, x_1) retournent à $S^1 \times [1/2, 1]$ après au moins n itérations (par hypothèse pour le premier point, et le second est en dessous de (ω_2, x_2)). On peut utiliser $n - 1$ fois le corollaire 1.5.5, et obtenir qu'en distance verticale, $\mathrm{d}_{\mathrm{vert}}(T^n(\omega_1, x_1), T^n(\omega_2, x_1)) \leqslant D|F^n\omega_1 - F^n\omega_2|$. En particulier,

$T^n_{\omega_2}(x_1) \geqslant T^n_{\omega_1}(x_1) - D\varepsilon_0 \geqslant 1/2 - D\varepsilon_0$. Ainsi, par définition de ε_0, $T^n(\omega_2, x_1) \in I_i(F^n\omega_2)$ pour $i = 0$ ou 1, donc $T^{n-1}(\omega_2, x_1) \in [X_3(F^{n-1}\omega_2), X_1(F^{n-1}\omega_2)]$. Notons que $T^{n-1}(\omega_2, x_2)$ appartient au même intervalle (en fait, $T^{n-1}_{\omega_2}(x_2) \in [X_2(F^{n-1}\omega_2), X_1(F^{n-1}\omega_2)]$). De plus, les T_α sont dilatants, donc $\mathrm{d}_{\mathrm{vert}}(T^{n-1}(\omega_2, x_1), T^{n-1}(\omega_2, x_2)) \geqslant |x_1 - x_2|$. On applique T une fois de plus, ce qui dilate d'au moins λ sur $[X_3(F^{n-1}\omega_2), X_1(F^{n-1}\omega_2)]$ par définition de λ, et on obtient $\mathrm{d}_{\mathrm{vert}}(T^n(\omega_2, x_1), T^n(\omega_2, x_2)) \geqslant \lambda|x_1 - x_2|$.

Finalement,

$$\begin{aligned} d'(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) &= a\,\mathrm{d}_{\mathrm{vert}}(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) + |F^n\omega_1 - F^n\omega_2| \\ &\geqslant a\,\mathrm{d}_{\mathrm{vert}}(T^n(\omega_2, x_1), T^n(\omega_2, x_2)) - a\,\mathrm{d}_{\mathrm{vert}}(T^n(\omega_1, x_1), T^n(\omega_2, x_1)) \\ &\quad + |F^n\omega_1 - F^n\omega_2| \\ &\geqslant a\lambda|x_1 - x_2| - aD|F^n\omega_1 - F^n\omega_2| + |F^n\omega_1 - F^n\omega_2|. \end{aligned}$$

La proposition sera prouvée si $(1 - aD)|F^n\omega_1 - F^n\omega_2| \geqslant \lambda|\omega_1 - \omega_2|$. Mais

$$(1 - aD)|F^n\omega_1 - F^n\omega_2| = (1 - aD)4^n|\omega_1 - \omega_2| \geqslant (1 - aD)4|\omega_1 - \omega_2| = \lambda|\omega_1 - \omega_2|. \qquad \square$$

Dans la suite, on aura également besoin du lemme suivant qui garantit que T vérifie l'hypothèses 4 de la définition des applications localement Gibbs-Markov.

Lemme 1.5.7. *Il existe une constante C telle que, pour tout ensemble $A_{s,n}$, pour tout $k \leqslant n$, pour presque tous $x, y \in A_{s,n}$, $d(T^k x, T^k y) \leqslant C d(T^n x, T^n y)$.*

Démonstration. Soient (ω_1, x_1) et $(\omega_2, x_2) \in A_{s,n}$ avec par exemple $x_2 \geqslant x_1$. Cela implique que, pour $0 \leqslant k \leqslant n$, $d(T^k(\omega_1, x_1), T^k(\omega_2, x_2)) \leqslant (1 + D)|F^k\omega_1 - F^k\omega_2|$. De plus, $d(T^k(\omega_2, x_1), T^k(\omega_2, x_2)) \leqslant d(T^n(\omega_2, x_1), T^n(\omega_2, x_2))$ (puisque, à ω fixé, l'application $T_{\alpha(\omega)}$ est dilatante).

Ainsi, pour $0 \leqslant k \leqslant n$,

$$\begin{aligned} d(T^k(\omega_1, x_1), T^k(\omega_2, x_2)) &\leqslant d(T^k(\omega_1, x_1), T^k(\omega_2, x_1)) + d(T^k(\omega_2, x_1), T^k(\omega_2, x_2)) \\ &\leqslant (1 + D)|F^k\omega_1 - F^k\omega_2| + d(T^n(\omega_2, x_1), T^n(\omega_2, x_2)) \\ &\leqslant (1 + D)|F^n\omega_1 - F^n\omega_2| + d(T^n(\omega_1, x_1), T^n(\omega_2, x_1)) \\ &\quad + d(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) \\ &\leqslant (1 + D)|F^n\omega_1 - F^n\omega_2| + (1 + D)|F^n\omega_1 - F^n\omega_2| \\ &\quad + d(T^n(\omega_1, x_1), T^n(\omega_2, x_2)) \\ &\leqslant (3 + 2D)d(T^n(\omega_1, x_1), T^n(\omega_2, x_2)). \end{aligned}$$

$\square$

1.5.3 Bornes sur la distorsion

Lemme 1.5.8. *Il existe une constante $E > 0$ telle que*

$$\begin{aligned}\forall n > 0, \forall \omega_1, \omega_2 \in S^1 \text{ avec } |\omega_1 - \omega_2| \leqslant \frac{\varepsilon_0}{4^n}, \forall x_1 \in J_n(\omega_1) \text{ avec } T^{n-1}_{\omega_2} x_1 \leqslant 1/2,\\ \left|\ln(T^n_{\omega_1})'(x_1) - \ln(T^n_{\omega_2})'(x_1)\right| \leqslant E|F^n\omega_1 - F^n\omega_2|.\end{aligned} \tag{1.31}$$

Démonstration. On utilise le corollaire 1.5.5 n fois et on obtient que $|T^k_{\omega_1} x_1 - T^k_{\omega_2} x_1| \leqslant D|F^k\omega_1 - F^k\omega_2|$ pour $0 \leqslant k \leqslant n$.

En particulier, pour $k = n$, $|T^n_{\omega_1} x_1| \geqslant 1/2$, donc $|T^n_{\omega_2} x_1| \geqslant 1/2 - D\varepsilon_0$. Ainsi, $T^n(\omega_2, x_1) \in I_i(F^n\omega_2)$ pour un certain $i \in \{0, 1\}$, par définition de ε_0. Une récurrence inverse donne $T^k(\omega_2, x_1) \in I_{n-k+i}(F^k\omega_2)$.

Pour $x \leqslant 1/2$ et $\omega \in S^1$, posons $G(\omega, x) = \ln T'_{\alpha(\omega)}(x) = \ln\left(1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)}\right)$. Alors

$$\frac{\partial G}{\partial x}(\omega, x) = \frac{(\alpha(\omega) + 1)\alpha(\omega)2^{\alpha(\omega)}x^{\alpha(\omega)-1}}{1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)}} \leqslant Cx^{\alpha_{\min}-1}$$

et

$$\left|\frac{\partial G}{\partial \omega}(\omega, x)\right| = \left|\frac{\alpha'(\omega)(2x)^{\alpha(\omega)} + (\alpha(\omega) + 1)\alpha'(\omega)\ln(2x)(2x)^{\alpha(\omega)}}{1 + (\alpha(\omega) + 1)(2x)^{\alpha(\omega)}}\right| \leqslant C.$$

Le lemme 1.5.2 et le fait que $T^k(\omega_1, x_1) \in I_{n-k}(F^k\omega_1)$ et $T^k(\omega_2, x_1) \in I_{n-k+i}(F^k\omega_2)$ avec $i \leqslant 1$ impliquent que les deuxièmes composantes de $T^k(\omega_1, x_1)$ et $T^k(\omega_2, x_1)$ sont $\geqslant \frac{1}{C(n-k+1)^{1/\alpha_{\min}}}$. Sur l'ensemble des points (ω, x) avec $x \geqslant \frac{1}{C(n-k+1)^{1/\alpha_{\min}}}$, les estimées sur les dérivées partielles de G montrent que cette fonction est $C(n-k+1)^{1/\alpha_{\min}-1}$-lipschitzienne, si bien que

$$\begin{aligned}|G(T^k(\omega_1, x_1)) - G(T^k(\omega_2, x_1))| &\leqslant C(n-k+1)^{1/\alpha_{\min}-1} d((T^k(\omega_1, x_1), T^k(\omega_2, x_1))\\ &\leqslant C(n-k+1)^{1/\alpha_{\min}-1}(1 + D)|F^k\omega_1 - F^k\omega_2|\\ &\leqslant C(n-k+1)^{1/\alpha_{\min}-1}(1 + D)4^k|\omega_1 - \omega_2|.\end{aligned}$$

Finalement,

$$\begin{aligned}\left|\ln(T^n_{\omega_1})'(x_1) - \ln(T^n_{\omega_2})'(x_1)\right| &\leqslant \sum_{k=0}^{n-1} |G(T^k(\omega_1, x_1)) - G(T^k(\omega_2, x_1))|\\ &\leqslant C4^n|\omega_1 - \omega_2| \sum_{k=0}^{n-1}(n-k+1)^{1/\alpha_{\min}-1}4^{k-n}\\ &\leqslant C|F^n\omega_1 - F^n\omega_2| \sum_{l=1}^{\infty}(l+1)^{1/\alpha_{\min}-1}4^{-l}.\end{aligned}$$

La dernière somme est finie, ce qui conclut la preuve. □

Posons $I_n^+(\omega) = [X_{n+2}(\omega), X_n(\omega)]$ et $J_n^+(\omega) = [Y_{n+2}(\omega), Y_n(\omega)]$ pour $n \geqslant 1$. Ainsi, $J_{n+1}^+(\omega)$ est la préimage de $I_n^+(F\omega)$. Ces intervalles vont apparaître naturellement dans les contrôles de distorsion, puisqu'on a vu dans la preuve du lemme 1.5.8 que, si on s'écarte horizontalement d'un point de $J_n(\omega_1)$, on trouve un point de $J_{n+i}(\omega_2)$ avec $i \in \{0, 1\}$, i.e. de $J_n^+(\omega_2)$.

Lemme 1.5.9. *Il existe une constante C telle que*

$$\forall n > 0, \forall \omega \in S^1, \forall x, y \in J_n^+(\omega),\ |\ln(T_\omega^n)'(x) - \ln(T_\omega^n)'(y)| \leqslant C|T_\omega^n(x) - T_\omega^n(y)|.$$

Démonstration. On rappelle que la dérivée schwarzienne d'un difféomorphisme croissant g de classe C^3 est $Sg(x) = \frac{g'''(x)}{g'(x)} - \frac{3}{2}\left(\frac{g''(x)}{g'(x)}\right)^2$. La composée de deux fonctions avec une dérivée schwarzienne négative a encore une dérivée schwarzienne négative.

Si $I \subset I'$ sont deux intervalles et $\tau > 0$, on dit que I' contient un voisinage τ-dilaté de I si les intervalles à gauche et à droite de I dans I' sont de taille au moins $\tau|I|$. Pour $\tau > 0$, le principe de Koebe ([dMvS93, théorème IV.1.2]) affirme que, si $Sg \leqslant 0$, et $J \subset J'$ sont deux intervalles tels que $g(J')$ contienne un voisinage τ-dilaté de $g(J)$, alors il existe une constante $K(\tau)$ telle que

$$\forall x, y \in J, |\ln g'(x) - \ln g'(y)| \leqslant K(\tau)\frac{|x-y|}{|J|}.$$

Ceci implique que la distorsion de g est bornée sur J, si bien qu'on peut remplacer la borne de droite par $K'(\tau)\frac{|g(x)-g(y)|}{|g(J)|}$.

Pour $0 < \alpha < 1$, la branche de gauche de T_α a une dérivée schwarzienne négative, puisque $T_\alpha''' < 0$ et $T_\alpha' > 0$. Soit g la composition des branches de gauche de $T_{\alpha(F^{n-1}\omega)}, \dots, T_{\alpha(F\omega)}$, et de la branche de droite de $T_{\alpha(\omega)}$. Alors, sur J_n^+, on a $T_\omega^n = g$, et g a une dérivée schwarzienne négative.

On veut voir que $|\ln(T_\omega^n)'(x) - \ln(T_\omega^n)'(y)| \leqslant C|T_\omega^n(x) - T_\omega^n(y)|$. Pour cela, on applique le principe de Koebe à $J = J_n^+$ et $J' = [1/2 + \delta, 2]$, pour δ très petit. Alors $g(J) = [X_2, 1]$ tandis que $g(J')$ contient $[\delta', 2]$ pour $\delta' > 0$, arbitrairement petit si δ est assez petit. Comme $X_2(\omega)$ est uniformément minoré, il existe une constante $\tau > 0$ (indépendante de ω et n) telle que $g(J')$ contienne un voisinage τ-dilaté de $g(J)$. Le principe de Koebe donne alors la conclusion recherchée. □

Proposition 1.5.10. *Il existe une constante C telle que, pour tout $A_{s,n}$, pour tous (ω_1, x_1) et $(\omega_2, x_2) \in A_{s,n}$,*

$$\left|\frac{\det DT^n(\omega_1, x_1)}{\det DT^n(\omega_2, x_2)} - 1\right| \leqslant Cd(T^n(\omega_1, x_1), T^n(\omega_2, x_2)).$$

Démonstration. La matrice $DT^n(\omega, x)$ est triangulaire supérieure, avec 4^n dans le coin en haut à gauche. Il faut donc montrer que

$$\left|\ln(T_{\omega_1}^n)'(x_1) - \ln(T_{\omega_2}^n)'(x_2)\right| \leqslant Cd(T^n(\omega_1, x_1), T^n(\omega_2, x_2)).$$

Supposons par exemple que $x_2 \geqslant x_1$, ce qui implique que $T^k_{\omega_2}(x_1) \leqslant 1/2$ pour $k = 0, \dots, n-1$. Le lemme 1.5.8 peut être appliqué à x_1, ω_1 et ω_2, et il donne en particulier que $x_1 \in J^+_n(\omega_2)$.

On a

$$\begin{aligned}\left|\ln(T^n_{\omega_1})'(x_1) - \ln(T^n_{\omega_2})'(x_2)\right| &\leqslant \left|\ln(T^n_{\omega_2})'(x_1) - \ln(T^n_{\omega_1})'(x_1)\right| + \left|\ln(T^n_{\omega_2})'(x_2) - \ln(T^n_{\omega_2})'(x_1)\right| \\ &\leqslant E|F^n\omega_2 - F^n\omega_1| + Cd(T^n(\omega_2, x_2), T^n(\omega_2, x_1))\end{aligned}$$

par les lemmes 1.5.8 et 1.5.9. Pour le second terme,

$$\begin{aligned}d(T^n(\omega_2, x_2), T^n(\omega_2, x_1)) &\leqslant d(T^n(\omega_2, x_2), T^n(\omega_1, x_1)) + d(T^n(\omega_1, x_1), T^n(\omega_2, x_1)) \\ &\leqslant d(T^n(\omega_2, x_2), T^n(\omega_1, x_1)) + (D+1)|F^n\omega_1 - F^n\omega_2|\end{aligned}$$

en utilisant les courbes admissibles.

Comme $|F^n\omega_1 - F^n\omega_2| \leqslant d(T^n(\omega_1, x_1), T^n(\omega_2, x_2))$, cela permet de conclure. □

1.5.4 Construction de la mesure invariante

Les estimées de distorsion et de dilatation qui précèdent montrent que T_Y est Gibbs-Markov, et qu'elle admet dont une mesure invariante, avec une densité lipschitzienne. En induisant, on trouve une mesure invariante pour T, dont la densité est lipschitzienne sur chaque ensemble $S^1 \times (X_{n+1}, X_n)$. Cependant, il n'est pas exclu qu'il y ait des discontinuités sur $S^1 \times X_n$, ce qui n'est pas surprenant puisque T lui-même a des discontinuités sur $S^1 \times \{1/2\}$, qui vont ensuite se propager aux autres X_n puisque la mesure est invariante.

Cependant, en dimension 1, Liverani, Saussol et Vaienti ([LSV99]) ont montré que la densité était continue partout, en la construisant comme élément d'un cône de fonctions continues (voir le paragraphe 1.4.1). On va démontrer un résultat analogue ici, par une méthode complètement différente :

Théorème 1.5.11. *L'application T admet une unique mesure de probabilité invariante absolument continue par rapport à la mesure de Lebesgue. De plus, cette mesure est ergodique. Finalement, la densité $h = \frac{\mathrm{dm}}{\mathrm{dLeb}}$ est lipschitzienne sur tout compact de $S^1 \times]0, 1]$.*

Démonstration. Considérons l'application T_Y induite par T sur $Y = S^1 \times]1/2, 1]$. Elle est markovienne pour la partition $\alpha = \{A_{s,n}\}$, et transitive puisque $T^2_Y(a) = Y$ pour tout $a \in \alpha$. De plus, elle est dilatante pour d' sur chaque élément de partition (proposition 1.5.6), et sa distorsion est lipschitzienne (proposition 1.5.10).

Ainsi, le corollaire 1.1.19 montre que T_Y admet une unique mesure de probabilité m_Y invariante absolument continue, qui est ergodique. De plus, sa densité h est lipschitzienne (pour la distance d', et dont pour la distance usuelle) sur chaque élément de la partition α_* engendrée par les éléments $T_Y(a)$, i.e. sur les ensembles $K_i = \left[\frac{i}{4^q}, \frac{i+1}{4^q}\right] \times [1/2, 1]$. De plus, par la proposition 1.1.24, h est majorée et minorée.

On obtient la mesure invariante pour m par la proposition 1.3.2. Pour vérifier que cette nouvelle mesure est de masse finie, on va montrer que $\sum m_Y(\varphi_Y > n) < \infty$. Comme $\mathrm{d}m_Y$ et dLeb sont équivalentes, il suffit de le vérifier pour dLeb. On a

$$\mathrm{Leb}(\varphi_Y > n) = \mathrm{Leb}(S^1 \times [1/2, Y_{n+1}]) = \frac{1}{2}\,\mathrm{Leb}(S^1 \times [0, X_n]) \leqslant \frac{1}{2}\frac{C}{n^{1/\,\alpha_{\max}}},$$

par le lemme 1.5.2. Comme $\alpha_{\max} < 1$, cette estimée est sommable.

On sait que h est lipschitzienne sur les ensembles $\left[\frac{i}{4^q}, \frac{i+1}{4^q}\right] \times [1/2, 1]$. Il n'est pas difficile de montrer la continuité en $\{i/4^q\} \times [1/2, 1]$: ces nombres $i/4^q$ sont artificiels, puisqu'ils dépendent du choix d'une partition de Markov pour F sur S^1. On peut faire la même construction en utilisant d'autres ensembles que les $A_{s,n}$. Par exemple, soient $A'_{s,n} = \left[\frac{1}{3} + \frac{s}{4^{q+n}}, \frac{1}{3} + \frac{s+1}{4^{q+n}}\right] \times J_n$, et $K'_i = \left[\frac{1}{3} + \frac{i}{4^q}, \frac{1}{3} + \frac{i+1}{4^q}\right]$. Comme $1/3$ est un point fixe de F, l'application T_Y est markovienne pour la partition $\{A'_{s,n}\}$, et chacun de ces ensembles est envoyé sur l'un des ensembles K'_i. Ainsi, les mêmes arguments s'appliquent, et montrent que h est lipschitzienne sur chaque ensemble K'_i. Comme les frontières des ensembles K_i et K'_i sont disjointes, on en déduit que h est en fait lipschitzienne sur $S^1 \times [1/2, 1]$.

On montre maintenant que h est lipschitzienne sur $S^1 \times [X_2, 1]$. Notons que h n'est définie que presque partout, si bien qu'il est légèrement incorrect de dire que h est lipschitzienne. Cependant, si on montre que $|h(\mathbf{x}) - h(\mathbf{y})| \leqslant Cd(\mathbf{x}, \mathbf{y})$ pour presque tous $\mathbf{x}$ et $\mathbf{y}$, alors il existera une unique version de h qui sera vraiment lipschitzienne. Ainsi, toutes les égalités qu'on écrira jusqu'à la fin de cette preuve ne seront vraies que presque partout, ce qui ne posera pas de problème.

Soit $A^+_{s,n} = \left[\frac{s}{4^{q+n}}, \frac{s+1}{4^{q+n}}\right] \times J^+_n : T^n$ est un difféomorphisme entre $A^+_{s,n}$ et $K^+_i = \left[\frac{i}{4^q}, \frac{i+1}{4^q}\right] \times [X_2, 1]$. On fixe K^+ égal à l'un des $K^+_i = I \times [X_2, 1]$, et on va montrer que h est lipschitzienne sur K^+. Notons $A^+_{s_1,n_1}, A^+_{s_2,n_2}, \ldots$ les ensembles $A^+_{s,n}$ dont l'image par T^n est K^+_i. Soit ensuite $U_j : K^+_i \to A^+_{s_j,n_j}$ l'inverse de la restriction de T^{n_j} à $A^+_{s_j,n_j}$. Soit T_Y l'application induite par T sur $Y = S^1 \times [1/2, 1]$. Alors $h\,\mathrm{dLeb}_{|Y}$ est invariante par T_Y, ce qui signifie que, pour tout $\mathbf{x} \in I \times [1/2, 1]$,

$$h(\mathbf{x}) = \sum JU_j(\mathbf{x}) h(U_j\mathbf{x})$$

où JU_j est le jacobien de U_j.

Soit $Z = S^1 \times [X_2, 1]$, et T_Z l'application induite par T sur Z. Comme $h\,\mathrm{dLeb}_{|Z}$ est également invariante par T_Z, on a le même type d'égalité que ci-dessus. Pour $\mathbf{x} \in I \times [X_2, 1/2]$, toutes ses préimages par T_Z sont dans $S^1 \times [1/2, 1]$, et l'invariance donne

$$h(\mathbf{x}) = \sum JU_j(\mathbf{x}) h(U_j\mathbf{x}).$$

On a montré que, pour tout $\mathbf{x} \in S^1 \times [X_2, 1]$,

$$h(\mathbf{x}) = \sum JU_j(\mathbf{x}) h(U_j\mathbf{x}).$$

Cela signifie que h est invariante par une sorte d'opérateur de transfert (même si ce n'est pas un vrai opérateur de transfert, puisque les images des applications U_j ne sont pas disjointes et puisqu'elles ne recouvrent pas tout l'espace). En particulier, les images des U_j sont incluses dans $S^1 \times [1/2, 1]$, et on sait déjà que h y est lipschitzienne.

Les bornes des paragraphes précédents s'appliquent encore à la distorsion des U_j et à leur dilatation. En particulier, $\left|1 - \frac{JU_j(\mathbf{y})}{JU_j(\mathbf{x})}\right| \leqslant Cd(\mathbf{x}, \mathbf{y})$ pour une constante C indépendante de j, et $|h(U_j\mathbf{x}) - h(U_j\mathbf{y})| \leqslant Cd(U_j\mathbf{x}, U_j\mathbf{y}) \leqslant C'd(\mathbf{x}, \mathbf{y})$ puisque h est lipschitzienne sur l'image de U_j. Ainsi,

$$\begin{aligned}
|h(\mathbf{x}) - h(\mathbf{y})| &\leqslant \sum |JU_j(\mathbf{x})h(U_j\mathbf{x}) - JU_j(\mathbf{y})h(U_j\mathbf{y})| \\
&\leqslant \sum |JU_j(\mathbf{x})| \left|1 - \frac{JU_j(\mathbf{y})}{JU_j(\mathbf{x})}\right| |h(U_j\mathbf{x})| + \sum |JU_j(\mathbf{y})||h(U_j\mathbf{x}) - h(U_j\mathbf{y})| \\
&\leqslant Cd(\mathbf{x}, \mathbf{y}) \sum |JU_j(\mathbf{x})| + C'd(\mathbf{x}, \mathbf{y}) \sum |JU_j(\mathbf{y})|.
\end{aligned}$$

Il reste à prouver que $\sum |JU_j(\mathbf{x})|$ est borné. Comme la distorsion est bornée, on a $JU_j(\mathbf{x}) \asymp \mathrm{Leb}(\mathrm{Im}\, U_j)$, si bien que $\sum JU_j(\mathbf{x}) \leqslant C \sum \mathrm{Leb}(\mathrm{Im}\, U_j)$, qui est fini puisque chaque point de $S^1 \times [1/2, 1]$ est dans l'image de au plus deux applications U_j.

On a montré que h était lipschitzienne sur $S^1 \times [X_2, 1]$, sauf peut-être sur $\{\frac{i}{4^q}\} \times [X_2, 1]$. Comme ci-dessus, en utilisant une autre partition de Markov, on montre que h est continue sur cet ensemble. Ainsi, h est lipschitzienne sur $S^1 \times [X_2, 1]$.

Pour montrer que h est lipschitzienne sur $S^1 \times [X_k, 1]$, on raisonne exactement de la même façon, sauf qu'on considère $[Y_{n+k}, Y_n]$ au lieu de $J_n^+ = [Y_{n+2}, Y_n]$. Comme ci-dessus, en notant $U_1, U_2, \dots$ les inverses des branches de T^n définies sur un ensemble $[\frac{s}{4^{n+q}}, \frac{s+1}{4^{n+q}}] \times [Y_{n+k}, Y_n]$ et dont l'image est $K' = [\frac{i}{4^q}, \frac{i+1}{4^q}] \times [X_k, 1] = I \times [X_k, 1]$, on montre que $h(\mathbf{x}) = \sum JU_j(\mathbf{x})h(U_j\mathbf{x})$ pour $\mathbf{x} \in K'$. En fait, pour $\mathbf{x} \in I \times [X_l, X_{l-1}]$, on utilise l'invariance de $h\,\mathrm{dLeb}$ par l'application induite par T sur $S^1 \times [X_l, 1]$. On conclut finalement comme plus haut, en utilisant le fait que h est lipschitzienne sur $S^1 \times [1/2, 1]$, qui contient l'image de tous les U_j.

Cela termine la preuve, puisque tout compact de $S^1 \times]0, 1]$ est inclus dans $S^1 \times [X_k, 1]$ pour k assez grand. □

1.5.5 Décorrélations multiples de l'application F, et bornitude L^p

Dans ce paragraphe, on expose divers résultats techniques concernant l'application $F : S^1 \to S^1$ donnée par $F(\omega) = 4\omega$, qui nous seront utiles plus tard pour estimer précisément X_n.

Théorème 1.5.12. *Soit $F : \omega \to 4\omega$ sur le cercle S^1. Alors, pour tout $p \in [1, \infty[$, il existe une constante K_p telle que, pour tout $n \in \mathbb{N}$, pour toutes fonctions $f_0, \dots, f_{n-1} : S^1 \to \mathbb{R}$*

bornées par 1, *de moyenne nulle et* 1*-lipschitziennes,*

$$\left\| \sum_{k=0}^{n-1} f_k \circ F^k \right\|_p \leqslant K_p \sqrt{n}.$$

Ce résultat a été essentiellement prouvé par Françoise Pène, dans un contexte beaucoup plus général. Sa preuve dépend d'une propriété de décorrélations multiples, qui est impliquée par le trou spectral de l'opérateur de transfert :

Lemme 1.5.13. *Soit* $\|f\|$ *la norme de Lipschitz de la fonction* f *sur* S^1*. Alors, pour tous* $m, m' \in \mathbb{N}$*, il existe* $C > 0$ *et* $\delta < 1$ *tels que, pour tout* $N \in \mathbb{N}$*, pour toutes suites croissantes* $(k_1, \dots, k_m)$ *et* $(l_1, \dots, l_{m'})$*, pour toutes fonctions lipschitziennes* $G_1, \dots, G_m, H_1, \dots, H_{m'}$*,*

$$\left| \mathrm{Cov}\left(\prod_{i=1}^{m} G_i \circ F^{k_i}, \prod_{j=1}^{m'} H_j \circ F^{N+l_j} \right) \right| \leqslant C \left(\prod_{i=1}^{m} \|G_i\| \right) \left(\prod_{j=1}^{m'} \|H_j\| \right) \delta^{N-k_m}. \tag{1.32}$$

Ici, $\mathrm{Cov}(u, v) = \int uv - \int u \int v$.

Démonstration. Soit $\widehat{F}$ l'opérateur de transfert associé à F, agissant sur les fonctions lipschitziennes. Il admet un trou spectral et ses itérés sont bornés, c'est-à-dire qu'il existe des constantes $M > 0$ et $\delta < 1$ telles que $\left\|\widehat{F}^n f\right\| \leqslant M \|f\|$, et $\left\|\widehat{F}^n f\right\| \leqslant M\delta^n \|f\|$ si $\int f = 0$.
On peut supposer que $N \geqslant k_m$ (sinon, $\delta^{N-k_m} \geqslant 1$, et l'inégalité (1.32) devient triviale). Alors, en écrivant $\varphi = \prod_{i=1}^{m} G_i \circ F^{k_i}$ et $\psi = \prod_{j=1}^{m'} H_j \circ F^{l_j}$, on obtient

$$\begin{aligned} \left|\mathrm{Cov}(\varphi, \psi \circ F^N)\right| &= \left| \int \left(\varphi - \int \varphi \right) \psi \circ F^N \right| = \left| \int \widehat{F}^N \left(\varphi - \int \varphi \right) \psi \right| \\ &\leqslant \left\| \widehat{F}^N \left(\varphi - \int \varphi \right) \right\| \|\psi\|_\infty . \end{aligned}$$

Mais

$$\begin{aligned} \widehat{F}^N(\varphi) &= \widehat{F}^N \left(\prod G_i^{k_i} \right) = \widehat{F}^{N-k_m}(G_m \widehat{F}^{k_m - k_{m-1}}(G_{m-1} \widehat{F}^{k_{m-1}-k_{m-2}}(\dots \widehat{F}^{k_2-k_1}(G_1)) \dots) \\ &=: \widehat{F}^{N-k_m}(\chi). \end{aligned}$$

Comme les itérés de $\widehat{F}$ sont bornés sur les fonctions lipschitziennes, on obtient une borne sur la norme de Lipschitz de χ : $\|\chi\| \leqslant M^{m-1} \prod \|G_i\|$. De plus, $\int \chi = \int \varphi$, donc

$$\begin{aligned} \left\| \widehat{F}^N \left(\varphi - \int \varphi \right) \right\| &= \left\| \widehat{F}^{N-k_m} \left(\chi - \int \chi \right) \right\| \leqslant M\delta^{N-k_m} \left\| \chi - \int \chi \right\| \\ &\leqslant M\delta^{N-k_m} M^{m-1} \prod \|G_i\| . \end{aligned}$$

□

Quand p est un entier pair, le théorème 1.5.12 est alors une conséquence de [Pèn02, lemme 2.3.4]. L'inégalité de Hölder permet d'en déduire le cas général.

Remarque 1.5.14. On a le même résultat pour des fonctions seulement hölderiennes, avec la même preuve.

On aura aussi besoin du résultat suivant :

Théorème 1.5.15. *Soient X un espace muni d'une mesure de probabilité et $T : X \to X$ une application qui préserve cette mesure. Soient $f : X \to \mathbb{R}$ et $p > 2$ tels que*

$$\exists C > 0, \forall n \in \mathbb{N}^*, \|S_n f\|_p \leqslant C\sqrt{n}.$$

Notons $M_n f(x) = \sup_{1\leqslant k\leqslant n} |S_k f(x)|$. Alors il existe une constante K telle que

$$\forall n \geqslant 2, \|M_n f\|_p \leqslant K(\ln n)^{\frac{p-1}{p}}\sqrt{n}.$$

Démonstration. Soit $n \in \mathbb{N}^*$. Soit $k < 2^n$, écrivons sa décomposition binaire $k = \sum_{j=0}^{n-1} \varepsilon_j 2^j$, avec $\varepsilon_j \in \{0, 1\}$. Soit $q_j = \sum_{l=j}^{n-1} \varepsilon_l 2^l$; en particulier, $q_0 = k$ et $q_n = 0$. Alors $S_k f = \sum_{j=0}^{n-1}(S_{q_j} f - S_{q_{j+1}} f)$. L'inégalité de convexité $(a_0 + \ldots + a_{n-1})^p \leqslant n^{p-1}(a_0^p + \ldots + a_{n-1}^p)$ donne alors

$$|S_k f|^p \leqslant n^{p-1} \sum_{j=0}^{n-1} |S_{q_j} f - S_{q_{j+1}} f|^p.$$

Notons que q_{j+1} est de la forme $\lambda 2^{j+1}$ avec $0 \leqslant \lambda \leqslant 2^{n-j-1} - 1$, et q_j est égal à q_{j+1} ou $q_{j+1} + 2^j$. Ainsi,

$$|S_k f|^p \leqslant n^{p-1} \sum_{j=0}^{n-1} \left(\sum_{\lambda=0}^{2^{n-j-1}-1} |S_{\lambda 2^{j+1}+2^j} f - S_{\lambda 2^{j+1}} f|^p \right).$$

Le membre de droite est indépendant de k. Il borne donc $|M_{2^n-1} f|^p$. De plus,

$$\int |S_{\lambda 2^{j+1}+2^j} f - S_{\lambda 2^{j+1}} f|^p = \int |S_{2^j} f|^p \leqslant C^p \sqrt{2^j}^p.$$

On obtient

$$\int |M_{2^n-1} f|^p \leqslant n^{p-1} \sum_{j=0}^{n-1} 2^{n-j-1} C^p 2^{pj/2} \leqslant K n^{p-1} 2^n 2^{(\frac{p}{2}-1)n} = K n^{p-1} \sqrt{2^n}^p.$$

Pour des temps de la forme $2^n - 1$, c'est une borne de la forme $\|M_t\|_p \leqslant K(\ln t)^{\frac{p-1}{p}}\sqrt{t}$. Pour obtenir la même estimée pour un temps t arbitraire, il suffit de choisir n avec $2^{n-1} \leqslant t < 2^n$, et de noter que $M_t \leqslant M_{2^n-1}$. □

Corollaire 1.5.16. *Soient $F : \omega \to 4\omega$ sur le cercle S^1, $\chi : S^1 \to \mathbb{R}$ une fonction hölderienne de moyenne nulle, et $p > 2$. Soit $M_n\chi(x) = \sup_{1\leqslant k\leqslant n} |S_k\chi(x)|$. Il existe alors une constante K telle que*

$$\forall n \geqslant 2, \|M_n\chi\|_p \leqslant K(\ln n)^{\frac{p-1}{p}}\sqrt{n}.$$

Démonstration. Le théorème 1.5.12 (ou plutôt la remarque 1.5.14, pour le cas hölderien) montre que $\|S_n\chi\|_p \leqslant C\sqrt{n}$. Par conséquent, le théorème 1.5.15 permet de conclure. □

1.5.6 Comportement asymptotique de X_n

Le but de cette section est d'estimer précisément le comportement de X_n. Cela permettra de trouver un équivalent de $m(\varphi_Y > n)$, ce qui sera utile pour obtenir des théorèmes limites (théorème 4.2.6).

Théorème 1.5.17. *On a*

$$\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\alpha_{min}} X_n(\omega) \to \frac{1}{\left(2^{\alpha_{min}}\,{\alpha_{min}}^{3/2}\sqrt{\frac{\pi}{2\alpha''(x_0)}}\right)^{1/\alpha_{min}}}$$

presque partout et dans L^1.

On en déduit le comportement asymptotique de $m(\varphi_Y > n)$. On notera

$$A = \frac{1}{4\left({\alpha_{\min}}^{3/2}\sqrt{\frac{\pi}{2\alpha''(x_0)}}\right)^{1/\alpha_{\min}}} \int_{S^1\times\{1/2\}} h \,\mathrm{dLeb}, \tag{1.33}$$

où h est la densité de la mesure de probabilité invariante m.

Proposition 1.5.18. *On a*

$$m(\varphi_Y > n) \sim \left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{min}} A.$$

Démonstration. On a

$$\begin{aligned} m(\varphi_Y > n) &= \int_{S^1}\int_{1/2}^{Y_{n+1}(\omega)} h(\omega,u)\,\mathrm{d}u\,\mathrm{d}\omega = \int_{S^1}\int_0^{X_n(F\omega)/2} h(\omega,1/2+u)\,\mathrm{d}u\,\mathrm{d}\omega \\ &= \int_{S^1}\frac{X_n(F\omega)}{2}h(\omega,1/2)\,\mathrm{d}\omega + \int_{S^1}\int_0^{X_n(F\omega)/2}\big[h(\omega,1/2+u)-h(\omega,1/2)\big]\,\mathrm{d}u\,\mathrm{d}\omega \\ &= I + II. \end{aligned}$$

Comme $\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\alpha_{\min}} X_n(F\omega) \to \frac{1}{\left(2^{\alpha_{\min}}\,{\alpha_{\min}}^{3/2}\sqrt{\frac{\pi}{2\alpha''(x_0)}}\right)^{1/\alpha_{\min}}}$ dans L^1 et presque partout (théorème 1.5.17) et $h(\omega,1/2)$ est borné, $I \sim \left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}} A$. De plus, si n est assez grand, $|h(\omega,1/2+u)-h(\omega,1/2)| \leqslant \varepsilon$, donc $II = o\left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}}$. □

La preuve du théorème 1.5.17 utilisera le lemme suivant.

Lemme 1.5.19. *On a*

$$E(e^{-(\alpha-\alpha_{min})w}) \sim \sqrt{\frac{\pi}{2\alpha''(x_0)}}\frac{1}{\sqrt{w}} \text{ quand } w \to \infty. \tag{1.34}$$

Ici, $E(e^{-(\alpha-\alpha_{\min})w})$ désigne $\int_{S^1} e^{-(\alpha(\omega)-\alpha_{\min})w}\,\mathrm{d}\omega$.

Démonstration. Soient $\beta = \alpha - \alpha_{\min}$, et $f(b) = \mathrm{Leb}\{\omega \mid \beta(\omega) \in [0,b)\}$. Dans un voisinage de x_0 l'unique point où α atteint son minimum $\alpha_{\min}$, α se comporte comme la parabole $\alpha_{\min} + \frac{\alpha''(x_0)}{2}(x-x_0)^2$, donc $f(b) \sim \sqrt{\frac{2}{\alpha''(x_0)}}\sqrt{b}$ quand $b \to 0$.
En notant P_β la distribution de β, une intégration par parties donne

$$\begin{aligned} E\left(e^{-(\alpha-\alpha_{\min})w}\right) &= \int_0^\infty e^{-bw}\,\mathrm{d}P_\beta(b) = w\int_0^\infty e^{-bw} f(b)\,\mathrm{d}b = \int_0^\infty e^{-u} f(u/w)\,\mathrm{d}u \\ &= \frac{1}{\sqrt{w}}\int_0^\infty e^{-u}\left(\sqrt{w} f(u/w)\right)\,\mathrm{d}u. \end{aligned}$$

Mais $e^{-u}\left(\sqrt{w} f(u/w)\right) \to e^{-u}\sqrt{\frac{2}{\alpha''(x_0)}}\sqrt{u}$ quand $w \to \infty$. Il existe une constante E telle que $f(u) \leqslant E\sqrt{u}$: c'est clair sur un voisinage de 0, et ailleurs car f est bornée. Ainsi, $e^{-u}\left(\sqrt{w} f(u/w)\right) \leqslant E e^{-u}\sqrt{u}$ intégrable. Par convergence dominée,

$$\int_0^\infty e^{-u}\left(\sqrt{w} f(u/w)\right)\,\mathrm{d}u \to \sqrt{\frac{2}{\alpha''(x_0)}}\int_0^\infty e^{-u}\sqrt{u}\,\mathrm{d}u = \sqrt{\frac{2}{\alpha''(x_0)}}\frac{\sqrt{\pi}}{2}.$$ □

Preuve du théorème 1.5.17. Comme dans la proposition 1.5.2, on écrit

$$\frac{1}{X_n(F\omega)^{\alpha_{\min}}} = \frac{1}{X_{n+1}(\omega)^{\alpha_{\min}}} - \alpha_{\min}\,2^{\alpha_{\min}}(2X_{n+1}(\omega))^{\alpha(\omega)-\alpha_{\min}} + O(X_{n+1}(\omega)^{2\alpha(\omega)-\alpha_{\min}}).$$

La proposition 1.5.2 donne

$$X_{n+1}(\omega)^{2\alpha(\omega)-\alpha_{\min}} \leqslant X_{n+1}(\omega)^{\alpha_{\min}} \leqslant \frac{C}{(n+1)^{\alpha_{\min}/\alpha_{\max}}} \leqslant \frac{C}{\sqrt{n+1}}$$

car $\alpha_{\min} / \alpha_{\max} \geqslant 1/2$ par hypothèse. Ainsi,

$$\frac{1}{X_{n+1}(\omega)^{\alpha_{\min}}} - \frac{1}{X_n(F\omega)^{\alpha_{\min}}} = 2^{\alpha_{\min}}\, \alpha_{\min} (2X_{n+1}(\omega))^{\alpha(\omega)-\alpha_{\min}} + O(1/\sqrt{n}).$$

En sommant de 1 à n, on obtient une constante P (indépendante de ω) telle que

$$\frac{1}{X_n(\omega)^{\alpha_{\min}}} \geqslant 2^{\alpha_{\min}}\, \alpha_{\min} \left[\sum_{k=1}^{n} (2X_k(F^{n-k}\omega))^{\alpha(F^{n-k}\omega)-\alpha_{\min}} - P\sqrt{n} \right]. \tag{1.35}$$

$$\frac{1}{X_n(\omega)^{\alpha_{\min}}} \leqslant 2^{\alpha_{\min}}\, \alpha_{\min} \left[\sum_{k=1}^{n} (2X_k(F^{n-k}\omega))^{\alpha(F^{n-k}\omega)-\alpha_{\min}} + P\sqrt{n} \right]. \tag{1.36}$$

L'équation (1.35) et la proposition 1.5.2 impliquent que

$$\frac{\sqrt{\ln n}}{n} \frac{1}{2^{\alpha_{\min}}\, \alpha_{\min}\, X_n(\omega)^{\alpha_{\min}}} \geqslant \frac{\sqrt{\ln n}}{n} \sum_{k=1}^{n} \left(\frac{2C^{-1}}{k^{1/\alpha_{\min}}} \right)^{\alpha(F^{n-k}\omega)-\alpha_{\min}} - P\sqrt{\frac{\ln n}{n}} =: A_n(\omega). \tag{1.37}$$

Étudions tout d'abord la convergence de A_n. Les fonctions α et $\alpha \circ F^{n-k}$ ont la même distribution puisque F préserve la mesure de Lebesgue. Ainsi, par le lemme 1.5.19,

$$E\left(\left(\frac{2C^{-1}}{k^{1/\alpha_{\min}}} \right)^{\alpha\circ F^{n-k}-\alpha_{\min}} \right) \sim \sqrt{\frac{\pi}{2\alpha''(x_0)}} \frac{1}{\sqrt{\ln(k^{1/\alpha_{\min}}) - \ln(2C^{-1})}} \sim \sqrt{\frac{\pi\, \alpha_{\min}}{2\alpha''(x_0)}} \frac{1}{\sqrt{\ln k}}.$$

En additionnant ces estimées, on obtient

$$E(A_n) \to C_1 := \sqrt{\frac{\pi\, \alpha_{\min}}{2\alpha''(x_0)}}, \tag{1.38}$$

puisque $\sum_{k=2}^{n} \frac{1}{\sqrt{\ln k}} \sim \frac{n}{\sqrt{\ln n}}$.

On aura besoin d'estimées L^p, pour $p \geqslant 1$. On notera $\|g\|$ la norme Lipschitz d'une fonction $g : S^1 \to \mathbb{R}$.

On définit $f_k(\omega) = \left(\frac{2C^{-1}}{k^{1/\alpha_{\min}}} \right)^{\alpha(\omega)-\alpha_{\min}}$, et $g_k = f_k - E(f_k)$. Ainsi, $A_n = \frac{\sqrt{\ln n}}{n} \sum_{k=1}^{n} f_k \circ F^{n-k} - P\sqrt{\frac{\ln n}{n}}$. Comme $g_k' = \ln\left(\frac{2C^{-1}}{k^{1/\alpha_{\min}}} \right) \alpha' f_k$, il existe une constante L telle que, pour tous $k \leqslant n$, $\|g_k\| \leqslant L \ln n$. Ainsi, le théorème 1.5.12 appliqué à $g_k/(L \ln n)$ donne

$$\|A_n - E(A_n)\|_p = \frac{\sqrt{\ln n}}{n} L \ln n \left\| \sum_{k=1}^{n} g_k \circ F^{n-k}/(L \ln n) \right\|_p \leqslant \frac{\sqrt{\ln n}}{n} L \ln n K_p \sqrt{n},$$

i.e.

$$\|A_n - E(A_n)\|_p \leqslant L_p \sqrt{\frac{\ln^3 n}{n}}. \tag{1.39}$$

Cela implique en particulier que A_n converge presque partout vers C_1. En effet, pour $\delta > 0$,

$$\mathrm{Leb}\{|A_n - E(A_n)| > \delta\} \leqslant \int \frac{|A_n - E(A_n)|^4}{\delta^4} \leqslant \frac{L_4^4}{\delta^4} \left(\frac{\ln^3 n}{n}\right)^{4/2}$$

qui est sommable, et $E(A_n) \to C_1$.
On a

$$\begin{aligned} A_n(\omega) &\geqslant \frac{\sqrt{\ln n}}{n} \left[\sum_{k=1}^{n} \left(\frac{2C^{-1}}{k^{1/\alpha_{\min}}}\right)^{\alpha_{\max} - \alpha_{\min}} - P\sqrt{n}\right] \geqslant \frac{\sqrt{\ln n}}{n}\left[K n^{2-\alpha_{\max}/\alpha_{\min}} - P\sqrt{n}\right] \\ &\geqslant K' \frac{\sqrt{\ln n}}{n} n^{2-\alpha_{\max}/\alpha_{\min}} \end{aligned}$$

car $\alpha_{\max}/\alpha_{\min} < 3/2$. Ainsi,

$$\left\| \frac{1}{A_n} \right\|_\infty \leqslant K'' \frac{n^{\alpha_{\max}/\alpha_{\min} - 1}}{\sqrt{\ln n}}. \tag{1.40}$$

Comme $E(A_n)$ tend vers $C_1 \neq 0$, $\frac{1}{E(A_n)}$ est borné, donc

$$\left\| \frac{1}{A_n} - \frac{1}{E(A_n)} \right\|_p \leqslant \left\| \frac{1}{A_n} \right\|_\infty \frac{1}{E(A_n)} \|A_n - E(A_n)\|_p \leqslant K''' \frac{n^{\alpha_{\max}/\alpha_{\min} - 1}}{\sqrt{\ln n}} L_p \sqrt{\frac{\ln^3 n}{n}} = M_p \frac{\ln n}{n^\kappa},$$

où $\kappa = \frac{3}{2} - \frac{\alpha_{\max}}{\alpha_{\min}} > 0$. En particulier, $\frac{1}{A_n}$ tend vers $\frac{1}{C_1}$ dans tout L^p. L'équation (1.37) montre que

$$\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\alpha_{\min}} X_n \leqslant \frac{1}{(2^{\alpha_{\min}} \alpha_{\min} A_n)^{1/\alpha_{\min}}}. \tag{1.41}$$

Le membre de droite tend vers

$$C_2 := \frac{1}{\left(2^{\alpha_{\min}} {\alpha_{\min}}^{3/2} \sqrt{\frac{\pi}{2\alpha''(x_0)}}\right)^{1/\alpha_{\min}}} \tag{1.42}$$

dans tout L^p, et en particulier dans L^1. Ainsi,

$$\overline{\lim}\, E\left(\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\alpha_{\min}} X_n\right) \leqslant C_2. \tag{1.43}$$

De plus, A_n converge presque partout vers C_1, donc (1.41) donne que, presque partout,

$$\overline{\lim} \left(\frac{n}{\sqrt{\ln n}}\right)^{1/\alpha_{\min}} X_n(\omega) \leqslant C_2. \tag{1.44}$$

Soit $Q = \sup_n \left(\frac{1}{E(A_n)}\right) + 1$, on estime Leb $\left\{\frac{1}{A_n} \geqslant Q\right\}$. Si $p \geqslant 1$,

$$\text{Leb}\left\{\frac{1}{A_n} \geqslant Q\right\} \leqslant \text{Leb}\left\{\left|\frac{1}{A_n} - \frac{1}{E(A_n)}\right| \geqslant 1\right\} \leqslant E\left(\left|\frac{1}{A_n} - \frac{1}{E(A_n)}\right|^p\right) \leqslant \left(M_p \frac{\ln n}{n^\kappa}\right)^p.$$

En particulier, si p est assez grand, on obtient

$$\text{Leb}\left\{\frac{1}{A_n} \geqslant Q\right\} \leqslant \frac{M}{n^5}.$$

Soit $Q' = \frac{Q}{2^{\alpha_{\min}}\,\alpha_{\min}}$. L'équation (1.41) implique alors que

$$\text{Leb}\left\{X_n \geqslant \left(\frac{Q'\sqrt{\ln n}}{n}\right)^{1/\,\alpha_{\min}}\right\} \leqslant \frac{M}{n^5}. \tag{1.45}$$

Ainsi, $U_n := \left\{\omega \mid \exists \sqrt{n} \leqslant k \leqslant n \text{ avec } X_k(F^{n-k}\omega) \geqslant \left(\frac{Q'\sqrt{\ln k}}{k}\right)^{1/\,\alpha_{\min}}\right\}$ a une mesure majorée par $\sum_{\sqrt{n}}^n \frac{M}{k^5} \leqslant \frac{M'}{n^2}$ (car Leb est invariante par F^{n-k}). Finalement, le lemme de Borel-Cantelli assure qu'il existe un ensemble de mesure pleine de S^1 sur lequel $\omega \notin U_n$ pour n assez grand. Soit

$$A'_n(\omega) = \frac{\sqrt{\ln n}}{n}\left[\sum_{k=1}^n \left(\frac{2(Q'\sqrt{\ln k})^{1/\,\alpha_{\min}}}{k^{1/\,\alpha_{\min}}}\right)^{\alpha(F^{n-k}\omega)-\alpha_{\min}} + (P+1)\sqrt{n}\right].$$

Comme pour A_n, on montre que $A'_n \to C_1$ dans tout L^p et presque partout.

Soit ω tel que $\omega \notin U_n$ pour n assez grand, et $A'_n(\omega) \to C_1$ (ces propriétés sont vraies presque partout). Pour n assez grand, l'équation (1.36) et le fait que $X_k(F^{n-k}\omega) \leqslant \left(\frac{Q'\sqrt{\ln k}}{k}\right)^{1/\,\alpha_{\min}}$ pour $\sqrt{n} \leqslant k \leqslant n$ donnent

$$\begin{aligned}\frac{1}{2^{\alpha_{\min}}\,\alpha_{\min}\,X_n(\omega)^{\alpha_{\min}}} &\leqslant \left[\sum_{k=1}^{\sqrt{n}} 1 + \sum_{k=\sqrt{n}}^n \left(\frac{2(Q'\sqrt{\ln k})^{1/\,\alpha_{\min}}}{k^{1/\,\alpha_{\min}}}\right)^{\alpha(F^{n-k}\omega)-\alpha_{\min}} + P\sqrt{n}\right] \\ &\leqslant \frac{n}{\sqrt{\ln n}} A'_n(\omega) \sim \frac{n}{\sqrt{\ln n}} C_1.\end{aligned}$$

Ainsi,

$$\underline{\lim}\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\,\alpha_{\min}} X_n(\omega) \geqslant C_2. \tag{1.46}$$

Les équations (1.44) et (1.46) montrent que $\left(\frac{n}{\sqrt{\ln n}}\right)^{1/\,\alpha_{\min}} X_n$ tend presque partout vers C_2. On obtient la convergence dans L^1 grâce à l'inégalité (1.43) et au lemme suivant. □

Lemme 1.5.20. *Soient f_n des fonctions positives ou nulles sur un espace de probabilité, avec $f_n \to f$ presque partout et $\overline{\lim}\, E(f_n) \leqslant E(f) < \infty$. Alors $f_n \to f$ dans L^1.*

Démonstration. Soit $g_n = f_n + f - |f - f_n| \geqslant 0$. Le lemme de Fatou donne $E(\underline{\lim}\, g_n) \leqslant \underline{\lim}\, E(g_n)$, donc

$$2E(f) \leqslant \overline{\lim}\, E(f_n) + E(f) - \overline{\lim}\, E(|f - f_n|).$$

Ainsi, les hypothèses impliquent que $\overline{\lim}\, E(|f - f_n|) \leqslant 0$. □

Chapitre 2

Décorrélation des tours localement Gibbs-Markov

Sommaire

2.1 Opérateurs de transfert de premier retour 78
2.1.1 Définition des opérateurs 78
2.1.2 Étude spectrale de R_n 80
2.2 Algèbres de Banach et résultats spectraux 82
2.2.1 Algèbres de Banach et lemme de Wiener 82
2.2.2 Sommabilité des T_n 84
2.2.3 Construction d'une approximation analytique 87
2.2.4 Classes de vitesse de décroissance 89
2.2.5 Estimées de T_n 94
2.3 Vitesse de décorrélation 96
2.3.1 Minoration de la vitesse de décorrélation 97
2.3.2 Majoration de la vitesse de décorrélation 98
2.3.3 Exemples 100
2.4 Raffinements sur la vitesse de décorrélation 104
2.4.1 Estimation à plusieurs termes de la décorrélation 104
2.4.2 Fonctions de moyenne nulle supportées par Y 109
2.4.3 Fonctions de moyenne nulle générales 111

Dans ce chapitre, qui reprend et étend les résultats de l'article [Gou04b], on s'intéresse à la vitesse de décroissance des corrélations dans les tours localement Gibbs-Markov mélangeantes. Comme on l'a vu dans le paragraphe 1.3.3, cela aura des conséquences pour des applications localement Gibbs-Markov générales puisqu'on peut toujours ramener leur étude à l'étude de tours localement Gibbs-Markov mélangeantes.

Soit (X, T, m) une tour localement Gibbs-Markov mélangeante, de base $Y \subset X$, préservant la mesure de probabilité m. Soient $u, v : X \to \mathbb{R}$ deux fonctions hölderiennes. On définit leurs corrélations par

$$\mathrm{Cor}(u, v \circ T^n) := \int u \cdot v \circ T^n \, \mathrm{d}m - \left(\int u \, \mathrm{d}m\right)\left(\int v \, \mathrm{d}m\right). \tag{2.1}$$

Lorsque T est Gibbs-Markov, ces corrélations décroissent exponentiellement vite (corollaire 1.1.21). Dans le cas localement Gibbs-Markov, on va voir que la vitesse est plus lente, et dépend en général de la mesure des points qui mettent "longtemps" à revenir à la base. Plus précisément, pour $y \in Y$, soit $\varphi_Y(y) = \inf\{n \geqslant 1, T^n(y) \in Y\}$ le temps de retour de y à Y. On va voir que la décroissance de $\mathrm{Cor}(f, g \circ T^n)$ est essentiellement dictée par la vitesse de décroissance de $m\{y \in Y \mid \varphi_Y(y) > n\}$. Des bornes supérieures (dans le cas des tours de Young) ont été démontrées dans [You99]), mais on va s'intéresser ici à des vitesses exactes, c'est-à-dire qu'on cherchera également des bornes inférieures sur la vitesse de décorrélation, ou même des équivalents. La méthode sera essentiellement spectrale, et utilisera les opérateurs de transfert de premier retour introduits par Sarig dans [Sar02].

2.1 Opérateurs de transfert de premier retour

Dans toute cette partie, (X, T, m) sera une tour localement Gibbs-Markov mélangeante préservant la mesure de probabilité m. Comme dans le paragraphe 1.3.3, on notera α la partition de X pour laquelle T est markovienne, et $\alpha_0 \subset \alpha$ un ensemble d'éléments de partition tel que l'application induite par T sur $Y := \bigcup_{a \in \alpha_0} a$ soit Gibbs-Markov.

Quitte à remplacer la distance d sur X par d_τ définie dans la remarque 1.3.9, avec τ assez proche de 1, les fonctions hölderiennes avec un certain exposant pour la distance d initiale deviendront lipschitziennes pour la nouvelle distance. On pourra donc sans perte de généralité se restreindre à l'étude des fonctions lipschitziennes.

2.1.1 Définition des opérateurs

On notera g l'inverse du jacobien de T, donné sur un élément a de α par $g(x) = \frac{\mathrm{d}m}{\mathrm{d}(m \circ T_{|a})}(x)$. De même, soit $g^{(n)}(x) = \prod_0^{n-1} g(T^i x)$. Ainsi, l'opérateur de transfert associé à T s'écrit $\widehat{T}(u)(x) = \sum_{Ty=x} g(y)u(y)$, et

$$\widehat{T}^n(u)(x) = \sum_{T^n y = x} g^{(n)}(y)u(y). \tag{2.2}$$

On travaillera également avec l'application T_Y induite par T sur Y. Soit g_Y l'inverse de son jacobien. Ainsi, pour $y \in Y$ de temps de retour $\varphi_Y(y)$, on a $g_Y(y) = g^{(\varphi_Y(y))}(y)$. Comme X

est une tour et que la mesure m est invariante, tous les jacobiens avant le premier retour à la base sont égaux à 1, et on a donc en fait $g_Y(y) = g(T^{\varphi_Y(y)-1}y)$.

Dans l'équation (2.2), on va décomposer la somme suivant les instants auxquels les itérés de y tombent dans Y. Plus précisément, on introduit les opérateurs suivants :

$$R_n u(x) = \sum_{\substack{T^n y = x \\ y\in Y, Ty,\dots,T^{n-1}y\notin Y, T^n y\in Y}} g^{(n)}(y)u(y),$$

$$T_n u(x) = \sum_{\substack{T^n y = x \\ y\in Y, T^n y\in Y}} g^{(n)}(y)u(y),$$

$$A_n u(x) = \sum_{\substack{T^n y = x \\ y\in Y, Ty,\dots,T^n y\notin Y}} g^{(n)}(y)u(y),$$

$$B_n u(x) = \sum_{\substack{T^n y = x \\ y,\dots,T^{n-1}y\notin Y, T^n y\in Y}} g^{(n)}(y)u(y),$$

$$C_n u(x) = \sum_{\substack{T^n y = x \\ y,\dots,T^n y\notin Y}} g^{(n)}(y)u(y),$$

Les interprétations de ces opérateurs sont les suivantes : R_n prend en compte les premiers retours à Y, tandis que T_n prend tous les retours en compte, si bien qu'on peut écrire $T_n(u) = 1_Y \widehat{T}^n(1_Y u)$. En décomposant une trajectoire suivant les instants où elle tombe dans Y, on obtient l'équation suivante, appelée *équation de renouvellement* :

$$T_n = \sum_{k_1+\dots+k_l=n} R_{k_1}\dots R_{k_l}. \tag{2.3}$$

Remarque 2.1.1. Cette équation est classique en probabilités, mais avec des nombres à la place des opérateurs : typiquement, R_n est la probabilité qu'une ampoule casse au bout d'un temps n, et T_n est la probabilité qu'on doive remplacer une ampoule au temps n (en ayant éventuellement fait plusieurs remplacements auparavant).

Les opérateurs B_n et A_n tiennent compte respectivement des débuts et des fins des trajectoires, hors de Y. Ainsi, si $x \in X$ est fixé et $y \in X$ satisfait $T^n(y) = x$, on peut considérer les b premiers itérés de y, jusqu'à ce qu'il rentre dans Y (ce qui correspond à B_b), puis un certain nombre de retours successifs à Y, durant un temps k (cela correspond à T_k), et finalement a itérés hors de Y (correspondant à A_a). Ainsi,

$$\widehat{T}^n = C_n + \sum_{a+k+b=n} A_a T_k B_b. \tag{2.4}$$

L'opérateur C_n prend en compte les trajectoires qui ne rentrent jamais dans Y.

Pour étudier les corrélations $\mathrm{Cor}(u, v \circ T^n)$, qui s'écrivent également sous la forme

$$\mathrm{Cor}(u, v \circ T^n) = \int \widehat{T}^n(u) \cdot v \, \mathrm{d}m - \left(\int u \, \mathrm{d}m\right)\left(\int v \, \mathrm{d}m\right)$$

par définition de l'opérateur de transfert, on va utiliser (2.4). Comme A_a et B_b sont quasiment explicites, tout le problème va être de comprendre T_k. On va pour cela utiliser l'équation (2.3). Notons que cette équation peut se réécrire

$$\sum_{n=0}^{\infty} T_n z^n = \left(I - \sum_{n=1}^{\infty} R_n z^n\right)^{-1}. \tag{2.5}$$

Pour tirer parti de cette égalité, on aura besoin d'une description spectrale fine des opérateurs R_n.

2.1.2 Étude spectrale de R_n

Les opérateurs R_n agissent sur les fonctions définies sur Y. Comme T_Y est Gibbs-Markov, on va pouvoir utiliser les techniques spectrales concernant les opérateurs Gibbs-Markov pour étudier R_n. On notera $[a_0, \dots, a_{n-1}]_Y$ les cylindres sur Y, pour l'application T_Y : ils sont définis par $\bigcap_0^{n-1} T_Y^{-i}(a_i)$, pour $a_0, \dots, a_{n-1} \in \alpha_0$. Lorsque $\underline{a}_Y = [a_0, \dots, a_{n-1}]_Y$ est un tel cylindre, rappelons que $M_{\underline{a}_Y}$ désigne l'opérateur donné par $M_{\underline{a}_Y} u(x) = g_Y^{(n)}(y) u(y)$, où y est l'unique préimage de x dans $\underline{a}_Y$ par T_Y^n, si ce point existe, et $M_{\underline{a}_Y} u(x) = 0$ sinon. Il vérifie

$$\left\|M_{\underline{a}_Y} u\right\|_{\mathcal{L}} \leqslant Bm(\underline{a}_Y)\left(\lambda^{-n}\|u\|_{\mathcal{L}} + \frac{1}{m(\underline{a}_Y)}\int_{\underline{a}_Y} |u| \, \mathrm{d}m\right) \tag{2.6}$$

d'après le lemme 1.1.13 et la propriété de grande image. Ici, $\mathcal{L}$ désigne l'espace des fonctions définies sur Y, bornées et lipschitziennes sur chaque élément de α_0, comme dans le paragraphe 1.1.5.

Lemme 2.1.2. *Les opérateurs R_n agissent continûment sur $\mathcal{L}$, avec $\|R_n\| = O(m[\varphi_Y = n])$.*

Démonstration. On a $R_n = \sum_{\underline{a}_Y = [a]_Y, \varphi_Y(a) = n} M_{\underline{a}_Y}$. Ainsi, (2.6) montre que $\|R_n\| \leqslant B(1 + \lambda^{-1}) \sum m[\underline{a}_Y] = (1 + \lambda^{-1}) Bm[\varphi_Y = n]$. □

Par (2.3), on en déduit que T_n agit également continûment sur $\mathcal{L}$. Comme $\sum m[\varphi_Y = n] = m(Y) < \infty$, on peut définir un opérateur $R(z) = \sum R_n z^n$, pour $z \in \overline{\mathbb{D}} = \{z \in \mathbb{C} \mid |z| \leqslant 1\}$, qui agit sur $\mathcal{L}$.

Lemme 2.1.3. *L'opérateur $R(1) : \mathcal{L} \to \mathcal{L}$ a une valeur propre simple isolée en 1. Le projecteur spectral correspondant est donné par $Pu(x) = \frac{1}{m[Y]} \int_Y u \, \mathrm{d}m$.*

Démonstration. Comme $R(1)$ tient compte de tous les retours à Y, il n'est pas difficile de vérifier que $R(1)$ est l'opérateur de transfert $\widehat{T}_Y$ associé à T_Y. Comme T_Y est transitive (car T est transitive), la proposition 1.1.17 montre que 1 est valeur propre simple isolée de $R(1)$, la fonction propre correspondante étant 1 puisque m est invariante. Enfin, l'expression du projecteur spectral découle du corollaire 1.1.20. □

Lemme 2.1.4. *$\forall z \in \overline{\mathbb{D}}\backslash\{1\}$, $I - R(z)$ est inversible sur $\mathcal{L}$.*

Démonstration. En sommant les estimées (2.6) sur tous les cylindres $\underline{a}_Y = [a_0, \dots, a_{n-1}]_Y$ de longueur n, on obtient

$$\|R(z)^n u\|_{\mathcal{L}} \leqslant \sum_{\underline{a}_Y} |z|^{l(\underline{a}_Y)} Bm(\underline{a}_Y) \left(\lambda^{-n} \|u\|_{\mathcal{L}} + \frac{1}{m(\underline{a}_Y)} \int_{\underline{a}_Y} |u| \, \mathrm{d}m \right),$$

où $l(\underline{a}_Y)$ désigne la longueur de $\underline{a}_Y$ vu comme cylindre pour T, et vérifie donc $l(\underline{a}_Y) \geqslant n$. Ainsi,

$$\|R(z)^n u\|_{\mathcal{L}} \leqslant B|z|^n (\lambda^{-n} \|u\|_{\mathcal{L}} + \|u\|_1). \tag{2.7}$$

Comme l'injection $\mathcal{L} \to L^1(m)$ est compacte (proposition 1.1.16), le théorème 1.1.11 implique que, $\forall z \in \overline{\mathbb{D}}$, le rayon spectral essentiel de $R(z)$ sur $\mathcal{L}$ est $\leqslant \lambda^{-1}|z| < 1$. Ainsi, pour obtenir l'inversibilité de $I - R(z)$, il suffit de vérifier que 1 n'est pas une valeur propre de $R(z)$. Il n'y a un problème possible que pour $|z| = 1$ puisque, autrement, le rayon spectral de $R(z)$ est majoré par $|z| < 1$, par (2.7).

Soit donc z de module 1, et supposons que $R(z)f = f$ pour un élément f non nul de $\mathcal{L}$. On notera, pour $u, v \in L^2(Y)$, $\langle u, v\rangle = \int \overline{u} v \, \mathrm{d}m$. Soit $W : L^\infty(Y) \to L^\infty(Y)$ défini par $Wu = \overline{z}^{\varphi_Y} u \circ T_Y$. Comme $R(z)v = R(1)(z^{\varphi_Y} v)$, cet opérateur W est l'adjoint de $R(z)$. Ainsi,

$$\begin{aligned} \|Wf - f\|_2^2 &= \|Wf\|_2^2 - 2\operatorname{Re}\langle Wf, f\rangle + \|f\|_2^2 = \|Wf\|_2^2 - 2\operatorname{Re}\langle f, R(z)f\rangle + \|f\|_2^2 \\ &= \|Wf\|_2^2 - 2\operatorname{Re}\langle f, f\rangle + \|f\|_2^2 = \|Wf\|_2^2 - \|f\|_2^2 . \end{aligned}$$

Comme T_Y préserve la mesure, on a $\|Wf\|_2^2 = \int |f|^2 \circ T_Y = \int |f|^2 = \|f\|_2^2$, si bien que $\|Wf - f\|_2^2 = 0$. En particulier,

$$z^{\varphi_Y(x)} f(x) = f(T_Y x) \tag{2.8}$$

presque partout.

On définit une fonction $\widetilde{f}$ sur X ainsi : soient $y \in Y$ et $i < \varphi_Y(y)$, posons $\widetilde{f}(T^i y) = z^i f(y)$. Alors, $\widetilde{f}(Tx) = z\widetilde{f}(x)$ sur X (par construction lorsque $Tx \notin Y$, et par (2.8) lorsque $Tx \in Y$). La fonction $\widetilde{f}$ est non nulle et vérifie $\widetilde{f} \circ T = z\widetilde{f}$. Le mélange de T (proposition 1.3.15) implique donc que $z = 1$. □

Comme $\sum nm[\varphi_Y = n] = m(X) = 1$ par la formule de Kac (proposition 1.3.4), on a $\sum n \|R_n\| < \infty$. Ainsi, on peut définir $R'(z) = \sum nR_n z^{n-1}$ pour $z \in \overline{\mathbb{D}}$.

Lemme 2.1.5. *On a* $PR'(1)P = \frac{1}{m[Y]}P$.

Démonstration. Comme $Pu = \frac{\int u}{m(Y)}$, un changement de variables donne que, pour tout cylindre $\underline{a}_Y$, $PM_{\underline{a}_Y}P = \frac{m(\underline{a}_Y)}{m(Y)}P$. En sommant, on obtient $PR_nP = \frac{m[\varphi_Y=n]}{m[Y]}P$. Le lemme en découle par la formule de Kac. □

2.2 Algèbres de Banach et résultats spectraux

Définition 2.2.1. *Soient* $\mathcal{L}$ *un espace de Banach, et* $(R_n)_{n\geqslant 1}$ *une suite d'opérateurs sur* $\mathcal{L}$. *On dit que c'est une* suite d'opérateurs de renouvellement *si*
- $\sum n\,\|R_n\| < \infty$. *Ainsi,* $R(z) = \sum R_n z^n$ *et* $R'(z) = \sum nR_n z^{n-1}$ *sont définis pour* $z \in \overline{\mathbb{D}}$.
- 1 *est une valeur propre simple isolée de* $R(1)$, *et le projecteur spectral associé* P *vérifie* $PR'(1)P = \mu P$ *avec* $\mu > 0$.
- *Pour tout* $z \in \overline{\mathbb{D}}\backslash\{1\}$, $I - R(z)$ *est inversible.*

Par exemple, lorsque les R_n sont les opérateurs définis dans le paragraphe 2.1.1, les lemmes 2.1.2, 2.1.3, 2.1.4 et 2.1.5 montrent qu'ils constituent une suite d'opérateurs de renouvellement.

Lorsque R_n est une suite d'opérateurs de renouvellement, la série $I - R(z)$ est analytique et inversible sur $\mathbb{D}$. Par [DS58, lemme VII.6.13], son inverse est encore analytique sur $\mathbb{D}$, et il peut donc s'écrire $\sum T_n z^n$. En fait, T_n peut s'obtenir directement par $T_n = \sum_{k_1+\ldots+k_l=n} R_{k_1}\cdots R_{k_l}$. Le but de cette partie va être de comprendre précisément le comportement de T_n en fonction de celui de R_n. L'outil technique principal de l'étude sera la description précise de $S(z) := \frac{I-R(z)}{1-z}$.

Dans le cas des opérateurs R_n associés à une tour localement Gibbs-Markov, ces T_n coïncident avec les T_n définis dans le paragraphe 2.1.1. Ainsi, les résultats de cette partie permettront d'obtenir des informations très précises sur la vitesse de décorrélation.

2.2.1 Algèbres de Banach et lemme de Wiener

Pour étudier précisément les suites d'opérateurs de renouvellement, on aura besoin d'utiliser des techniques d'algèbres de Banach.

Rappelons qu'une algèbre de Banach $\mathcal{A}$ est un espace de Banach complexe muni d'un produit associatif $\mathcal{A}\times\mathcal{A}\to\mathcal{A}$ tel que $\|AB\| \leqslant \|A\|\cdot\|B\|$ et admettant un élément neutre. L'ensemble des éléments inversibles de $\mathcal{A}$ est alors un ouvert de $\mathcal{A}$, sur lequel l'inversion est continue et même localement lipschitzienne.

Soit $\mathcal{A}$ une algèbre de Banach. On notera $A(\mathcal{A})$ l'ensemble des fonctions continues $f : S^1 \to \mathcal{A}$ dont les coefficients de Fourier, définis par

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} f(e^{i\theta})\, \mathrm{d}\theta,$$

sont sommables en norme.

Proposition 2.2.2. *Si $f : S^1 \to \mathcal{A}$ est continue et satisfait $\sum \|c_n(f)\| < +\infty$, alors $f(e^{i\theta}) = \sum c_n(f) e^{in\theta}$, la série convergeant normalement.*

Démonstration. Quitte à remplacer f par $f - \sum c_n(f) e^{in\theta}$, on peut supposer que $c_n(f) = 0$ pour tout n, et on veut montrer que $f = 0$.

Supposons qu'il existe $z \in S^1$ tel que $f(z) \neq 0$. Il existe alors une forme linéaire continue ξ sur $\mathcal{A}$, avec $\xi(f(z)) \neq 0$. Comme ξ est linéaire, $c_n(\xi \circ f) = \xi(c_n(f)) = 0$ pour tout n. Comme $\xi \circ f$ est continue et à valeurs complexes, un résultat classique (prouvé par exemple en utilisant l'égalité de Parseval) montre que $\xi \circ f = 0$, ce qui est absurde. □

On vérifie alors aisément que, si les coefficients de Fourier de f et g sont sommables en norme, il en va de même de ceux de fg. Plus précisément, si on munit $A(\mathcal{A})$ de la norme $\|f\|_{A(\mathcal{A})} = \sum_{n\in\mathbb{Z}} \|c_n(f)\|$, alors $A(\mathcal{A})$ est une algèbre de Banach.

Pour $f : S^1 \to \mathcal{A}$, on notera aussi $\|f\|_{\mathcal{A}} = \sup\{\|f(z)\| \mid z \in S^1\}$. On a $\|f\|_{\mathcal{A}} \leqslant \|f\|_{A(\mathcal{A})}$, mais l'inégalité inverse est fausse.

Dans le cas où $A = \mathbb{C}$, le lemme de Wiener affirme que, si $f : S^1 \to \mathbb{C}$ appartient à $A(\mathbb{C})$ et ne s'annule pas sur S^1, alors $1/f$ appartient aussi à $A(\mathbb{C})$. On a un résultat analogue pour une algèbre de Banach $\mathcal{A}$ générale :

Théorème 2.2.3. *Soit $f \in A(\mathcal{A})$ telle que, pour tout $z \in S^1$, $f(z)$ est inversible dans l'algèbre de Banach $\mathcal{A}$. Alors, en posant $g(z) = f(z)^{-1}$, on a aussi $g \in A(\mathcal{A})$.*

Démonstration. Lorsque $\mathcal{B}$ est une algèbre de Banach commutative, les caractères de $\mathcal{B}$ sont les morphismes d'algèbre de $\mathcal{B}$ dans $\mathbb{C}$. Le théorème de Gelfand ([Rud91, théorème 11.5 (c)]) affirme que, si un élément b de $\mathcal{B}$ vérifie $\chi(b) \neq 0$ pour tout caractère χ, alors b est inversible dans $\mathcal{B}$.

Pour $\lambda \in S^1$, on définit un caractère χ_λ sur $A(\mathbb{C})$ par $\chi_\lambda(f) = f(\lambda)$. Ce sont en fait les seuls caractères de $A(\mathbb{C})$ ([Rud91, théorème 11.6]). Ainsi, le théorème de Gelfand permet de conclure dans le cas $\mathcal{A} = \mathbb{C}$.

On en déduit le résultat pour $\mathcal{A}$ générale grâce à [BP42]. □

Dans la suite, on utilisera d'autres algèbres de Banach que $A(\mathcal{A})$, prenant en compte la vitesse de décroissance des coefficients de Fourier. Pour obtenir un analogue du lemme de Wiener

dans ces algèbres, il suffira donc d'identifier les caractères de l'algèbre obtenue en remplaçant $\mathcal{A}$ par $\mathbb{C}$, par le même argument.

On peut déduire du théorème précédent un résultat plus général. Soient Ω un ouvert de $\mathbb{C}$ et $h : \Omega \to \mathbb{C}$ une fonction holomorphe. Si $x \in \mathcal{A}$ est tel que son spectre $\sigma(x)$ est inclus dans Ω, on peut définir $h(x) \in \mathcal{A}$ en utilisant la formule de Cauchy, par $h(x) := \frac{1}{2i\pi} \int_\gamma \frac{h(u)}{uI-x}\,\mathrm{d}u$, où γ est un chemin entourant $\sigma(x)$ dans Ω. Cette expression est indépendante de γ (cf. [DS58, VII.3]). Par exemple, si $h(z) = z^n$ avec $n \in \mathbb{Z}$, alors $h(x) = x^n$ (pour la multiplication dans l'algèbre $\mathcal{A}$). Pour un autre exemple, soient $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ où $\mathcal{L}$ est un espace de Banach, et $x \in \mathcal{A}$ pour lequel 1 est isolé dans $\sigma(x)$. Si h est égale à 1 dans un petit voisinage de 1 et à 0 hors de ce voisinage, alors $h(x)$ est le projecteur spectral associé à la valeur propre 1 de x.

Théorème 2.2.4. *Soient Ω un ouvert de $\mathbb{C}$ et $h : \Omega \to \mathbb{C}$ holomorphe. Soit $f \in A(\mathcal{A})$ telle que, pour tout $z \in S^1$, $\sigma(f(z)) \subset \Omega$. En notant $g(z) = h(f(z))$, on a encore $g \in A(\mathcal{A})$.*

Le lemme de Wiener précédent est le cas particulier correspondant à $h(z) = 1/z$.

Démonstration. Comme le spectre dépend semi-continûment de l'opérateur, il existe un chemin γ dans Ω qui entoure $\sigma(f(z))$ pour tout $z \in S^1$. Ainsi, $g(z) = \frac{1}{2i\pi} \int_\gamma \frac{h(u)}{uI-f(z)}\,\mathrm{d}u$. Pour $u \in \gamma$ fixé, $M_u : z \mapsto \frac{h(u)}{uI-f(z)}$ appartient à $A(\mathcal{A})$ par le théorème 2.2.3. Comme l'inversion est continue dans l'algèbre de Banach $A(\mathcal{A})$, $u \mapsto M_u$ est continue sur le compact γ. Comme $g = \int_\gamma M_u\,\mathrm{d}u$, cela permet de conclure. □

Ce résultat s'applique en particulier aux projecteurs spectraux.

2.2.2 Sommabilité des T_n

Dans ce paragraphe, on va démontrer le résultat suivant :

Théorème 2.2.5. *Soit R_n une suite d'opérateurs de renouvellement, et soit $S(z) = \frac{I-R(z)}{1-z}$ pour $z \in \overline{\mathbb{D}}\backslash\{1\}$. Alors $S(z)^{-1}$ se prolonge continûment en 1 par $\frac{1}{\mu}P$, et ses coefficients de Fourier sont sommables.*

En particulier, si T_n est donné par $\sum T_n z^n = (I - R(z))^{-1}$, alors $\sum \|T_n - T_{n-1}\| < \infty$.

Démonstration. On notera $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$. Le problème principal de la preuve est le suivant : $R(z)$ a une valeur propre $\lambda(z)$ proche de 1 pour z proche de 1, par la proposition 1.1.10, mais il n'y a pas de raison que cette valeur propre puisse se définir pour tout $z \in S^1$, si bien qu'on ne peut pas lui appliquer directement de technique d'analyse de Fourier (et en particulier le lemme de Wiener). L'idée va être de perturber $R(z)$ en $\widetilde{R}(z)$ de manière à obtenir une valeur propre définie sur tout S^1.

Étape 1 : $S(z)^{-1}$ s'étend continûment à $\overline{\mathbb{D}}$, et se prolonge par $\frac{1}{\mu}P$ en 1.

Le problème est l'extension à 1 puisque, ailleurs, $I - R(z)$ est inversible par hypothèse. Pour $z \in \overline{\mathbb{D}}$ proche de 1, soit $\lambda(z)$ la valeur propre proche de 1 de $R(z)$. On note $P(z)$ le projecteur spectral correspondant, et $Q(z) = I - P(z)$. Alors

$$S(z)^{-1} = \frac{1-z}{1-\lambda(z)} P(z) + (1-z)(I - R(z)Q(z))^{-1} Q(z). \tag{2.9}$$

Comme $I - R(z)Q(z)$ est inversible, le second terme s'étend continûment en 1, par 0. Comme $(I - R(1))P(1) = 0$,

$$\frac{1-\lambda(z)}{1-z} P(z) = P(z) \frac{R(1) - R(z)}{1-z} P(z) + P(z)(I - R(1)) \frac{P(z) - P(1)}{1-z}. \tag{2.10}$$

Notons que

$$\frac{R(z) - R(1)}{z-1} = \sum_{n=0}^{\infty} \left(\sum_{k=n+1}^{\infty} R_k \right) z^n.$$

Comme $\sum_{k>n} \|R_k\|$ est sommable, cette somme converge en norme. En particulier, elle converge en 1 vers $R'(1)$.

Si $\delta > 0$ est assez petit, alors, pour z assez proche de 1,

$$P(z) = \frac{1}{2i\pi} \int_{|u-1|=\delta} \frac{1}{uI - R(z)} \, \mathrm{d}u.$$

Ainsi,

$$\frac{P(z) - P(1)}{z-1} = \frac{1}{2i\pi} \int_{|u-1|=\delta} \frac{1}{uI - R(z)} \frac{R(z) - R(1)}{z-1} \frac{1}{uI - R(1)} \, \mathrm{d}u. \tag{2.11}$$

Comme $\frac{R(z)-R(1)}{z-1}$ a une limite en 1, on obtient que $\frac{P(z)-P(1)}{z-1}$ converge également en 1, vers une limite notée $P'(1)$.

Dans (2.10), quand z tend vers 1, le terme de droite tend vers $P(1)R'(1)P(1) - P(1)(I - R(1))P'(1) = P(1)R'(1)P(1)$ puisque $P(1)(I - R(1)) = 0$. Par hypothèse, $P(1)R'(1)P(1) = \mu P$ pour un certain $\mu > 0$.

Soit ξ une forme linéaire continue sur $\mathrm{Hom}(\mathcal{L}, \mathcal{L})$ telle que $\xi(P(1)) = 1$. On applique ξ aux deux membres de (2.10) et on fait tendre z vers 1. On obtient

$$\frac{1-\lambda(z)}{1-z} \to \mu. \tag{2.12}$$

Comme $\mu \neq 0$, cela montre que, dans (2.9), le premier terme de droite converge vers $\frac{1}{\mu} P$.

Finalement, dans (2.9), tous les termes à droite ont une limite en 1, ce qui conclut.

Étape 2 : Construction d'une fonction $\widetilde{R}(z)$ sur S^1, qui coïncide avec $R(z)$ pour z proche de 1, qui a partout une valeur propre simple $\widetilde{\lambda}(z)$ proche de 1, différente de 1 pour $z \neq 1$, et avec $\frac{\widetilde{R}(z)-\widetilde{R}(1)}{z-1} \in A(\mathcal{A})$.

On construit $\widetilde{R}$ en trois étapes : on définit d'abord deux approximations F et G, et on les recolle pour obtenir $\widetilde{R}$.

Soit $\gamma > 0$ très petit. Soit $\varphi + \psi$ une partition de l'unité C^∞ sur S^1, subordonnée aux ensembles $\{\theta \in [0, \gamma)\}$ et $\{\theta \in (\gamma - \eta, \pi/2]\}$, où θ désigne l'angle sur S^1 et η est positif et beaucoup plus petit que γ. On définit $F(z) = \varphi(z)R(z) + \psi(z)R(e^{i\gamma})$ sur $\{\theta \in [0, \pi/2]\}$: F est égal à R sur $\{\theta \in [0, \gamma - \eta]\}$ et à $R(e^{i\gamma})$ sur $\{\theta \in [\gamma, \pi/2]\}$. En particulier, le spectre de $F(z)$ est "presque le même" que celui de $R(1)$, si γ est assez petit.

On définit de la même façon F sur $\{\theta \in [-\pi/2, 0]\}$, égal à $R(e^{-i\gamma})$ sur $\{\theta \in [-\pi/2, -\gamma]\}$ et à R sur $\{\theta \in [-\gamma + \eta, 0]\}$.

Finalement, on étend F au demi-cercle restant en symétrisant, i.e. $F(e^{i(\pi/2+a)}) = F(e^{i(\pi/2-a)})$, pour s'assurer que tout se raccorde bien.

Si γ est assez petit, $F(z)$ a une unique valeur propre proche de 1 pour tout z, qui dépend continûment de z ; on la note $\rho(z)$. Le problème serait résolu si on avait $\rho(z) \neq 1$ pour $z \neq 1$, ce qui n'est pas le cas puisque $\rho(-1) = \rho(1) = 1$. Il faut donc perturber un peu ρ. Soit ν une fonction C^∞ à valeurs complexes sur $\{\theta \in [\pi/2, 3\pi/2]\}$, arbitrairement proche de ρ et qui ne prend pas la valeur 1 (par transversalité). Sur $\{\theta \in [\pi/2 + \eta, 3\pi/2 - \eta]\}$, on définit $G(z) = \frac{\nu(z)}{\rho(z)}F(z)$: sa valeur propre proche de 1 est $\nu(z) \neq 1$. Finalement, pour obtenir $\widetilde{R}$, on raccorde F et G sur $\{\theta \in [\pi/2, \pi/2 + \eta]\}$ et $\{\theta \in [3\pi/2 - \eta, 3\pi/2]\}$ en utilisant une partition de l'unité comme plus haut. Comme le spectre de $F(e^{i\pi/2}) = R(e^{i\gamma})$ ne contient pas 1, cela ne donnera pas de valeur propre 1 si η est assez petit et ν est assez proche de ρ.

Finalement, comme $\frac{R(z)-R(1)}{z-1} \in A(\mathcal{A})$, on a aussi $\frac{\widetilde{R}(z)-\widetilde{R}(1)}{z-1} \in A(\mathcal{A})$: pour obtenir $\widetilde{R}(z)$, on a juste multiplié par des fonctions C^∞, qui sont dans $A(\mathbb{C})$.

Étape 3 : Soit $\widetilde{S}(z) = \frac{I-\widetilde{R}(z)}{1-z}$ *sur* S^1. *Alors* $\widetilde{S}(z)^{-1} \in A(\mathcal{A})$.

Par construction, $\frac{\widetilde{R}(z)-\widetilde{R}(1)}{z-1} \in A(\mathcal{A})$, et $\widetilde{R}(z) \in A(\mathcal{A})$. L'équation (2.11) (avec des tildes partout) implique alors que $\frac{\widetilde{P}(z)-\widetilde{P}(1)}{z-1} \in A(\mathcal{A})$ (en utilisant d'abord le lemme de Wiener 2.2.3 sur $uI - \widetilde{R}(z)$, puis en multipliant et en intégrant, comme dans la preuve du théorème 2.2.4). En particulier, $\widetilde{P}(z) \in A(\mathcal{A})$.

Soit ξ une forme linéaire continue sur $\mathcal{A}$, avec $\xi(P(1)) = 1$ et, $\forall z \in S^1$, $\xi(\widetilde{P}(z)) \neq 0$ (on peut d'abord fixer ξ avec $\xi(P(1)) = 1$, puis choisir γ assez petit dans la construction de $\widetilde{R}(z)$ pour assurer que $\xi(\widetilde{P}(z)) \neq 0$ pour tout z). Alors $\xi(\widetilde{P}(z)) \in A(\mathbb{C})$, donc $1/\xi(\widetilde{P}(z)) \in A(\mathbb{C})$ par le lemme de Wiener 2.2.3. On applique ξ aux deux membres de (2.10) (avec des tildes), et on divise par $\xi(\widetilde{P}(z))$, ce qui donne $\frac{1-\widetilde{\lambda}(z)}{1-z} \in A(\mathbb{C})$. Comme cette fonction ne s'annule pas, le lemme de Wiener donne encore que $\frac{1-z}{1-\widetilde{\lambda}(z)} \in A(\mathbb{C})$.

Finalement, on utilise (2.9) (avec des tildes). Par le lemme de Wiener, $(I - \widetilde{R}(z)\widetilde{Q}(z))^{-1} \in A(\mathcal{A})$, puisque $I - \widetilde{R}(z)\widetilde{Q}(z)$ est dans $A(\mathcal{A})$, et partout inversible. Ainsi, le membre de droite de cette équation est dans $A(\mathcal{A})$, ce qui conclut.

Étape 4 : $S(z)^{-1} \in A(\mathcal{A})$.

Soit $\varphi + \psi$ une partition de l'unité C^∞ de S^1 telle que $\widetilde{R} = R$ sur le support de φ. Alors

$$\left(\frac{I - R(z)}{1-z}\right)^{-1} = \varphi(z)\left(\frac{I - \widetilde{R}(z)}{1-z}\right)^{-1} + \psi(z)\left(\frac{I - R(z)}{1-z}\right)^{-1}.$$

Le premier terme du membre de droite est dans $A(\mathcal{A})$, par l'étape 3. Pour le deuxième terme, on peut modifier $\frac{I-R(z)}{1-z}$ hors du support de ψ de manière à le rendre partout inversible (en utilisant les mêmes techniques que dans la construction de $\widetilde{R}$). D'après le lemme de Wiener, son inverse est alors dans $A(\mathcal{A})$, ce qui conclut la preuve.

Étape 5 : conclusion.

Par définition de T_n, on a $(I - R(z))^{-1} = \sum T_n z^n$ sur $\mathbb{D}$. En multipliant par $1 - z$, on obtient $S(z)^{-1} = \sum (T_n - T_{n-1})z^n$. Ainsi, pour $r < 1$, $T_n - T_{n-1} = \frac{1}{2\pi r^n} \int e^{-in\theta} S(re^{i\theta})^{-1} \, \mathrm{d}\theta$. Comme $S(z)^{-1}$ est continue sur $\overline{\mathbb{D}}$ (première étape), on peut faire tendre r vers 1 et obtenir que $T_n - T_{n-1}$ est le n-ième coefficient de Fourier de $S(z)^{-1}$ sur S^1. D'après l'étape 4, ces coefficients de Fourier sont sommables. □

Remarque 2.2.6. La preuve du théorème 2.2.5 montre essentiellement que $S(z)^{-1} \in A(\mathcal{A})$ dès lors que $\frac{R(z)-R(1)}{z-1} \in A(\mathcal{A})$. La preuve marche en fait pour des algèbres de Banach plus générales que $A(\mathcal{A})$, et on utilisera plus loin le même résultat pour des algèbres dont la définition contiendra en plus des informations sur la vitesse de décroissance des suites considérées.

Remarque 2.2.7. L'énoncé du théorème 2.2.5 ne fait intervenir que des fonctions définies sur tout $\overline{\mathbb{D}}$, et pas seulement sur S^1. Cependant, dans la preuve, il a été important de sortir de ce cadre analytique et de travailler sur S^1 pour pouvoir utiliser des partitions de l'unité (et obtenir $\widetilde{R}(z)$ ayant une valeur propre définie partout pour appliquer des techniques d'analyse de Fourier à cette valeur propre).

2.2.3 Construction d'une approximation analytique

Dans ce paragraphe, on démontre le lemme suivant, dû à Sarig ([Sar02]) :

Lemme 2.2.8. *Soient R_n une suite d'opérateurs de renouvellement, et $R(z) = \sum R_n z^n$. Il existe alors un polynôme $R_B : \mathbb{C} \to \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ tel que :*
- *$R_B(1) = R(1)$ et $R'_B(1) = R'(1)$.*
- *$\frac{R_B(1)-R_B(z)}{1-z}$ et $\frac{1}{1-z}\left(\frac{R_B(1)-R_B(z)}{1-z} - R'(1)\right)$ sont des polynômes.*
- *$\frac{I-R_B(z)}{1-z}$ est inversible pour tout $z \in \overline{\mathbb{D}}\backslash\{1\}$, et son inverse se prolonge continûment en 1 par $\frac{1}{\mu}P$.*

– $\forall z \in \mathbb{D}$, $\left(\frac{I-R_B(z)}{1-z}\right)^{-1} = \frac{1}{\mu}P + (1-z)\sum_{n\geqslant 0} z^n E_n$, *où* $E_n \in \mathrm{Hom}(\mathcal{L},\mathcal{L})$ *vérifie* $\|E_n\| = O(\kappa^n)$ *pour un certain* $\kappa < 1$.

Démonstration. Le théorème 2.2.5 montre que $S(z)^{-1}$ est continu sur $\overline{\mathbb{D}}$. Soit M qui borne sa norme. Soit N assez grand pour que $2\sum_N^\infty n\,\|R_n\| < \frac{1}{M}$.

On pose $R_B(z) = \sum_0^{N-1} R_n z^n + \sum_N^\infty R_n + (z-1)\sum_N^\infty nR_n$. Ainsi, $R_B(1) = R(1)$ et $R'_B(1) = R'(1)$. De plus, $\frac{R_B(1)-R_B(z)}{1-z}$ et $\frac{1}{1-z}\left(\frac{R_B(1)-R_B(z)}{1-z} - R'(1)\right)$ sont des polynômes.

Pour $z \in \overline{\mathbb{D}}\backslash\{1\}$, montrons que $S_B(z) := \frac{I-R_B(z)}{1-z}$ est inversible. On a

$$\frac{I-R_B(z)}{1-z} = \frac{I-R(z)}{1-z}\left(I + \left(\frac{I-R(z)}{1-z}\right)^{-1}\frac{R(z)-R_B(z)}{1-z}\right).$$

Mais

$$\frac{R(z)-R_B(z)}{1-z} = \sum_N^\infty nR_n - \sum_N^\infty R_n(1+z+\ldots+z^{n-1}). \tag{2.13}$$

Ainsi, sa norme est bornée par $2\sum_N^\infty n\,\|R_n\| < \frac{1}{M}$, donc $\left\|\left(\frac{I-R(z)}{1-z}\right)^{-1}\frac{R(z)-R_B(z)}{1-z}\right\| < 1$, si bien que $I + \left(\frac{I-R(z)}{1-z}\right)^{-1}\frac{R(z)-R_B(z)}{1-z}$ est inversible. Finalement,

$$S_B(z)^{-1} = \left(I + S(z)^{-1}\frac{R(z)-R_B(z)}{1-z}\right)^{-1} S(z)^{-1},$$

ce terme étant défini sur tout $\overline{\mathbb{D}}\backslash\{1\}$.

L'opérateur $R_B(1) = R(1)$ a une valeur propre simple en 1 si bien que, pour $z \in \mathbb{C}$ proche de 1, $R_B(z)$ a encore une valeur propre $\lambda_B(z)$ proche de 1, qui dépend holomorphiquement de z. En notant $P_B(z)$ et $Q_B(z)$ les projecteurs spectraux de $R_B(z)$ correspondant respectivement à $\lambda_B(z)$ et au reste du spectre, on a comme en (2.9)

$$S_B(z)^{-1} = \frac{1-z}{1-\lambda_B(z)}P_B(z) + (1-z)(I - R_B(z)Q_B(z))^{-1}Q_B(z).$$

Ainsi, $S_B(z)^{-1}$ admet une extension méromorphe à un voisinage de 1 dans $\mathbb{C}$; cette extension est en fait holomorphe, puisque $S_B(z)^{-1}$ converge en 1 vers $\frac{1}{\mu}P$, par les mêmes arguments que dans la première étape de la preuve du théorème 2.2.5. De plus, comme les inversibles de $\mathrm{Hom}(\mathcal{L},\mathcal{L})$ sont ouverts et comme $S_B(z)$ est inversible pour $z \in \overline{\mathbb{D}}\backslash\{1\}$, on obtient finalement un prolongement holomorphe de $S_B(z)^{-1}$ à une couronne $\{|z| < \alpha\}$ pour un certain $\alpha > 1$.

Enfin, $\frac{1}{1-z}\left(S_B(z)^{-1} - \frac{1}{\mu}P\right)$ est encore holomorphe dans cette couronne. Il s'écrit donc sous la forme $\sum E_n z^n$, avec $\|E_n\| = O(\kappa^n)$ pour tout $\kappa > 1/\alpha$. □

Le théorème 2.2.5 montre que T_n est une suite de Cauchy, donc convergente. Le lemme 2.2.8 va permettre d'identifier sa limite :

Proposition 2.2.9. *Soit R_n une suite d'opérateurs de renouvellement. Alors T_n converge vers $\frac{1}{\mu}P$.*

Démonstration. On écrit $S(z)^{-1} = S_B(z)^{-1} + S_B(z)^{-1}(S_B(z) - S(z))S(z)^{-1}$, et $S_B(z)^{-1} = \frac{1}{\mu}P + (1-z)E(z)$ où les coefficients de $E(z)$ tendent vers 0 (et décroissent même exponentiellement vite). Comme $\sum T_n z^n = \frac{1}{1-z}S(z)^{-1}$, on obtient

$$\sum T_n z^n = \frac{1}{1-z}\frac{1}{\mu}P + E(z) + S_B(z)^{-1}\frac{S_B(z)-S(z)}{1-z}S(z)^{-1}.$$

Si on montre que le coefficient de z^n dans $S_B(z)^{-1}\frac{S_B(z)-S(z)}{1-z}S(z)^{-1}$ tend vers 0, cela démontrera donc le résultat annoncé.

On a $S_B(z) - S(z) = R'(1) - \frac{R(z)-R(1)}{z-1} + (1-z)H(z)$ où H est un polynôme, par le lemme 2.2.8. De plus, $R'(1) - \frac{R(z)-R(1)}{z-1} = \sum_n \left(\sum_{n+1}^{\infty} R_p\right)(1-z^n)$, donc

$$\frac{1}{1-z}\left(R'(1) - \frac{R(z)-R(1)}{z-1}\right) = \sum_n \left(\sum_{p=n+1}^{\infty}\sum_{q=p+1}^{\infty} R_q\right) z^n.$$

Comme $\sum n\,\|R_n\| < \infty$, le coefficient de z^n dans cette série tend vers 0 quand $n \to \infty$. Ainsi, le coefficient de z^n dans $\frac{S_B(z)-S(z)}{1-z}$ tend vers 0.

On sait aussi que les coefficients de $S_B(z)^{-1}$ et de $S(z)^{-1}$ sont sommables par le théorème 2.2.5 et le lemme 2.2.8. Pour conclure, il suffit donc de montrer que, si $A_n \to 0$ et B_n est sommable, alors le coefficient $(AB)_n$ de z^n dans $\left(\sum A_n z^n\right)\left(\sum B_n z^n\right)$ tend vers 0 quand $n \to \infty$.

On a $(AB)_n = \sum_{k=0}^n A_k B_{n-k}$. Soient $\varepsilon > 0$, et N tel que, si $k \geqslant N$, $\|A_k\| \leqslant \varepsilon$. Si K majore $\|A_k\|$ et $\sum \|B_k\|$, on obtient pour $n \geqslant N$

$$\|(AB)_n\| \leqslant \sum_{k=0}^{N-1} \|A_k\|\,\|B_{n-k}\| + \sum_{k=N}^{n} \varepsilon\,\|B_{n-k}\| \leqslant K\sum_{l=n-N}^{\infty}\|B_l\| + K\varepsilon.$$

Pour n assez grand, c'est majoré par $2K\varepsilon$. □

2.2.4 Classes de vitesse de décroissance

Dans ce paragraphe, on introduit des ensembles de suites, qui seront les vitesses de décroissance de $\|R_n\|$ pour lesquelles on pourra décrire précisément le comportement de T_n.

Pour u_n, v_n deux suites réelles, on notera $u \star v$ leur produit de convolution, défini par $(u\star v)_n = \sum_{k=0}^n u_k v_{n-k}$.

La prochaine définition regroupe essentiellement les propriétés qui seront utiles pour faire marcher les démonstrations qui suivront. Cependant, ces propriétés ne sont pas si artificielles qu'il y paraît, puisqu'elles sont satisfaites par beaucoup de suites, comme on le verra plus loin.

Définition 2.2.10. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}_+^*$ sommable. On dit qu'elle est* convolutive *si*

1. *Il existe $C > 0$ tel que, pour tous $n \leqslant m$, $u_m \leqslant Cu_n$ (i.e. "u_n est presque décroissante").*
2. *Pour tout $\varepsilon > 0$, $(1-\varepsilon)^n = O(u_n)$.*
3. *Pour tous I_1, I_2 intervalles ouverts de S^1 avec $I_1 \cup I_2 = S^1$, il existe des fonctions $f_1, f_2 : S^1 \to [0,1]$ avec $f_1 + f_2 = 1$ et $C > 0$ tels que, pour $i = 1, 2$, f_i soit nulle hors de I_i, et $\forall n \in \mathbb{Z}$ $|c_n(f_i)| \leqslant Cu_{|n|}$.*
4. *Il existe $C > 0$ tel que, pour tout $n \in \mathbb{N}$, $(u \star u)_n \leqslant Cu_n$.*
5. *Pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que la suite $v_n = 1_{n\geqslant N}u_n$ vérifie : $\forall n \in \mathbb{N}, (v\star v)_n \leqslant \varepsilon u_n$.*
6. *Soit $w_n = \sum_{n+1}^{\infty} u_p$. Alors $w \star w + w \star (nu_n) = o(nw_n)$, et $u_n = o(w_n)$.*

On dit que u_n est fortement convolutive *si on a en plus $w\star w + w\star(nu_n) = O(\max(nu_n, w_n))$.*

La troisième condition, d'existence de partitions de l'unité, est essentiellement une condition de décroissance non exponentielle de u_n, analogue à la deuxième condition : si u_n décroissait exponentiellement vite, alors une fonction non nulle f vérifiant $|c_n(f)| \leqslant Cu_{|n|}$ serait analytique, et ne pourrait donc être nulle sur un intervalle non trivial. Par contre, de nombreuses vitesses "presque exponentielles" laissent assez de liberté pour avoir des partitions de l'unité (voir par exemple la preuve de la proposition 2.2.13).

Définition 2.2.11. *Soit $(u_n)_{n\in\mathbb{N}}$ une suite de $\mathbb{R}_+^*$. On dit qu'elle est à décroissance polynomiale s'il existe $C > 0$ tel que, pour tous $n/2 \leqslant k$, $u_k \leqslant Cu_n$.*

Cela implique que u_n ne décroît pas trop vite : il existe $\gamma > 0$ tel que $\frac{1}{n^\gamma} = O(u_n)$ (par exemple $\gamma = \frac{\log C}{\log 2}$). Les suites $u_n = \frac{1}{(n+1)^\gamma}$ avec $\gamma > 0$ sont à décroissance polynomiale. Plus généralement, si L est une fonction à variation lente en $+\infty$ et $\gamma > 0$, alors $\frac{L(n)}{(n+1)^\gamma}$ est à décroissance polynomiale.

Proposition 2.2.12. *Une suite sommable u_n à décroissance polynomiale est convolutive. De plus, si $w_n = \sum_{n+1}^{\infty} u_p$ est sommable, alors u_n est fortement convolutive.*

Démonstration. Soit u_n sommable à décroissance polynomiale. Par définition, elle est presque décroissante. De plus, des partitions de l'unité C^∞ vérifient $|c_n(f)| \leqslant \frac{C_p}{|n|^p}$ pour tout p, donc en particulier $|c_n(f)| \leqslant Cu_{|n|}$.

Il reste à vérifier les conditions sur la convolution de u_n. Soit C tel que, $\forall n/2 \leqslant k, u_k \leqslant Cu_n$. Alors, pour $n \in \mathbb{N}$,

$$(u \star u)_n = \sum_0^n u_k u_{n-k} \leqslant 2 \sum_{0\leqslant k\leqslant n/2} u_k u_{n-k} \leqslant 2 \sum_{0\leqslant k\leqslant n/2} u_k C u_n \leqslant 2C \left(\sum_0^\infty u_k\right) u_n.$$

Soit maintenant $\varepsilon > 0$. Soit N tel que $2C\sum_N^\infty u_k \leqslant \varepsilon$. Le même calcul montre que $v_n = 1_{n\geqslant N} u_n$ satisfait $(v \star v)_n \leqslant \varepsilon u_n$ pour tout $n \in \mathbb{N}$.

Soit $w_n = \sum_{n+1}^\infty u_p$. Alors w_n est encore à décroissance polynomiale. Ainsi,

$$(w \star w)_n \leqslant 2 \sum_{0\leqslant k\leqslant n/2} w_k w_{n-k} \leqslant 2C \left(\sum_{0\leqslant k\leqslant n/2} w_k\right) w_n. \tag{2.14}$$

Comme $w_n \to 0$, on a $\sum_{0\leqslant k\leqslant n/2} w_k = o(n)$, donc $w\star w = o(nw_n)$. De plus, $nu_n \leqslant C\sum_{p=n+1}^{2n} u_k$ par décroissance polynomiale, donc $nu_n = O(w_n)$, ce qui implique que $w \star (nu_n) = o(nw_n)$. Cela montre que u est convolutive.

Finalement, si w_n est sommable, la somme $\sum_{0\leqslant k\leqslant n/2} w_k$ est bornée dans (2.14), si bien que $w \star w = O(w_n)$. □

La suite $u_n = e^{-n}$ n'est pas convolutive, puisqu'elle n'autorise pas de partition de l'unité. Cependant, les suites qui tendent un peu moins rapidement vers 0 peuvent être convolutives :

Proposition 2.2.13. *Soient $0 < \eta < 1$ et $c > 0$. Alors la suite $u_n = e^{-cn^\eta}$ est fortement convolutive.*

Démonstration. Montrons tout d'abord l'existence de partitions de l'unité. Soient I_1, I_2 deux intervalles ouverts recouvrant S^1, on va utiliser le résultat suivant ([Man52, théorème 4.1.VI]).

Soit $\varepsilon_1, \varepsilon_2, \dots$ une suite strictement positive satisfaisant $\sum \varepsilon_i < \infty$. Il existe alors une partition de l'unité f_1, f_2 subordonnée aux intervalles I_1 et I_2 et une constante $C > 0$ telles que, $\forall n \in \mathbb{N}$, $\forall i = 1, 2$, $\left\|f_i^{(n)}\right\|_\infty \leqslant \frac{C^n}{\varepsilon_1 \cdots \varepsilon_n}$.

Soit $1 < \gamma < 1/\eta$. On applique ce théorème à $\varepsilon_n = \frac{1}{n^\gamma}$, et on obtient f_1, f_2 et C. Soient $f = f_1$ ou f_2, et $n \in \mathbb{N}^*$, estimons $|c_n(f)|$. Soit $p = \left\lfloor \left(\frac{n}{C}\right)^{1/\gamma} \right\rfloor$. Alors

$$|c_n(f)| \leqslant \frac{\left\|f^{(p)}\right\|_\infty}{n^p} \leqslant \frac{(p!)^\gamma C^p}{n^p} \leqslant D\sqrt{p}^\gamma \frac{(p/e)^{\gamma p} C^p}{n^p} = D\sqrt{p}^\gamma \left(\frac{p^\gamma C}{n}\right)^p e^{-\gamma p}.$$

Par définition de p, $\frac{p^\gamma C}{n} \leqslant 1$, si bien qu'on obtient une majoration

$$|c_n(f)| \leqslant D'\sqrt{n} e^{-\frac{\gamma}{C^{1/\gamma}} n^{1/\gamma}}.$$

Comme $1/\gamma > \eta$, cela montre que $|c_n(f)| = O(e^{-cn^\eta})$. Comme $|c_{-n}(f)| = |c_n(f)|$ puisque f est à valeurs réelles, on obtient la même estimée pour $n < 0$.

Vérifions finalement les inégalités de convolution. Soient $\varepsilon > 0$ et N entier. Notons $v_n = 1_{n\geqslant N} u_n$, on va montrer que $(v\star v)_n \leqslant \varepsilon u_n$ si N est assez grand. La démonstration de $(u\star u)_n \leqslant Cu_n$ est analogue, et même plus facile.

Notons que, sur $[0, 1/2]$, la fonction $(x^\eta + (1-x)^\eta - 1)/x^\eta$ est continue (elle se prolonge par 1 en 0) et strictement positive, donc plus grande qu'une constante $\mu > 0$. Ainsi, $x^\eta + (1-x)^\eta \geqslant 1 + \mu x^\eta$.

Pour $n < 2N$, on a $(v \star v)_n = 0$. Soit donc $n \geqslant 2N$. Alors

$$(v \star v)_n \leqslant 2 \sum_{N\leqslant k\leqslant n/2} e^{-ck^\eta} e^{-c(n-k)^\eta} = 2 \sum_{N\leqslant k\leqslant n/2} e^{-cn^\eta((k/n)^\eta + (1-k/n)^\eta)}.$$

Le réel $x = k/n$ appartient à $[0, 1/2]$, donc

$$(v \star v)_n \leqslant 2 \sum_{N\leqslant k\leqslant n/2} e^{-cn^\eta(1+\mu(k/n)^\eta)} \leqslant 2e^{-cn^\eta} \sum_{k\geqslant N} e^{-c\mu k^\eta}.$$

Il suffit donc de prendre N assez grand pour que $2\sum_{k\geqslant N} e^{-c\mu k^\eta} \leqslant \varepsilon$ pour conclure.

Soit $w_n = \sum_{n+1}^\infty u_p$. En comparant à une intégrale, on vérifie que $w_n \sim Cn^{1-\eta}u_n$. Le même argument que ci-dessus pour estimer les convolutions donne alors $w\star w = O(w_n)$ et $w\star(nu_n) = O(nu_n)$, ce qui montre que u_n est fortement convolutive. □

Lorsque u_n est convolutive, on peut lui associer une algèbre de Banach :

Définition 2.2.14. *Soient u_n une suite convolutive et $\mathcal{A}$ une algèbre de Banach. On note $\mathcal{O}_u(\mathcal{A})$ l'ensemble des fonctions continues $f : S^1 \to \mathcal{A}$ dont les coefficients de Fourier sont sommables et satisfont $c_n(f) = O(u_n)$ lorsque $n \to +\infty$. On le munit de la norme $\|f\|_u = \sum_{n\in\mathbb{Z}} \|c_n(f)\| + \sup_{n\geqslant 0} \frac{\|c_n(f)\|}{u_n}$.*

Proposition 2.2.15. *Il existe une norme équivalente à $\|\ \|_u$ sur $\mathcal{O}_u(\mathcal{A})$, qui fait de $\mathcal{O}_u(\mathcal{A})$ une algèbre de Banach pour le produit des fonctions (ou, de manière équivalente, pour la convolution des coefficients de Fourier).*

Démonstration. Notons $P_u(f) = \sup_{n\geqslant 0} \frac{\|c_n(f)\|}{u_n}$ et $\|f\|_A = \sum \|c_n(f)\|$. Pour $f, g \in \mathcal{O}_u(\mathcal{A})$ et $k \geqslant 0$, notons $f_k = \|c_k(f)\|$ et $g_k = \|c_k(g)\|$. Comme u est presque décroissante, on a si $n \geqslant 0$

$$\begin{aligned}
\frac{\|c_n(fg)\|}{u_n} &\leqslant \frac{\sum_{k\in\mathbb{Z}} f_k g_{n-k}}{u_n} \leqslant \frac{1}{u_n}\sum_0^n f_k g_{n-k} + \sum_{k=-\infty}^{-1} f_k P_u(g)\frac{u_{n-k}}{u_n} + \sum_{k=-\infty}^{-1} g_k P_u(f)\frac{u_{n-k}}{u_n} \\
&\leqslant \frac{(u\star u)_n}{u_n} P_u(f)P_u(g) + \|f\|_A P_u(g)C + \|g\|_A P_u(f)C \\
&\leqslant C\big[P_u(f)P_u(g) + \|f\|_A P_u(g) + \|g\|_A P_u(f)\big].
\end{aligned}$$

On a $\|fg\|_A \leqslant \|f\|_A \|g\|_A$. On peut supposer que $C \geqslant 1$, et on obtient $\|fg\|_u \leqslant C \|f\|_u \|g\|_u$. Soit finalement $\|f\|' = C\|f\|_u$, elle vérifie bien $\|fg\|' \leqslant \|f\|' \|g\|'$, et est équivalente à la norme initiale. □

Ces algèbres vérifient encore un lemme de Wiener :

Théorème 2.2.16. *Soient u_n une suite convolutive et $\mathcal{A}$ une algèbre de Banach. Soit $f : S^1 \to \mathcal{A}$ appartenant à $\mathcal{O}_u(\mathcal{A})$ telle que $\forall z \in S^1, f(z)$ est inversible dans $\mathcal{A}$. Alors la fonction $g : S^1 \to \mathcal{A}$ donnée par $g(z) = f(z)^{-1}$ appartient aussi à $\mathcal{O}_u(\mathcal{A})$.*

Démonstration. On suit le même schéma de preuve que pour le théorème 2.2.3. Il suffit donc de montrer que tous les caractères de l'algèbre $\mathcal{O}_u(\mathbb{C})$ sont de la forme $f \mapsto f(\lambda)$ pour un certain $\lambda \in S^1$, car [BP42] s'appliquera alors. On suit la preuve de [Fre87, théorème 1.2.12]. Rappelons que, si χ est un caractère, il vérifie automatiquement $|\chi(f)| \leqslant \|f\|$ pour toute norme d'algèbre sur $\mathcal{O}_u(\mathbb{C})$, par [Rud91, théorème 10.7 (c)]. En particulier, il existe K tel que $|\chi(f)| \leqslant K\|f\|_u$ pour tout $f \in \mathcal{O}_u(\mathbb{C})$.

Soit χ un caractère sur $\mathcal{O}_u(\mathbb{C})$, et notons $\lambda = \chi(z \mapsto z)$. Comme χ est un morphisme d'algèbre, on obtient $\chi(f) = f(\lambda)$ pour tout $f(z) = \sum_{-M}^{N} a_n z^n$. Si on avait $|\lambda| < 1$, alors $\chi(z \mapsto z^{-n})$ tendrait vers $+\infty$ alors que $z \mapsto z^{-n}$ reste borné dans $\mathcal{O}_u(\mathbb{C})$, ce qui est absurde. Si on avait $|\lambda| > 1$, alors $|\lambda|^n = |\chi(z \mapsto z^n)| \leqslant K\|z \mapsto z^n\|_u \sim K/u_n$, donc u_n tendrait exponentiellement vite vers 0, ce qui contredit la définition des suites convolutives. Ainsi, $\lambda \in S^1$.

On veut montrer que $\chi(f) = f(\lambda)$ pour tout $f \in \mathcal{O}_u(\mathbb{C})$, sachant que cette formule est vraie pour les sommes finies. Soit donc $f(z) = \sum_{-\infty}^{\infty} a_n z^n \in \mathcal{O}_u(\mathbb{C})$. Notons $f_N(z) = \sum_{-\infty}^{N} a_n z^n$. C'est la limite dans $\mathcal{O}_u(\mathbb{C})$ de $\sum_{-M}^{N} a_n z^n$ quand $M \to \infty$ donc, par continuité des caractères, $\chi(f_N) = f_N(\lambda)$. On va montrer que $\chi(f_N) \to \chi(f)$ quand $N \to \infty$, ce qui conclura.

Le point 5 de la définition des suites convolutives assure que $P_u((f - f_N)^2) \leqslant \varepsilon P_u(f)$ si N est assez grand. Ainsi, $\|(f - f_N)^2\|_u \to 0$. Mais

$$|\chi(f - f_N)|^2 = |\chi((f - f_N)^2)| \leqslant K\left\|(f - f_N)^2\right\|_u \to 0,$$

donc $|\chi(f) - \chi(f_N)| \to 0$. □

On pourrait introduire de nombreuses autres classes de suites, et les algèbres de Banach correspondantes (voir [Fre87] pour une multitude de possibilités). Dans l'optique de l'application à la décroissance des corrélations, les suites convolutives semblent cependant les plus utiles. On pourrait certainement affaiblir certaines hypothèses dans la définition des suites convolutives, au prix de complications dans les démonstrations ; cela ne semble pas utile, puisque toutes les applications dans la suite feront intervenir des suites à décroissance polynomiale ou de la forme e^{-cn^η}.

Donnons pour finir un lemme élémentaire qui nous servira plus loin :

Lemme 2.2.17. *Soient u_n une suite convolutive, et $w_n = \sum_{n+1}^{\infty} u_p$. Alors $(w \star u)_n = O(w_n)$.*

Démonstration. On a

$$(w \star u)_n = \sum_{p+q>n, q\leqslant n} u_p u_q \leqslant \sum_{l=n+1}^{\infty} \left(\sum_{p+q=l} u_p u_q \right) = \sum_{l=n+1}^{\infty} (u \star u)_l \leqslant C \sum_{l=n+1}^{\infty} u_l = C w_n. \qquad \square$$

2.2.5 Estimées de T_n

Lorsque $\|R_n\|$ est estimé par une suite convolutive, on peut obtenir un développement asymptotique de T_n :

Proposition 2.2.18. *Soit R_n une suite d'opérateurs de renouvellement. On suppose que $\sum_{n+1}^{\infty} \|R_p\| = O(u_n)$, où u_n est une suite convolutive. Notons*

$$G(z) = R'(1) - \frac{R(z) - R(1)}{z-1} = \sum_{n=0}^{\infty} \left(\sum_{n+1}^{\infty} R_p \right) (1 - z^n).$$

Alors, pour tout $N \in \mathbb{N}$, il existe une suite $F_n \in \mathrm{Hom}(\mathcal{L}, \mathcal{L})$ avec $\|F_n\| = O(u_n)$ telle que

$$\sum T_n z^n = \frac{1}{1-z} \sum_{k=0}^{N-1} \left[\frac{1}{\mu} P \cdot G(z) \right]^k \frac{1}{\mu} P + \frac{1}{1-z} \left[\frac{1}{\mu} P \cdot G(z) \right]^N S(z)^{-1} + \sum F_n z^n. \quad (2.15)$$

Le terme en $k = 0$ donne $\frac{1}{\mu} P$, qui est la limite de T_n. Les autres termes tendent vers 0, puisqu'on vérifie aisément que $\frac{G(z)}{1-z}$ a ses coefficients qui tendent vers 0 car $G(1) = 0$. L'équation (2.15) donne donc un développement perturbatif de T_n, qui est d'autant plus précis que N est grand. Les termes provenant de $\frac{1}{1-z} \left[\frac{1}{\mu} P \cdot G(z) \right]^N S(z)^{-1} + \sum F_n z^n$ donnent le terme d'erreur.

Démonstration. Soit $S_B(z) = \frac{I - R_B(z)}{1-z}$, où R_B est donné par le lemme 2.2.8. Alors

$$S(z)^{-1} = \sum_{k=0}^{N-1} \left[S_B(z)^{-1} (S_B(z) - S(z)) \right]^k S_B(z)^{-1} + \left[S_B(z)^{-1} (S_B(z) - S(z)) \right]^N S(z)^{-1}. \quad (2.16)$$

Mais $S_B(z)^{-1} = \frac{1}{\mu} P + (1-z) E(z)$, où les coefficients de $E(z)$ décroissent exponentiellement vite, par le lemme 2.2.8. En notant $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$, on a donc $E(z) \in \mathcal{O}_u(\mathcal{A})$. De plus, $S_B(z) - S(z) = G(z) + (1-z) H(z)$ où H est un polynôme, par définition de S_B. Par hypothèse, $G(z) \in \mathcal{O}_u(\mathcal{A})$.

Dans le membre de droite de (2.16), on remplace tous les $S_B(z)-S(z)$ par $G(z)+(1-z)H(z)$, et tous les $S_B(z)^{-1}$ par $\frac{1}{\mu}P+(1-z)E(z)$. En développant, on obtient

$$S(z)^{-1}=\sum_{k=0}^{N-1}\left[\frac{1}{\mu}P\cdot G(z)\right]^k\frac{1}{\mu}P+\left[\frac{1}{\mu}P\cdot G(z)\right]^N S(z)^{-1}+(1-z)F(z)$$

où $F(z)$ est une combinaison linéaire de produits d'éléments de $\mathcal{O}_u(\mathcal{A})$, et est donc encore dans $\mathcal{O}_u(\mathcal{A})$ puisque c'est une algèbre. On en déduit (2.15) en divisant par $1-z$. □

La proposition qui suit est essentielle pour estimer le terme d'erreur qui fait intervenir $S(z)^{-1}$ dans (2.15).

Proposition 2.2.19. *Soit R_n une suite d'opérateurs de renouvellement. On suppose que $\sum_{n+1}^{\infty}\|R_p\|=O(u_n)$, où u_n est une suite convolutive. Alors $S(z)^{-1}\in\mathcal{O}_u(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$.*

Démonstration. On va utiliser la preuve du théorème 2.2.5 : conformément à la remarque 2.2.6, cette preuve fonctionne dans d'autres algèbres de Banach. Plus précisément, les seules propriétés utilisées dans la preuve sont le lemme de Wiener, et l'existence de partitions de l'unité à coefficients dans cette algèbre. Ces propriétés sont vraies dans $\mathcal{O}_u(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$, d'après le théorème 2.2.16 pour le lemme de Wiener, et la définition des suites convolutives pour la partition de l'unité.

Dans la deuxième étape de la preuve, il est également important de s'assurer que la fonction auxiliaire ν est à coefficients dans $\mathcal{O}_u(\mathbb{C})$. Il suffit pour cela de prendre un polynôme pour ν. □

En utilisant le développement de la proposition 2.2.18 pour $N=2$, on obtient :

Théorème 2.2.20. *Soit R_n une suite d'opérateurs de renouvellement, avec $\sum_{n+1}^{\infty}\|R_p\|=O(u_n)$ où u_n est une suite convolutive. Alors*

$$T_n=\frac{1}{\mu}P+\frac{1}{\mu^2}\sum_{p=n+1}^{\infty}P\left(\sum_{p+1}^{\infty}R_q\right)P+o\left(\sum_{n+1}^{\infty}u_p\right). \tag{2.17}$$

De plus, quand u_n est fortement convolutive, on peut remplacer dans cette équation le terme d'erreur $o(\sum_{n+1}^{\infty}u_p)$ par $O\left(\max\left(u_n,\frac{1}{n}\sum_{n+1}^{\infty}u_p\right)\right)$.

Quand $\sum_{n+1}^{\infty}R_p$ est vraiment de l'ordre de u_n, alors le terme $\sum_{n+1}^{\infty}P\left(\sum_{p+1}^{\infty}R_q\right)P$ est de l'ordre de $\sum_{n+1}^{\infty}u_p$, donc (2.17) donne vraiment le premier terme du développement asymptotique de T_n.

Démonstration. On utilise (2.15) pour $N = 2$. Le terme en $k = 0$ donne $\frac{1}{\mu}P$. Le terme en $k = 1$ donne $\frac{1}{\mu}P \cdot \frac{G(z)}{1-z} \cdot \frac{1}{\mu}P$. Mais $G(z) = \sum_n \left(\sum_{n+1}^{\infty} R_p\right)(1 - z^n)$, donc

$$\frac{G(z)}{1-z} = \sum_n \left(\sum_{p=n+1}^{\infty} \left(\sum_{p+1}^{\infty} R_q \right) \right) z^n,$$

et on obtient le deuxième terme de (2.17).

Il reste à estimer le terme d'erreur, correspondant à $H(z) = \frac{1}{\mu^2(1-z)}(PG(z))^2 S(z)^{-1}$. Lorsque $K(z) = \sum K_n z^n$, on notera $K(z) = O(v_n)$ si $\|K_n\| = O(v_n)$. Soit $\widetilde{G}(z) = \frac{1}{\mu}PG(z)$: ainsi, $\widetilde{G}(1) = 0$ et $\widetilde{G}(z) = O(u_n)$, donc $\frac{\widetilde{G}(z)}{1-z} = O(w_n)$ où $w_n = \sum_{n+1}^{\infty} u_p$.

On dérive l'égalité $H(z) = \frac{1}{1-z}\widetilde{G}(z)^2 S(z)^{-1}$:

$$H'(z) = \left(\frac{\widetilde{G}(z)}{1-z}\right)^2 S(z)^{-1} + \frac{\widetilde{G}(z)}{1-z}\widetilde{G}'(z)S(z)^{-1} + \widetilde{G}'(z)\frac{\widetilde{G}(z)}{1-z}S(z)^{-1} + \frac{\widetilde{G}(z)}{1-z}\widetilde{G}(z)\left(S(z)^{-1}\right)'.$$

On ne peut pas regrouper les termes à cause de la non-commutativité, mais ce n'est pas un problème. On a $\frac{\widetilde{G}(z)}{1-z} = O(w_n)$, $\widetilde{G}(z) = O(u_n)$ donc $\widetilde{G}'(z) = O(nu_n)$, et $S(z)^{-1} = O(u_n)$ (par la proposition 2.2.19) donc $(S(z)^{-1})' = O(nu_n)$. En convolant, on obtient que $H'(z) = O(w \star w \star u + w \star (nu_n) \star u)$. Mais $w \star u = O(w)$ par le lemme 2.2.17, et on obtient après intégration que

$$H(z) = O\left(\frac{(w \star w)_n}{n} + \frac{(w \star (ku_k))_n}{n}\right). \tag{2.18}$$

Par définition des suites convolutives, on obtient $H(z) = o(w_n)$, ce qui conclut.

Finalement, dans le cas fortement convolutif, le terme d'erreur est en $O\left(\max(u_n, w_n/n)\right)$. □

Remarque 2.2.21. Même lorsque u_n n'est pas fortement convolutive, la preuve donne une estimée sur le terme d'erreur : il est en $O\left(\frac{(w\star w)_n}{n} + \frac{(w\star(ku_k))_n}{n}\right)$ d'après (2.18).

2.3 Vitesse de décorrélation

Dans cette partie, on estime précisément la vitesse de décorrélation des tours localement Gibbs-Markov, en utilisant les informations sur T_n données par le théorème 2.2.20, et l'équation (2.4).

Dans toute cette partie, $T : X \to X$ désignera une tour localement Gibbs-Markov mélangeante, qui préserve une mesure de probabilité m. On utilisera les notations et les résultats de la partie 2.1 ; en particulier, Y désigne la base de la tour X et φ_Y le temps de retour à Y. Les opérateurs R_n définis dans la partie 2.1 forment alors une suite d'opérateurs de

renouvellement (avec $\mu = m(Y)$ et $Pu = \frac{1}{m(Y)} \int_Y u$), ce qui permet d'appliquer les résultats de la partie 2.2 pour obtenir des informations sur $T_n = 1_Y \widehat{T}^n 1_Y$ et, de là, sur la vitesse de décorrélation.

2.3.1 Minoration de la vitesse de décorrélation

Théorème 2.3.1. *On suppose que $m[\varphi_Y > n] = O(u_n)$, où u_n est une suite convolutive. Alors, pour toutes fonctions $f : X \to \mathbb{R}$ lipschitzienne sur chaque élément de α et bornée, et $g : X \to \mathbb{R}$ intégrable, toutes deux nulles hors de Y, on a*

$$\int f \cdot g \circ T^n - \left(\int f\right)\left(\int g\right) = \left(\sum_{p=n+1}^{\infty} m[\varphi_Y > p]\right)\left(\int f\right)\left(\int g\right) + o\left(\sum_{p=n+1}^{\infty} u_p\right).$$

De plus, quand la suite u_n est fortement convolutive, on peut remplacer le terme d'erreur par $O\left(\max\left(u_n, \frac{1}{n}\sum_{p=n+1}^{\infty} u_p\right)\right)$.

Démonstration. Comme f et g sont supportées par Y, on a

$$\int f \cdot g \circ T^n = \int \widehat{T}^n(f) \cdot g = \int 1_Y \widehat{T}^n(1_Y f) \cdot g = \int T_n(f) \cdot g.$$

Par hypothèse, $f \in \mathcal{L}$, si bien que le théorème 2.2.20 s'applique et donne

$$T_n(f) = \frac{1}{\mu} P(f) + \frac{1}{\mu^2} \sum_{p=n+1}^{\infty} P\left(\sum_{p+1}^{\infty} R_q\right) P(f) + o\left(\sum_{n+1}^{\infty} u_p\right), \tag{2.19}$$

où $P(u) = \frac{1}{m(Y)} \int_Y u$ et $\mu = \frac{1}{m(Y)}$ d'après les lemmes 2.1.3 et 2.1.5. Ainsi, le premier terme du développement de $T_n(f)$ donne $(\int f) 1_Y$. Comme $PR_nP = \frac{m[\varphi_Y = n]}{m(Y)} P$ (voir la preuve du lemme 2.1.5), le second terme donne $\left(\sum_{p=n+1}^{\infty} m[\varphi_Y > p]\right)\left(\int f\right) 1_Y$. En multipliant par g et en intégrant, on obtient le résultat annoncé.

Dans le cas fortement convolutif, on a un meilleur terme d'erreur dans (2.19), qui donne un meilleur terme d'erreur dans l'estimée finale. □

En particulier, quand $m[\varphi_Y > p]$ est du même ordre de grandeur que u_p, le terme d'erreur est négligeable par rapport à $\sum_{n+1}^{\infty} m[\varphi_Y > p]$. On obtient donc un équivalent de la vitesse de décorrélation (et en particulier une minoration) lorsque $\int f$ et $\int g$ sont non nulles. Il faut noter que, pour des fonctions d'intégrale nulle, la décorrélation est plus rapide !

Remarque 2.3.2. Même dans le cas où u_n est seulement convolutive, on peut majorer le terme d'erreur par $O\left(\frac{(w \star w)_n}{n} + \frac{(w \star (ku_k))_n}{n}\right)$, d'après la remarque 2.2.21.

2.3.2 Majoration de la vitesse de décorrélation

Le théorème 2.3.1 n'est valable que pour des fonctions supportées dans Y. Dans ce paragraphe, on va étendre ces résultats à des fonctions lipschitziennes sur tout l'espace (mais on n'obtiendra plus que des bornes supérieures). On va pour cela utiliser (2.4), et il faut donc d'abord estimer les opérateurs A_a, B_b et C_n.

Lemme 2.3.3. *Pour $u : X \to \mathbb{R}$ bornée, on a : $\forall a \in \mathbb{N}$, $\|A_a(u)\|_1 \leqslant m[\varphi_Y > a] \|u\|_\infty$.*

Démonstration. Rappelons que l'opérateur A_a est défini par

$$A_a u(x) = \sum_{\substack{T^a y = x \\ y \in Y, Ty, \dots, T^a y \notin Y}} g^{(a)}(y) u(y).$$

En particulier, $A_a u$ est supporté par les points d'altitude a dans la tour. De plus, tous les jacobiens $g^{(a)}(y)$ qui interviennent dans l'expression de $A_a u(x)$ sont égaux à 1. Ainsi, $\|A_a u\|_\infty \leqslant \|u\|_\infty$, et $A_a u$ est supporté par un ensemble de mesure au plus $m[\varphi_Y > a]$. □

Rappelons que $\mathcal{L}$ désigne l'espace des fonctions lipschitziennes sur chaque élément de partition, et bornées.

Lemme 2.3.4. *Il existe une constante $C > 0$ telle que, pour tout $b \in \mathbb{N}$, $\|B_b\|_{\mathcal{L}} \leqslant C m[\varphi_Y > b]$. De plus, pour toute fonction $u : X \to \mathbb{R}$ appartenant à $\mathcal{L}$, on a $\sum_{b=0}^\infty \int B_b f = \int f$.*

Démonstration. Soit Λ_b l'ensemble des points de X qui entrent dans Y après exactement b itérations. Comme

$$B_b u(x) = \sum_{\substack{T^b y = x \\ y, \dots, T^{b-1} y \notin Y, T^b y \in Y}} g^{(b)}(y) u(y),$$

$B_b u$ est supporté par Y, et ne dépend que des valeurs de u sur Λ_b. Comme $g^{(b)}$ est l'inverse du jacobien de T^b, un changement de variables donne $\int B_b u = \int_{\Lambda_b} u$. Comme $\bigsqcup \Lambda_b = X$, on obtient bien $\sum \int B_b f = \int f$.

Finalement, le lemme 1.1.13 et la propriété de grande image des retours à la base assurent que $\|B_b\|_{\mathcal{L}} \leqslant C m[\Lambda_b] = C m[\varphi_Y > b]$. □

Lemme 2.3.5. *Pour $u : X \to \mathbb{R}$ bornée, on a : $\forall n \in \mathbb{N}$, $\|C_n(u)\|_1 \leqslant \sum_{p>n} m[\varphi_Y > p] \|u\|_\infty$.*

Démonstration. Notons $X_{n+1} = X \setminus \bigcup_{i=0}^n T^{-i} Y$ et $Z_{n+1} = T^n(X_{n+1})$. Alors $C_n(u)$ est nulle hors de Z_{n+1}. Soit $x \in Z_{n+1}$ et y sa préimage par T^n. Alors $|C_n(u)(x)| = |u(y)| \leqslant \|u\|_\infty$. Ainsi, $\|C_n(u)\|_1 \leqslant m(Z_{n+1}) \|u\|_\infty$. Finalement, $m[Z_{n+1}] = m[X_{n+1}] = \sum_{p=n+1}^\infty m[\varphi_Y > p]$. □

Théorème 2.3.6. *On suppose que $m[\varphi_Y > n] = O(u_n)$, où u_n est une suite convolutive. Alors il existe une constante $C > 0$ telle que, pour toute fonction $f : X \to \mathbb{R}$ appartenant à $\mathcal{L}$, pour tout $n \in \mathbb{N}$, $\left\| \widehat{T}^n(f) - \int f \right\|_1 \leqslant C \left(\sum_{n+1}^{\infty} u_p\right) \|f\|_{\mathcal{L}}$.*
En particulier, pour $f \in \mathcal{L}$ et $g \in L^\infty$,

$$\left| \int f \cdot g \circ T^n - \left(\int f\right)\left(\int g\right) \right| \leqslant C \left(\sum_{n+1}^{\infty} u_p\right) \|f\|_{\mathcal{L}} \|g\|_\infty .$$

En particulier, quand $\sum_{n+1}^{\infty} u_p$ est sommable, les corrélations sont sommables, si bien que les fonctions hölderiennes satisfont un théorème central limite, d'après le théorème de Gordin–Liverani cité dans l'introduction.

Démonstration. On peut supposer que $\int f = 0$.
D'après le théorème 2.2.20, on peut écrire $T_n = \frac{1}{\mu} P + E_n$, où $\|E_n\| = O(w_n)$ avec $w_n = \sum_{n+1}^{\infty} u_p$, et $\frac{1}{\mu} P(u) = \left(\int_Y u\right) 1_Y$ par les lemmes 2.1.3 et 2.1.5. Ainsi, par (2.4),

$$\widehat{T}^n(f) = C_n f + \sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f) + \sum_{a+k+b=n} A_a E_k B_b(f). \tag{2.20}$$

D'après le lemme 2.3.5, on a

$$\|C_n f\|_1 \leqslant \sum_{p=n+1}^{\infty} m[\varphi_Y > p] \|f\|_\infty \leqslant C w_n \|f\|_\infty .$$

Majorons ensuite le dernier terme de (2.20), en utilisant les lemmes 2.3.3 et 2.3.4 :

$$\begin{aligned} \left\| \sum_{a+k+b=n} A_a E_k B_b(f) \right\|_1 &\leqslant \sum_{a+k+b=n} m[\varphi_Y > a] \|E_k B_b(f)\|_\infty \\ &\leqslant \sum_{a+k+b=n} m[\varphi_Y > a] \|E_k\|_{\mathcal{L}} \|B_b\|_{\mathcal{L}} \|f\|_{\mathcal{L}} . \end{aligned}$$

On a $\|E_k\|_{\mathcal{L}} = O(w_k)$ et $\|B_b\|_{\mathcal{L}} = O(u_b)$, donc ce terme est majoré par $(u \star w \star u)_n$, qui est un $O(w_n)$ par le lemme 2.2.17.

Soit maintenant $Z \subset X$, et majorons $\int_Z \left(\sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f)\right)$. Notons que $\frac{1}{\mu} P B_b(f) = \left(\int_Y B_b(f)\right) 1_Y = \left(\int_{\Lambda_b} f\right) 1_Y$, où Λ_b est l'ensemble des points qui rentrent dans Y après b itérations. Soit $\lambda_b = \int_{\Lambda_b} f$, cette suite vérifie donc $|\lambda_b| \leqslant m(\Lambda_b) \|f\|_\infty = O(u_b)$. De plus,

$\sum \lambda_b = \int f = 0$. Notons également $\alpha_a = \int_Z A_a(1_Y)$: on a $\alpha_a = O(u_a)$ par le lemme 2.3.3. Finalement,

$$\int_Z \left(\sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f) \right) = \sum_{a+k+b=n} \alpha_a \lambda_b = \sum_{a=0}^{n} \alpha_a \left(\sum_{b=0}^{n-a} \lambda_b \right) = -\sum_{a=0}^{n} \alpha_a \left(\sum_{b=n-a+1}^{\infty} \lambda_b \right), \tag{2.21}$$

puisque $\sum \lambda_b = 0$. Comme $\sum_{b=n-a+1}^{\infty} \lambda_b = O(w_{n-a})$, (2.21) est donc majoré en valeur absolue par $(u \star w)_n = O(w_n)$ par le lemme 2.2.17.

On applique ce résultat à $Z_1 = \left\{ x \in X \mid \left(\sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f) \right)(x) > 0 \right\}$, puis à $Z_2 = \left\{ x \in X \mid \left(\sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f) \right)(x) < 0 \right\}$. En sommant, on obtient finalement

$$\left\| \sum_{a+k+b=n} A_a \frac{1}{\mu} P B_b(f) \right\|_1 \leqslant C w_n \|f\|_\infty ,$$

ce qui termine la preuve. □

Remarque 2.3.7. Dans le cas où $m[\varphi_Y > n]$ décroît exponentiellement vite, on peut aussi obtenir une décroissance exponentielle de $\left\| \widehat{T}^n f - \int f \right\|_1$, même si les suites e^{-cn} ne sont pas convolutives.

En effet, les arguments utilisés pour estimer $S_B(z)^{-1}$ dans le lemme 2.2.8 montrent que $S(z)^{-1}$ s'étend alors holomorphiquement à une couronne, puis que $T_n = \frac{1}{\mu} P + E_n$ où E_n décroît exponentiellement vite. En suivant la preuve du théorème 2.3.6, on obtient bien une décroissance exponentielle, en $O(\kappa^n)$ pour un certain $\kappa < 1$.

La différence principale avec le cas convolutif est qu'on n'a pas de contrôle *a priori* sur la vitesse finale, i.e. sur le coefficient κ, à cause des pôles qui peuvent apparaître dans la fonction $S(z)^{-1}$.

2.3.3 Exemples

Dans le cas des tours de Young, la borne supérieure sur la décorrélation était déjà connue dans de nombreux cas, par une méthode complètement différente : par des méthodes de couplage probabiliste, Young a traité dans [You99] le cas où $m[\varphi_Y > n] = O(1/n^\gamma)$ lorsque $\gamma > 1$, ce qui donne une vitesse de décorrélation en $O(1/n^{\gamma-1})$. En fait, sa méthode de preuve s'adapte aux suites à décroissance polynomiale. En revanche, dans le cas où $m[\varphi_Y > n] = O(e^{-cn^\eta})$, i.e. le cas polynomial gauche, Young obtient une décorrélation en $O(e^{-n^{\eta'}})$ pour tout $\eta' < \eta$, tandis que le théorème 2.3.6 montre que la décorrélation est majorée par $n^{1-\eta} e^{-cn^\eta}$: on ne perd non seulement pas au niveau de l'exposant, mais pas non plus au niveau de la constante c.

Les premiers résultats généraux de borne inférieure ont été obtenus par Sarig dans [Sar02] : il traitait le cas où $u_n = O(1/n^\gamma)$ avec $\gamma > 2$, sa méthode ne permettant pas d'atteindre des vitesses plus lentes car elle n'utilisait pas de techniques d'algèbres de Banach et reposait sur des arguments plus élémentaires. Il ne pouvait pas non plus obtenir de borne inférieure dans des cas plus que polynomiaux, par exemple dans le cas où $u_n = e^{-cn^\eta}$ avec $0 < \eta < 1$.

On va maintenant s'intéresser à des exemples en dimension 1, plus précisément aux applications LSV et à leurs généralisations par Holland, introduites respectivement dans les paragraphes 1.4.1 et 1.4.2.

Soit

$$T(x) = \begin{cases} x(1 + x^\alpha \rho(x)) & \text{si } 0 \leqslant x \leqslant 1/2, \\ 2x - 1 & \text{si } 1/2 < x \leqslant 1 \end{cases} \tag{2.22}$$

une application de Holland, i.e. $\alpha > 0$ et $\rho : \mathbb{R}_+^* \to \mathbb{R}_+^*$ est une fonction à variation lente en 0, de classe C^2, avec $x\rho'(x) = o(\rho(x))$ et $x^2\rho''(x) = o(\rho(x))$, normalisée de telle sorte que $T(1/2) = 1$, avec de plus, pour tout $x \neq 0$, $T'(x) > 1$.

On suppose également que $\int_0^1 \frac{1}{x^{1/\alpha}\rho(x)} < \infty$, si bien que T admet une unique mesure de probabilité m invariante absolument continue par rapport à la mesure de Lebesgue. De plus, l'application induite par T sur $Y =]1/2, 1]$ est Gibbs-Markov et, d'après le théorème 1.4.10 et la proposition 1.4.12,

$$m[\varphi_Y > n] \sim \frac{h(1/2)}{2(\alpha n)^{1/\alpha} l^\#(n^{1/\alpha})}, \tag{2.23}$$

où $l(x) = \frac{1}{\rho(1/x)^{1/\alpha}}$, $l^\#$ est la conjuguée de de Bruijn de l, et h est la densité de la mesure m.

Notons $u_n = \frac{1}{n^{1/\alpha} l^\#(n^{1/\alpha})}$: c'est une suite à décroissance polynomiale, sommable par la proposition 1.4.12. Elle est donc convolutive d'après la proposition 2.2.12. Les résultats de la partie précédente sur la vitesse de décorrélation donnent alors :

Théorème 2.3.8. *Soit T une application de Holland (2.22) admettant une mesure de probabilité invariante absolument continue m de densité h. Soit $0 < \nu \leqslant 1$. Alors il existe une constante $C(\nu) > 0$ telle que, pour toutes fonctions f ν-hölderienne et g bornée,*

$$\left| \int f \cdot g \circ T^n \, \mathrm{d}m - \left(\int f \, \mathrm{d}m \right) \left(\int g \, \mathrm{d}m \right) \right| \leqslant C(\nu) \, \|f\|_\nu \, \|g\|_\infty \sum_{p=n+1}^{\infty} \frac{1}{p^{1/\alpha} l^\#(p^{1/\alpha})} \tag{2.24}$$

où $l(x) = \frac{1}{\rho(1/x)^{1/\alpha}}$.

De plus, si f et g tendent vers 0 en 0, alors

$$\begin{aligned} &\int f \cdot g \circ T^n \, \mathrm{d}m - \left(\int f \, \mathrm{d}m \right) \left(\int g \, \mathrm{d}m \right) \\ &\qquad = \left(\sum_{p=n+1}^{\infty} \frac{h(1/2)}{2(\alpha p)^{1/\alpha} l^\#(p^{1/\alpha})} \right) \left(\int f \, \mathrm{d}m \right) \left(\int g \, \mathrm{d}m \right) + o\left(\sum_{p=n+1}^{\infty} \frac{1}{p^{1/\alpha} l^\#(p^{1/\alpha})} \right). \end{aligned}$$

Démonstration. Soit $Y =]1/2, 1]$, on induit sur Y. La mesure $m[\varphi_Y > n]$ est estimée par (2.23), et l'application induite par T sur Y est Gibbs-Markov. Ainsi, T admet une extension qui est une tour localement Gibbs-Markov, d'après la proposition 1.3.13. La borne supérieure sur la vitesse de décorrélation est alors une conséquence du théorème 2.3.6, appliqué dans cette extension. Dans le cas où f et g sont nulles hors de Y, la borne inférieure provient du théorème 2.3.1.

Montrons maintenant la deuxième égalité lorsque f et g sont nulles sur un voisinage de 0. Il existe $N > 0$ tel que f et g soient supportées dans $Y' = [x_N, 1]$. Comme l'application induite par T sur Y' est encore Gibbs-Markov, le théorème 2.3.1 implique encore qu'il existe une suite $w_n(N)$ (égale à $\sum_{n+1}^{\infty} m[\varphi_{Y'} > p]$) telle que $\mathrm{Cor}(f, g \circ T^n) = w_n(N) \left(\int f\right) \left(\int g\right) + o(w_n(N))$. Il reste à montrer que $w_n(N) \sim w_n(1)$ pour conclure. On utilise pour cela deux fonctions f, g supportées dans $]1/2, 1]$ et d'intégrale 1. Alors $\mathrm{Cor}(f, g \circ T^n) \sim w_n(1)$ en utilisant le résultat sur Y, et $\mathrm{Cor}(f, g \circ T^n) \sim w_n(N)$ en utilisant le résultat sur Y'. Cela montre que $w_n(N) \sim w_n(1)$.

Soient finalement f et g qui tendent vers 0 en 0. Soit $\varepsilon > 0$ très petit. Soit

$$a_\varepsilon(x) = \begin{cases} f(x) & \text{si } x \leqslant \delta \\ f(2\delta - x) & \text{si } \delta \leqslant x \leqslant 2\delta \\ 0 & \text{sinon} \end{cases}$$

où δ est assez petit pour que $\int |a_\varepsilon| \leqslant \varepsilon$. Soit aussi b_ε définie de la même manière à partir de g. Soit $\nu' \in]0, \nu[$. Quitte à diminuer δ, on peut aussi supposer que $\|a_\varepsilon\|_{\nu'} \leqslant \varepsilon$, $\|b_\varepsilon\|_\infty \leqslant \varepsilon$ et $\int |b_\varepsilon| \leqslant \varepsilon$.

Les fonctions $f_\varepsilon = f - a_\varepsilon$ et $g_\varepsilon = g - b_\varepsilon$ sont alors nulles sur un voisinage de 0, donc les résultats précédents s'appliquent. Soit $w_n = \sum_{p=n+1}^{\infty} \frac{h(1/2)}{2(\alpha p)^{1/\alpha} l^\#(p^{1/\alpha})}$. Pour n assez grand, on a donc

$$\begin{aligned} \int f_\varepsilon \cdot g_\varepsilon \circ T^n &= \left(\int f_\varepsilon\right)\left(\int g_\varepsilon\right) + w_n \left(\int f_\varepsilon\right)\left(\int g_\varepsilon\right) \pm \varepsilon w_n \\ &= \left(\int f_\varepsilon\right)\left(\int g_\varepsilon\right) + w_n \left(\int f\right)\left(\int g\right) \pm C\varepsilon w_n, \end{aligned}$$

où la notation $\pm\varepsilon w_n$ indique un terme d'erreur compris entre $-\varepsilon w_n$ et εw_n. D'après (2.24) appliqué dans l'espace des fonctions ν'-hölderiennes, on a

$$\begin{aligned} \int f \cdot g \circ T^n &= \int f_\varepsilon \cdot g_\varepsilon \circ T^n + \int a_\varepsilon \cdot g_\varepsilon \circ T^n + \int f \cdot b_\varepsilon \circ T^n \\ &= \int f_\varepsilon \cdot g_\varepsilon \circ T^n + \left(\int a_\varepsilon\right)\left(\int g_\varepsilon\right) + \left(\int f\right)\left(\int b_\varepsilon\right) \pm C\varepsilon w_n. \end{aligned}$$

Si n est assez grand, on obtient donc

$$\int f \cdot g \circ T^n = \left(\int f\right)\left(\int g\right) + w_n \left(\int f\right)\left(\int g\right) \pm C\varepsilon w_n. \qquad \square$$

Cela montre que la vitesse de décorrélation est exactement $w_n := \sum_{n+1}^{\infty} \frac{h(1/2)}{2(\alpha p)^{1/\alpha} l\#(p^{1/\alpha})}$. Lorsque $\alpha < 1$, cette suite est équivalente à $\frac{h(1/2)}{2(1/\alpha-1)\alpha^{1/\alpha} n^{1/\alpha-1} l\#(n^{1/\alpha})}$, par [BGT87, proposition 1.5.10]. Ainsi, on peut de cette manière obtenir toutes les vitesses de décorrélation $\frac{1}{n^\gamma L(n)}$ où $\gamma > 0$ et L est à variation lente.

Dans le cas où $\alpha = 1$, la fonction $L(x) = \int_x^\infty \frac{1}{tl\#(t)}$ est à variation lente et tend vers 0, par [BGT87, proposition 1.5.9b], et w_n est équivalent à $CL(n)$. Par exemple, prenons $\rho(x) = 1/l(1/x)$ sur un voisinage de 0, avec $l(x) = \left(\prod_1^{s-1} \log_k(x)\right) \log_s(x)^{1+\varepsilon}$, où $\log_k$ désigne le logarithme itéré k fois. Alors $\frac{1}{x\rho(x)}$ est intégrable en 0, donc l'application de Holland correspondante T admet une mesure de probabilité invariante. De plus, la conjuguée de de Bruijn de l est $1/l$ (voir la remarque 1.4.15), et on obtient $w_n \sim \frac{C}{(\log_s n)^\varepsilon}$. Ainsi, les vitesses de mélange possibles sont arbitrairement lentes.

Donnons par exemple le résultat obtenu dans le cas des applications LSV, en précisant le terme d'erreur grâce à la remarque 2.3.2 :

Proposition 2.3.9. *Soient $0 < \alpha < 1$ et T l'application LSV correspondante donnée par (1.25). Soit h la densité de sa mesure de probabilité invariante. Alors, pour toutes fonctions f hölderienne et g bornée,*

$$\int f \cdot g \circ T^n - \left(\int f\right)\left(\int g\right) = O\left(\frac{1}{n^{1/\alpha-1}}\right).$$

De plus, si f et g sont nulles sur un voisinage de 0,

$$\int f \cdot g \circ T^n - \left(\int f\right)\left(\int g\right) = \frac{h(1/2)}{4\left(\frac{1}{\alpha}-1\right)\alpha^{1/\alpha} n^{1/\alpha-1}} \left(\int f\right)\left(\int g\right) + O(E_\alpha(n))$$

où $E_\alpha(n) = \frac{1}{n^{1/\alpha}}$ si $\alpha < 1/2$, $\frac{\log n}{n^2}$ si $\alpha = 1/2$, et $\frac{1}{n^{2/\alpha-2}}$ si $\alpha > 1/2$.
Enfin, si f et g tendent vers 0 en 0,

$$\int f \cdot g \circ T^n - \left(\int f\right)\left(\int g\right) = \frac{h(1/2)}{4\left(\frac{1}{\alpha}-1\right)\alpha^{1/\alpha} n^{1/\alpha-1}} \left(\int f\right)\left(\int g\right) + o\left(\frac{1}{n^{1/\alpha-1}}\right).$$

En particulier, pour $\alpha < 1/2$, une fonction hölderienne a ses corrélations sommables et satisfait donc un théorème central limite, mais ce résultat ne dit rien sur les théorèmes limites pour $\alpha \geqslant 1/2$. On verra plus loin (corollaire 2.4.6) qu'une fonction d'intégrale nulle et nulle au voisinage de 0 a encore ses corrélations sommables, et on étendra même ce résultat aux fonctions lipschitziennes et nulles en 0 grâce au théorème 2.4.13. Ainsi, ces fonctions vérifient également un théorème central limite. Dans le chapitre 4, on décrira précisément (par une méthode complètement différente) toutes les lois limites qu'on peut obtenir.

Finalement, pour les produits gauches de la partie 1.5, on a $m[\varphi_Y > n] \sim A\left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}}$ d'après la proposition 1.5.18. Ainsi, les corrélations des fonctions hölderiennes décroissent en

$O\left(\frac{(\ln n)^{\frac{1}{2\alpha_{\min}}}}{n^{\frac{1}{\alpha_{\min}}-1}}\right)$, et elles décroissent exactement à cette vitesse pour des fonctions d'intégrale non nulle et qui s'annulent sur $S^1 \times \{0\}$.

2.4 Raffinements sur la vitesse de décorrélation

Dans cette partie, on décrit plusieurs raffinements possibles du théorème 2.2.20 et de ses conséquences sur la décorrélation des fonctions. Ces raffinements vont dépendre de la vitesse de décroissance de $\|R_n\|$, ou des particularités de la fonction f considérée.

2.4.1 Estimation à plusieurs termes de la décorrélation

Jusqu'à maintenant, on n'a utilisé que les deux premiers termes dans l'estimation de T_n donnée par la proposition 2.2.18, c'est-à-dire que l'on n'a utilisé cette proposition que pour $N = 2$. Cela suffit pour obtenir des bornes supérieures générales sur la décorrélation, ainsi que des bornes inférieures pour les fonctions nulles hors de Y, mais le terme d'erreur obtenu n'est pas toujours optimal. Par exemple, dans la proposition 2.3.9, on n'obtient un terme d'erreur en $O(1/n^{1/\alpha})$ que dans le cas où $\alpha < 1/2$, tandis que $\|T_n - T_{n+1}\| = O(1/n^{1/\alpha})$ dans tous les cas d'après la proposition 2.2.19, ce qui permet en général d'espérer un développement jusqu'à une précision $O(1/n^{1/\alpha})$. Il est donc naturel de chercher à utiliser la proposition 2.2.18 avec plus de termes, pour obtenir un développement asymptotique plus précis de T_n. C'est ce qu'on va faire dans cette partie, en se restreignant au cas où $m[\varphi_Y > n] = O(1/n^\beta)$ avec $\beta > 1$, ce qui traite le cas des applications LSV (pour $\beta = 1/\alpha$).

Pour estimer le terme d'erreur provenant de $\frac{1}{1-z}\left[\frac{1}{\mu}P \cdot G(z)\right]^N S(z)^{-1}$ dans la proposition 2.2.18, on va utiliser la proposition suivante. Les seules hypothèses utiles seront que $G(1) = 0$ et $G(z) = O(1/n^\beta)$ (i.e. le coefficient de z^n dans $G(z)$ est un $O(1/n^\beta)$). Cela implique que $\frac{G(z)}{1-z} = O(1/n^{\beta-1})$.

Proposition 2.4.1. *Soit $G(z) = \sum G_n z^n$ une série formelle à coefficients dans une algèbre de Banach $\mathcal{A}$, telle que $G(z) = O(1/n^\beta)$ pour un certain $\beta > 1$ et $G(1) = 0$. Alors, pour tout $p \in \mathbb{N}$, il existe une constante C telle que, pour tous $H_1, \dots, H_{p-1} \in \mathcal{A}$,*

$$\left\| \left(\frac{G(z)H_1G(z)\dots H_{p-1}G(z)}{1-z} \right)_n \right\| \leqslant C \|H_1\| \dots \|H_{p-1}\| \cdot \begin{cases} \frac{1}{n^\beta} & \text{si } \beta < p(\beta-1), \\ \frac{\log n}{n^\beta} & \text{si } \beta = p(\beta-1), \\ \frac{1}{n^{p(\beta-1)}} & \text{si } \beta > p(\beta-1). \end{cases}$$

Ici, la notation $(\ \ )_n$ désigne le coefficient de z^n dans la série formelle entre les parenthèses.

Dans la preuve de la proposition, on utilisera les deux lemmes techniques suivants, dont les preuves reposent sur les mêmes arguments que la preuve de la proposition 2.2.12 :

Lemme 2.4.2. *Si $a_n = O(1/n^\alpha)$ et $b_n = O(1/n^\beta)$, avec $\alpha \leqslant \beta \in \mathbb{R}$, alors*

$$(a \star b)_n = \begin{cases} O(1/n^\alpha) & \text{si } \beta > 1, \\ O(\log n/n^\alpha) & \text{si } \beta = 1, \\ O(1/n^{\alpha+\beta-1}) & \text{si } \beta < 1. \end{cases} \tag{2.25}$$

En particulier, pour $\alpha > 1$ ou $\beta > 1$ (sans supposer $\alpha \leqslant \beta$), $(a \star b)_n = O(1/n^\alpha) + O(1/n^\beta)$.

Lemme 2.4.3. *Si $a_n = O\left(\frac{(\log n)^u}{n^\alpha}\right)$ et $b_n = O\left(\frac{(\log n)^v}{n^\beta}\right)$ avec $\alpha \leqslant 1, \beta \leqslant 1$ et $u, v \geqslant 0$, alors $(a \star b)_n = O\left(\frac{(\log n)^{u+v+1}}{n^{\alpha+\beta-1}}\right)$.*

En fait, le facteur $(\log n)^{u+v+1}$ peut être remplacé par $(\log n)^{u+v}$ dès que $\alpha < 1$ et $\beta < 1$, mais on n'en aura pas besoin.

Démonstration de la proposition 2.4.1. Dans cette preuve, si $L(z) = \sum L_n z^n$ est une série entière à coefficients dans $\mathcal{A}$, on écrira $\|L\|$ ou $\|L(z)\|$ pour la suite $(\|L_n\|)_{n\in\mathbb{N}}$. Si $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites de nombres positifs ou nuls tels que $\forall n, a_n \leqslant b_n$, on écrira $a \leqslant b$. Avec ces notations, si $L_0(z)$ et $L_1(z)$ sont des séries entières, alors $\|L_0 L_1\| \leqslant \|L_0\| \star \|L_1\|$, où $\star$ désigne la convolution des suites. Notons aussi

$$F_{\beta,p}(n) = \begin{cases} \frac{1}{n^\beta} & \text{si } \beta < p(\beta-1), \\ \frac{\log n}{n^\beta} & \text{si } \beta = p(\beta-1), \\ \frac{1}{n^{p(\beta-1)}} & \text{si } \beta > p(\beta-1). \end{cases}$$

On va démontrer le lemme par récurrence sur p. L'argument central pour se débarrasser des problèmes de non-commutativité sera le suivant : si $L_0(z), \dots, L_r(z)$ sont des séries entières pour un certain $r < p$, considérons $\frac{L_0(z)G(z)L_1(z)\dots L_{r-1}(z)G(z)L_r(z)}{1-z}$. L'hypothèse de récurrence donne $\left\|\frac{G(z)H_1G(z)\dots H_{r-1}G(z)}{1-z}\right\| \leqslant CF_{\beta,r}(n)\,\|H_1\|\dots\|H_{r-1}\|$ pour une certaine constante C. Alors

$$\left\|\frac{L_0(z)G(z)L_1(z)\dots L_{r-1}(z)G(z)L_r(z)}{1-z}\right\| \leqslant CF_{\beta,r} \star \|L_0\| \star \dots \star \|L_r\|\,. \tag{2.26}$$

En effet, si on développe dans $\frac{L_0(z)G(z)L_1(z)\dots L_{r-1}(z)G(z)L_r(z)}{1-z}$ chaque facteur $L_i(z)$ en série, on obtient des termes de la forme $\frac{H_0G(z)H_1\dots H_{r-1}G(z)H_r}{1-z}$, auxquels on peut appliquer l'hypothèse de récurrence.

On prouve la proposition 2.4.1 par récurrence sur p. Le résultat est immédiat pour $p = 1$, et il découle facilement de la preuve du théorème 2.2.20 pour $p = 2$: le même argument marche encore quand on ajoute un terme H_1. Supposons donc $p > 2$.

Si $\beta < (p-1)(\beta-1)$, l'hypothèse de récurrence donne

$$\left\|\frac{G(z)H_1G(z)\dots H_{p-2}G(z)}{1-z}\right\| \leqslant \frac{C}{n^\beta}\,\|H_1\|\dots\|H_{p-2}\|\,.$$

Comme $G(z) \in O(1/n^\beta)$, une multiplication à droite par H_{p-1} et une convolution par $G(z)$ donnent le résultat, en utilisant le lemme 2.4.2. On peut donc supposer que $\beta \geqslant (p-1)(\beta-1)$. Comme $p \geqslant 3$, cela implique en particulier que $\beta \leqslant 2$.

En dérivant $p-1$ fois $F(z) = \frac{G(z)H_1G(z)\dots H_{p-1}G(z)}{1-z}$, on obtient, pour certaines constantes $C_{i,i_1,\dots,i_p}$,

$$F^{(p-1)}(z) = \sum_{\substack{i+i_1+\dots+i_p=p-1 \\ i_1,\dots,i_p\geqslant 0}} C_{i,i_1,\dots,i_p} \frac{G(z)^{(i_1)}H_1\dots H_{p-1}G(z)^{(i_p)}}{(1-z)^{i+1}},$$

où $G(z)^{(k)}$ désigne la dérivée k-ème de $G(z)$.

Soit $T(z) = T_{i,i_1,\dots,i_p}(z)$ un des termes de cette somme. Comme $i + \sum_{j=1}^p i_j = p-1$, le nombre de facteurs $G(z)$ qui ne sont pas dérivés (i.e. pour lesquels $i_j = 0$) est de la forme $i+r$ avec $r \geqslant 1$. Attribuons un facteur $1/(1-z)$ à i facteurs $G(z)$ non dérivés, on obtient :

$$T(z) = \frac{L_0(z)G(z)L_1(z)\dots G(z)L_r(z)}{1-z},$$

où les facteurs $L_i(z)$ sont des produits de facteurs de la forme H_j, ou $G^{(i_j)}$ avec $i_j \geqslant 1$, ou $\frac{G(z)}{1-z}$. L'équation (2.26) donne alors

$$\begin{aligned} \|T(z)\| &\leqslant C F_{\beta,r} \star \|L_0\| \star \dots \star \|L_r\| \\ &\leqslant C \|H_1\| \dots \|H_{p-1}\| \cdot F_{\beta,r} \star \left\|G^{(i_1)}\right\| \star \dots \star \left\|G^{(i_p)}\right\| \star \left\|\frac{G(z)}{1-z}\right\|^{\star i}, \end{aligned} \tag{2.27}$$

où $\left\|G^{(i_j)}\right\|$ n'est écrit dans le produit que lorsque $i_j > 0$.

On distingue trois cas :

1. Si $(p-1)(\beta-1) \leqslant \beta < p(\beta-1)$.

   Dans chaque terme $T(z) = T_{i,i_1,\dots,i_p}(z)$ (dans lequel $i+r$ facteurs $G(z)$ ne sont pas dérivés), on a $r < p$: sinon, $i + r \leqslant p$ impliquerait que $i = 0$, et rien ne serait dérivé. Ainsi, $r(\beta-1) \leqslant \beta$, et on est dans le deuxième ou le troisième cas de la récurrence.

   On n'est en fait dans le deuxième cas que si $r = p-1$, ce qui signifie que $i = 0$ et que l'un des i_j est égal à $p-1$ tandis que les autres sont nuls. Dans (2.27), $F_{\beta,p-1}(n) = O\left(\frac{\log n}{n^{(p-1)(\beta-1)}}\right) = O(1/n^\gamma)$ pour un certain $\gamma > 1$ (car $p(\beta-1) > \beta$ implique que $(p-1)(\beta-1) > 1$). Quand on convole avec $\left\|G(z)^{(p-1)}\right\| = O(\frac{1}{n^{\beta-p+1}})$ (où $\beta - p + 1 \leqslant 1$ puisque $\beta \leqslant 2$), le lemme 2.4.2 montre qu'on obtient un élément de $O(\frac{1}{n^{\beta-p+1}})$.

   Considérons maintenant un autre terme $T(z) = T_{i,i_1,\dots,i_p}(z)$ avec $r < p-1$. Comme $(p-1)(\beta-1) \leqslant \beta$, on obtient $(p-2)(\beta-1) \leqslant 1$, donc $r(\beta-1) \leqslant 1$. Notons que $F_{\beta,r}(n) = O(1/n^{r(\beta-1)})$ et $\left\|G^{(i_j)}\right\| \in O(1/n^{\beta-i_j})$ (avec $\beta - i_j \leqslant 1$ lorsque $i_j > 0$ puisque $\beta \leqslant 2$) et $\frac{G(z)}{1-z} \in O(1/n^{\beta-1})$ (avec $\beta - 1 \leqslant 1$). Ainsi, le lemme 2.4.3 appliqué $p-r$ fois

montre que le terme de droite de l'équation (2.27) est dans $O(\frac{(\log n)^u}{n^\nu})$ pour un certain $u \geqslant 0$ et

$$\nu = \sum_{i_j>0}(\beta - i_j) + i(\beta-1) + r(\beta-1) - (p-r) = p(\beta-1) - (p-1).$$

Comme $p(\beta-1) > \beta$, on a en fait $\nu > \beta - (p-1)$, et on obtient donc une majoration de $\|T(z)\|$ de la forme $\frac{C}{n^{\beta-(p-1)}}$.

En ajoutant tous les termes, on obtient $\left\|F(z)^{(p-1)}\right\| \leqslant C\,\|H_1\|\dots\|H_{p-1}\|\,\frac{1}{n^{\beta-p+1}}$. En intégrant $p-1$ fois, on obtient le résultat désiré.

2. Si $\beta = p(\beta-1)$.

   Ici, on a $\beta < 2$, donc $\beta - i_j < 1$ et $\beta - 1 < 1$, ce qui implique qu'on va pouvoir utiliser le lemme 2.4.2 à la place du lemme 2.4.3.

   On utilise le même raisonnement que dans le premier cas. Pour les termes avec $r = p-1$, on a $F_{\beta,p-1}(n) = O(1/n)$, ce qui implique que le terme de droite de (2.27) est majoré par $(\frac{1}{n}) \star (\frac{1}{n^{\beta-p+1}})$, et donc par $O(\frac{\log n}{n^{\beta-p+1}})$ d'après le lemme 2.4.2.

   Lorsque $r < p-1$, on utilise le lemme 2.4.2 $p-r-1$ fois pour estimer $\left\|G^{(i_1)}\right\|\star\dots\left\|G^{(i_p)}\right\|\star \left\|\frac{G(z)}{1-z}\right\|^{\star i}$, puisque tous les exposants sont < 1. On obtient que c'est dans $O(1/n^\delta)$, où $\delta = (p-r)(\beta-1) - p + 2$. En convolant une fois de plus avec $F_{\beta,r}(n) = 1/n^{r(\beta-1)}$, on obtient un terme dans $O(1/n^\nu)$ où $\nu = p(\beta-1) - p + 1 = \beta - p + 1$. On additionne les différents termes puis on intègre, ce qui donne le résultat.

3. Si $\beta > p(\beta-1)$.

   On n'a plus besoin de traiter à part les termes avec $r = p-1$: dans tous les termes, tous les exposants sont < 1. En convolant, on obtient des termes dans $O(\frac{1}{n^{p(\beta-1)-p+1}})$, ce qui donne le résultat recherché après intégration. □

Proposition 2.4.4. *Soit R_n une suite d'opérateurs de renouvellement, avec $\sum_{n+1}^\infty \|R_p\| = O(1/n^\beta)$ pour un certain $\beta > 1$. Alors, pour tout $N \in \mathbb{N}$, il existe une suite $F_n \in \mathrm{Hom}(\mathcal{L},\mathcal{L})$ avec*

$$\|F_n\| = \begin{cases} O(1/n^\beta) & \text{si } N(\beta-1) > \beta, \\ O(\log n/n^\beta) & \text{si } N(\beta-1) = \beta, \\ O(1/n^{N(\beta-1)}) & \text{si } N(\beta-1) < \beta, \end{cases}$$

telle que

$$\sum T_n z^n = \frac{1}{1-z}\sum_{k=0}^{N-1}\left[\frac{1}{\mu}P\cdot G(z)\right]^k \frac{1}{\mu}P + \sum F_n z^n,$$

où $G(z) = R'(1) - \frac{R(z)-R(1)}{z-1} = \sum_{n=0}^\infty \left(\sum_{p=n+1}^\infty R_p\right)(1-z^n)$.

Démonstration. C'est une conséquence directe des propositions 2.2.18, 2.2.19 et 2.4.1. □

Notons que, pour tout $\beta > 1$, il est possible de choisir N de telle sorte que $N(\beta-1) > \beta$. Ainsi, on peut théoriquement obtenir un développement de T_n jusqu'à la précision $1/n^\beta$, même lorsque $\beta < 2$: on a donc bien un raffinement du théorème 2.2.20, puisqu'il ne donnait un développement que jusqu'à la précision $1/n^{2\beta-2}$ lorsque $\beta < 2$. Notons quand même que, dans le cas $\beta > 2$, ce raffinement n'apporte rien, puisqu'on avait déjà une précision $1/n^\beta$.

Il reste ensuite à développer $\left[\frac{1}{\mu}P \cdot G(z)\right]^k \frac{1}{\mu}P$, pour obtenir l'expression des termes suivants dans le développement asymptotique de T_n. Le terme obtenu pour $k = 2$ est

$$\frac{1}{\mu^3}\Big(\sum_{k,l>n} P_k P_l - \sum_{\substack{k,l\leqslant n\\ k+l>n}} P_k P_l\Big)$$

où $P_n = \sum_{p=n+1}^{\infty} R_p$.
Pour $k = 3$, on trouve (après des calculs laborieux)

$$\sum_{k,l,m>n} - \sum_{\substack{0<k,l\leqslant n\\ k+l>n\\ m>n}} - \sum_{\substack{0<k,m\leqslant n\\ k+m>n\\ l>n}} - \sum_{\substack{0<l,m\leqslant n\\ l+m>n\\ k>n}} - \sum_{\substack{0<k,l,m\leqslant n\\ k+l>n\\ k+m>n}}$$
$$- \sum_{\substack{0<k,l,m\leqslant n\\ l+k>n\\ l+m>n}} - \sum_{\substack{0<k,l,m\leqslant n\\ m+k>n\\ m+l>n}} + \sum_{\substack{0<k,l,m\leqslant n\\ k+l>n\\ k+m>n\\ l+m>n}} + \sum_{\substack{k+l\leqslant n\\ k+m\leqslant n\\ l+m\leqslant n\\ k+l+m>n}} \frac{1}{\mu^4} P_k P_l P_m.$$

Corollaire 2.4.5. *Soit R_n une suite d'opérateurs de renouvellement, avec $\sum_{n+1}^{\infty} \|R_p\| = O(1/n^\beta)$ pour un certain $\beta > 1$. Il existe alors une constante $C > 0$ telle que, pour toute $f \in \mathcal{L}$ avec $Pf = 0$, pour tout $n \in \mathbb{N}^*$,*

$$\|T_n f\| \leqslant \frac{C}{n^\beta} \|f\|.$$

Démonstration. Soit N assez grand pour que le terme d'erreur dans la proposition 2.4.4 soit un $O(1/n^\beta)$. Il suffit alors de noter que tous les termes du développement asymptotique donnés par cette proposition se factorisent par P, et sont donc nuls quand on les applique à f. □

Corollaire 2.4.6. *Soit T l'application LSV de paramètre $\alpha \in]0,1[$. Alors, pour toutes fonctions f hölderienne d'intégrale nulle, et g bornée, qui s'annulent toutes deux au voisinage de 0, on a*

$$\int f \cdot g \circ T^n = O\left(\frac{1}{n^{1/\alpha}}\right).$$

En particulier, une fonction hölderienne d'intégrale nulle et nulle sur un voisinage de 0 a ses corrélations sommables, ce qui permet de montrer qu'elle satisfait un théorème central

limite, même dans le cas $1/2 < \alpha < 1$. C'est un résultat plus fin que la proposition 2.3.9, qui ne peut se démontrer par des majorations élémentaires de la vitesse de décorrélation comme la méthode de couplage de Young, puisque des phénomènes subtils de compensation sont à l'œuvre.

La preuve du corollaire repose de façon importante sur la forme de la suite $u_n = 1/n^\beta$, mais on va voir dans le paragraphe suivant que ce phénomène est en fait général, et valable pour toutes les suites convolutives.

2.4.2 Fonctions de moyenne nulle supportées par Y

Théorème 2.4.7. *Soit R_n une suite d'opérateurs de renouvellement. Alors*

$$\sum \|T_n(I-P)\| < \infty.$$

En particulier, pour toute fonction $f \in \mathcal{L}$ vérifiant $P(f) = 0$, on a $\sum_{n=0}^\infty \|T_n(f)\| < \infty$.

Démonstration. On va modifier légèrement la preuve du théorème 2.2.5. Soit $\widetilde{R}(z)$ la perturbation de $R(z)$ définie dans cette preuve, montrons tout d'abord que $(I - \widetilde{R}(z))^{-1}(I-P) \in A(\mathcal{A})$. On a

$$\begin{aligned}(I-\widetilde{R}(z))^{-1}(I-P) &= \frac{1}{1-z}\widetilde{S}(z)^{-1}(I-P)\\ &= \frac{1-z}{1-\widetilde{\lambda}(z)}\frac{\widetilde{P}(z)}{1-z}(I-P) + (I-\widetilde{R}(z)\widetilde{Q}(z))^{-1}\widetilde{Q}(z)(I-P).\end{aligned}$$

Comme $P(1)(I-P) = 0$ par définition, on obtient

$$(I-\widetilde{R}(z))^{-1}(I-P) = \frac{1-z}{1-\widetilde{\lambda}(z)}\frac{\widetilde{P}(z)-\widetilde{P}(1)}{1-z}(I-P) + (I-\widetilde{R}(z)\widetilde{Q}(z))^{-1}\widetilde{Q}(z)(I-P).$$

Mais $\frac{1-z}{1-\widetilde{\lambda}(z)} \in A(\mathbb{C})$ et $\frac{\widetilde{P}(z)-\widetilde{P}(1)}{1-z} \in A(\mathcal{A})$ (étape 3 de la preuve du théorème 2.2.5). De même, $(I-\widetilde{R}(z)\widetilde{Q}(z))^{-1} \in A(\mathcal{A})$ par le lemme de Wiener. En sommant, on obtient bien $(I-\widetilde{R}(z))^{-1}(I-P) \in A(\mathcal{A})$.

Comme dans l'étape 4 de la preuve du théorème 2.2.5, on utilise alors une partition de l'unité pour en déduire que $\sum T_n(I-P)z^n = (I-R(z))^{-1}(I-P) \in A(\mathcal{A})$, i.e. $\sum \|T_n(I-P)\| < \infty$. □

Remarque 2.4.8. Le même argument montre que $\sum \|(I-P)T_n\| < \infty$. Comme $T_n - PT_nP = T_n(I-P) + (I-P)T_nP$, on obtient

$$\sum \|T_n - PT_nP\| < \infty.$$

On en déduit le théorème suivant :

Théorème 2.4.9. *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante et préservant une mesure de probabilité m, de base Y. Alors, pour toute fonction $f : X \to \mathbb{R}$ de $\mathcal{L}$ supportée par Y et d'intégrale nulle, $\sum \|\widehat{T}^n(f)\|_1 < \infty$.*

Démonstration. Comme f est supportée dans Y, on a $B_b(f) = 0$ pour $b > 0$ et $C_n f = 0$ pour $n > 0$. L'équation (2.4) donne donc $\widehat{T}^n f = \sum_{a+k=n} A_a T_k(f)$. Ainsi, par le lemme 2.3.3,

$$\|\widehat{T}^n f\|_1 \leqslant \sum_{a+k=n} m[\varphi_Y > a] \, \|T_k(f)\|_{\mathcal{L}} \, .$$

En sommant, on obtient

$$\sum_{n=0}^{\infty} \|\widehat{T}^n f\|_1 \leqslant \left(\sum_{a=0}^{\infty} m[\varphi_Y > a] \right) \cdot \left(\sum_{k=0}^{\infty} \|T_k f\| \right) .$$

Les deux sommes de droite sont finies, d'après la formule de Kac pour la première et d'après le théorème 2.4.7 pour la seconde. Cela conclut la preuve. □

Par le théorème de Gordin-Liverani, cela implique que f satisfait un théorème central limite de variance $\sigma^2 = \int f + 2 \sum_1^\infty \int f \cdot f \circ T^n$. C'est très différent du cas où f est d'intégrale non nulle, puisqu'on a vu que la vitesse de décroissance des corrélations pouvait être arbitrairement lente. L'hypothèse de support est également très importante : soit f d'intégrale non nulle supportée par Y, dont les corrélations peuvent donc décroître lentement. Alors la fonction $g = f - \int f$ est hölderienne et d'intégrale nulle, mais ses corrélations, égales à celles de f, ne sont pas nécessairement sommables.

Les théorèmes de ce paragraphe ont été démontrés en travaillant dans l'algèbre de Banach des fonctions de coefficients de Fourier sommables. En travaillant dans une autre algèbre de Banach, par exemple $\mathcal{O}_u(\mathrm{Hom}(\mathcal{L}, \mathcal{L}))$ où u_n est une suite convolutive, on obtient les théorèmes suivants (avec exactement les mêmes preuves).

Théorème 2.4.10. *Soit R_n une suite d'opérateurs de renouvellement, avec $\sum_{n+1}^\infty \|R_p\| = O(u_n)$, où u_n est une suite convolutive. Alors, pour toute fonction $f \in \mathcal{L}$ vérifiant $P(f) = 0$, on a $\|T_n(f)\| = O(u_n)$.*

Remarque 2.4.11. Comme dans la remarque 2.4.8, on montre en fait que

$$\|T_n - P T_n P\| = O(u_n).$$

On en déduit le théorème suivant, qui généralise le corollaire 2.4.6 au cas d'une suite convolutive u_n générale.

Théorème 2.4.12. *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante et préservant une mesure de probabilité m, de base Y. On suppose que $m[\varphi_Y > n] = O(u_n)$, où u_n est une suite convolutive. Alors il existe une constante $C > 0$ telle que, pour toute fonction $f : X \to \mathbb{R}$ appartenant à $\mathcal{L}$, de moyenne nulle et supportée par Y, pour tout $n \in \mathbb{N}$, $\left\|\widehat{T}^n(f)\right\|_1 \leqslant C u_n \|f\|_{\mathcal{L}}$.*

2.4.3 Fonctions de moyenne nulle générales

Pour des fonctions qui ne sont pas supportées dans Y, on a le résultat suivant :

Théorème 2.4.13. *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante et préservant une mesure de probabilité m, de base Y. On suppose que $m[\varphi_Y > n] = O(u_n)$, où u_n est une suite convolutive. Soit $X_n \subset X$ l'ensemble des points qui entrent dans Y après n itérations, i.e. $X_n = X \setminus \bigcup_0^{n-1} T^{-i}(Y)$.*

Alors il existe une constante $C > 0$ telle que, pour toute fonction $f \in \mathcal{L}$ de moyenne nulle, pour tout $n \in \mathbb{N}$,

$$\left\|\widehat{T}^n(f)\right\|_1 \leqslant C u_n \|f\|_{\mathcal{L}} + C \sum_{k=0}^{n} u_k \left|\int_{X_{n-k}} f\right| + C \int_{X_{n+1}} |f|.$$

Dans le cas où f est de moyenne nulle et supportée par Y, les intégrales $\int_{X_l} f$ sont toutes nulles et on retrouve le théorème 2.4.12. Dans le cas où f est arbitraire, on peut juste dire que $\left|\int_{X_n} f\right| \leqslant (\sum_n^\infty u_p) \|f\|_\infty$, et on retrouve la majoration du théorème 2.3.6 grâce au lemme 2.2.17. C'est dans les situations intermédiaires, où la fonction f est "petite dans le haut de la tour", que ce théorème va donner des résultats plus précis.

Démonstration. Soit $Z_k = PT_kP$ et $Y_k = T_k - Z_k$. Ainsi, $\|Y_k\| = O(u_k)$ par la remarque 2.4.11, et on peut écrire $Z_k(u) = z_k \left(\int_Y u\right) 1_Y$ pour une certaine constante z_k.

Soit Λ_b l'ensemble des points qui rentrent dans Y après exactement b itérations, i.e. $\Lambda_b = X_b \backslash X_{b+1}$. Notons aussi $\lambda_b = \int B_b f = \int_{\Lambda_b} f$, et $s_b = \sum_b^\infty \lambda_i = \int_{X_b} f$. En particulier, $s_0 = \int f = 0$.

Soient $K \subset X$ et $\alpha_a = \int_K A_a(1_Y) \leqslant m[\varphi_Y > a]$. Alors, par (2.4), on a

$$\begin{aligned}\int_K \widehat{T}^n(f) &= \int_K C_n(f) + \sum_{a+k+b=n} \int_K A_a T_k B_b(f) \\ &= \int_K C_n(f) + \sum_{a+k+b=n} \alpha_a z_k \lambda_b + \int_K \sum_{a+k+b=n} A_a Y_k B_b(f).\end{aligned}$$

Le dernier terme est majoré par $C\|f\|_{\mathcal{L}} (u \star u \star u)_n$, par les lemmes 2.3.3 et 2.3.4 et l'inégalité $\|Y_k\| = O(u_k)$. Comme u_n est convolutive, il est donc majoré par $C u_n \|f\|_{\mathcal{L}}$. Le premier terme s'écrit $\int_K C_n(f) = \int_{T^{-n}K \cap X_{n+1}} f \leqslant \int_{X_{n+1}} |f|$.

Pour le terme restant, estimons tout d'abord $z \star \lambda$. On a

$$(z \star \lambda)_n = \sum_{b=0}^{n} \lambda_b z_{n-b} = \sum_{b=0}^{n} (s_b - s_{b+1}) z_{n-b} = \sum_{b=1}^{n} s_b (z_{n-b} - z_{n-b+1}) - s_{n+1} z_0.$$

On a $z_k = \frac{1}{m(Y)} \int_Y T_k(1_Y)$. Comme $\|T_k - T_{k+1}\| = O(u_k)$ par le théorème 2.2.19, $z_k - z_{k+1} = O(u_k)$. Ainsi, il existe C tel que $|(z \star \lambda)_n| \leqslant C(u \star |s|)_n + C|s_{n+1}|$. De plus,

$$|s_{n+1}| \leqslant |s_n| + \int_{\Lambda_{n+1}} |f| \leqslant |s_n| + \|f\|_{\mathcal{L}} \, m(\Lambda_{n+1}) \leqslant |s_n| + \|f\|_{\mathcal{L}} \, u_{n+1} \leqslant |s_n| + C \|f\|_{\mathcal{L}} \, u_n$$

car u_n est convolutive donc presque décroissante. Ainsi, en convolant par $\alpha_a = O(u_a)$, on obtient

$$\left| \sum_{a+k+b=n} \alpha_a z_k \lambda_b \right| \leqslant C(u \star u \star |s|)_n + C \|f\|_{\mathcal{L}} \, (u \star u)_n \leqslant C(u \star |s|)_n + C \|f\|_{\mathcal{L}} \, u_n$$

puisque u_n est convolutive.

On a donc obtenu la majoration désirée pour $\left| \int_K \widehat{T}^n(f) \right|$. Il suffit de l'appliquer à $K = \{\widehat{T}^n f(x) > 0\}$ puis à $K = \{\widehat{T}^n f(x) < 0\}$ pour majorer $\left\| \widehat{T}^n f \right\|_1$. □

Par exemple, dans le cas des applications LSV, on peut appliquer ce théorème aux fonctions nulles en 0 :

Théorème 2.4.14. *Soient T l'application LSV de paramètre $\alpha \in]0,1[$, et $\beta = 1/\alpha$. Soit f une fonction hölderienne de moyenne nulle avec $f(0) = 0$, satisfaisant $|f(x)| \leqslant Cx^\gamma$ pour un certain $\gamma > 0$. Alors*

$$\left\| \widehat{T}^n f \right\|_1 = O\left(\frac{1}{n^{\min(\beta, \beta(1+\gamma)-1)}} \right).$$

Ce théorème précise le corollaire 2.4.6. En particulier, si $\gamma > 2\alpha - 1$, alors cette estimée est sommable donc f satisfait un théorème central limite. On verra plus loin, par une méthode complètement différente (qui ne passe pas par la vitesse de décorrélation) qu'on a en fait un théorème central limite dès que $\gamma > \alpha - 1/2$ (théorème 4.2.5).

Démonstration. On travaille dans l'extension X' de X en forme de tour localement Gibbs-Markov, donnée par la proposition 1.3.13. Soit $\Lambda_n = T'^{-n}(Y') \setminus \bigcup_0^{n-1} T'^{-i}(Y')$. Cet ensemble se projette dans X sur l'intervalle $[x_{n+1}, x_n]$, sur lequel f est majorée par $Cx_n^\gamma \leqslant \frac{C}{n^{\gamma\beta}}$. Comme $m(\Lambda_n) = O(1/n^\beta)$, on obtient $\int_{\Lambda_n} |f| \leqslant \frac{C}{n^{(1+\gamma)\beta}}$, puis $\int_{X'_n} |f| \leqslant \frac{C}{n^{(1+\gamma)\beta-1}}$. Le théorème 2.4.13 (appliqué avec $u_n = 1/n^\beta$) donne alors la conclusion. □

Pour les applications de Holland, on trouve de la même façon que, dans tous les cas, les corrélations des fonctions lipschitziennes et qui s'annulent en 0 décroissent en $O(m[\varphi_Y > n])$, et sont donc toujours sommables. Ainsi, ces fonctions vérifient un théorème central limite.

Chapitre 3

Construction de structures markoviennes

Sommaire

3.1 Résultats 114
3.1.1 Décorrélations et dilatation asymptotique 114
3.1.2 L'application de Viana 116
3.1.3 Décorrélations et dilatation en temps fini 117
3.1.4 Méthode de preuve 119
3.2 Temps hyperboliques 119
3.2.1 Fréquence des temps hyperboliques 120
3.2.2 Constructions aux temps hyperboliques 121
3.3 Description de la partition 125
3.3.1 Description de la construction 127
3.3.2 Mesure des points interdits de nombreuses fois 128
3.3.3 Preuve du théorème 3.3.1 135
3.4 Tour de Young 136

Dans le chapitre 2, on a vu comment étudier la vitesse de décorrélation dans les tours localement Gibbs-Markov. Par conséquent, on sait estimer la vitesse de mélange des applications pour lesquelles on sait construire des extensions qui sont localement Gibbs-Markov, à condition d'avoir un bon contrôle sur la mesure des points qui mettent longtemps à revenir à la base de la tour.

Les exemples d'applications localement Gibbs-Markov qu'on a vus jusqu'à maintenant sont explicites, ou presque explicites, si bien qu'il est aisé de construire une partition de Markov. Dans ce chapitre, on va montrer comment on peut construire une telle structure markovienne pour des applications générales, en utilisant seulement quelques hypothèses abstraites de dilatation asymptotique.

Comme conséquence de ces techniques, on obtiendra en particulier une structure markovienne

pour les application d'Alves-Viana, et on en déduira que, pour ces applications, les corrélations des fonctions hölderiennes décroissent au moins en $e^{-c\sqrt{n}}$ (théorème 3.1.2).Les résultats de ce chapitre correspondent à l'article [Gou06].

3.1 Résultats

3.1.1 Décorrélations et dilatation asymptotique

Récemment, [ALP05] a montré comment étudier la vitesse de décroissance des corrélations de certaines applications non-uniformément dilatantes, en mélangeant les techniques de temps hyperboliques ([Alv00]) et les tours de Young ([You99]). Une conséquence de leur résultat est que les corrélations de l'application d'Alves-Viana ([Via97]) décroissent plus rapidement que tous les polynômes, ce qui permet d'obtenir un théorème central limite. Cependant, les estimées de [Via97] sont toutes en $O(e^{-c\sqrt{n}})$, ce qui est plus fort. L'objectif de ce chapitre est d'étendre les résultats de [ALP05] (en utilisant une méthode substantiellement différente) à des décroissances (entre autres) en $e^{-c\sqrt{n}}$. En fait, on va construire une tour localement Gibbs-Markov, à laquelle on pourra appliquer les résultats de décorrélation du chapitre 2.

Soient M une variété riemannienne compacte (éventuellement à bords) et $T : M \to M$ une application. On suppose qu'il existe une partie fermée de mesure de Lebesgue nulle $S \subset M$ (le lieu critique, contenant éventuellement des discontinuités ou des points critiques de T, avec $\partial M \subset S$), hors de laquelle T est un difféomorphisme local C^2, non-uniformément dilatant : il existe $\lambda > 0$ tel que, pour Lebesgue presque tout $x \in M$,

$$\liminf_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \log \left\| DT(T^k x)^{-1} \right\|^{-1} \geqslant \lambda. \tag{3.1}$$

On fait également des hypothèses de non-dégénérescence sur l'ensemble critique, analogues à celles de [ABV00] ou [ALP05] : on suppose qu'il existe $B > 1$ et $\beta > 0$ tels que, pour tout $x \in M \backslash S$ et pour tout $v \in T_x M \backslash \{0\}$,

$$\frac{1}{B} \operatorname{dist}(x, S)^{\beta} \leqslant \frac{\|DT(x)v\|}{\|v\|} \leqslant B \operatorname{dist}(x, S)^{-\beta}. \tag{3.2}$$

On suppose aussi que, pour tous $x, y \in M$ avec $\operatorname{dist}(x, y) < \operatorname{dist}(x, S)/2$,

$$\left| \log \left\| DT(x)^{-1} \right\| - \log \left\| DT(y)^{-1} \right\| \right| \leqslant B \frac{\operatorname{dist}(x, y)}{\operatorname{dist}(x, S)^{\beta}} \tag{3.3}$$

et

$$\left| \log |\det DT(x)^{-1}| - \log |\det DT(y)^{-1}| \right| \leqslant B \frac{\operatorname{dist}(x, y)}{\operatorname{dist}(x, S)^{\beta}}, \tag{3.4}$$

i.e. $\log \|DT^{-1}\|$ et $\log |\det DT^{-1}|$ sont localement lipschitziennes, la constante de Lipschitz étant contrôlée par la distance à l'ensemble critique. Cela signifie que les singularités sont au plus polynomiales, et exclut par exemple les points critiques plats.

On suppose que les points s'approchent de S à une vitesse sous-exponentielle dans le sens suivant. Pour $\delta > 0$, soit $\mathrm{dist}_\delta(x, S) = \mathrm{dist}(x, S)$ si $\mathrm{dist}(x, S) < \delta$, et $\mathrm{dist}_\delta(x, S) = 1$ sinon. On suppose alors que, pour tout $\varepsilon > 0$, il existe $\delta(\varepsilon) > 0$ tel que, pour Lebesgue-presque tout $x \in M$,

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} -\log \mathrm{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant \varepsilon. \tag{3.5}$$

On aura besoin de contrôler plus précisément les vitesses de convergence dans (3.1) et dans (3.5). Pour cela, on considère comme [ALP05] la fonction

$$h^1_{(\varepsilon_1,\varepsilon_2)}(x) = \inf\Bigg\{ N \in \mathbb{N}^* \mid \forall n \geqslant N, \frac{1}{n}\sum_{k=0}^{n-1} \log \left\| DT(T^k x)^{-1} \right\|^{-1} \geqslant \frac{\lambda}{2} \\ \text{et pour } i = 1, 2, \frac{1}{n}\sum_{k=0}^{n-1} -\log \mathrm{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i \Bigg\}.$$

Il est important d'avoir deux indices ε_1 et ε_2 pour garantir l'existence de temps hyperboliques (voir le lemme 3.2.2). Pour alléger les notations, on notera $\varepsilon = (\varepsilon_1, \varepsilon_2)$. Les points satisfaisant $h^1_\varepsilon(x) = n$ sont donc "bons" à partir du temps n. Ainsi, le défaut d'expansion asymptotique du système au temps n est mesuré par

$$\mathrm{Leb}\{x \mid h^1_\varepsilon(x) > n\}, \tag{3.6}$$

et il est donc naturel de chercher à estimer la vitesse de décroissance des corrélations en fonction de cette quantité. C'est ce que fait [ALP05], dans le cas polynomial : si (3.6) $= O(1/n^\gamma)$ pour un certain $\gamma > 1$, alors les corrélations des fonctions hölderiennes décroissent comme $1/n^{\gamma-1}$.

Notons $\Lambda = \bigcap_{n\geqslant 0} T^n(M)$. On dira que T est topologiquement transitif sur l'attracteur Λ si, pour tout couple d'ouverts non vides U, V de Λ, il existe n tel que $T^{-n}(U) \cap V$ contienne un ouvert non vide (il faut faire attention à la formulation à cause des discontinuités éventuelles de T sur S).

Rappelons la définition 2.2.11 : une suite $(u_n)_{n\in\mathbb{N}}$ de réels strictement positifs est à décroissance polynomiale s'il existe $C > 0$ tel que, pour tous $n/2 \leqslant k$, $0 < u_k \leqslant C u_n$. Cela implique en particulier que u_n ne décroît pas trop vite : il existe $\gamma > 0$ tel que $1/n^\gamma = O(u_n)$ (par exemple $\gamma = \frac{\log C}{\log 2}$).

Théorème 3.1.1. *Supposons que tous les itérés de T sont topologiquement transitifs sur Λ et que, pour tout $\varepsilon = (\varepsilon_1, \varepsilon_2)$, il existe une suite $u_n(\varepsilon)$ avec $\sum u_n(\varepsilon) < +\infty$ et $\mathrm{Leb}\{x \mid h^1_\varepsilon(x) > n\} = O(u_n(\varepsilon))$. On suppose que $u_n(\varepsilon)$ vérifie une des propriétés suivantes :*

1. *$u_n(\varepsilon)$ est à décroissance polynomiale.*
2. *Il existe $c(\varepsilon) > 0$ et $\eta(\varepsilon) \in]0,1]$ tels que $u_n(\varepsilon) = e^{-c(\varepsilon)n^{\eta(\varepsilon)}}$.*

Alors T préserve une unique mesure μ absolument continue par rapport à la mesure de Lebesgue. De plus, cette mesure est une mesure de probabilité, mélangeante, et son bassin contient Lebesgue-presque tout point de M.

Finalement, il existe ε^0 tel que, si $f, g : M \to \mathbb{R}$ sont deux fonctions avec f hölderienne et g bornée, leurs corrélations $\mathrm{Cor}(f, g \circ T^n) = \int f \cdot g \circ T^n \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu \int g \,\mathrm{d}\mu$ décroissent à la vitesse suivante :

1. *$|\mathrm{Cor}(f, g \circ T^n)| \leqslant C \sum_{p=n}^{\infty} u_p(\varepsilon^0)$ dans le cas 1.*
2. *Il existe $c' > 0$ tel que $|\mathrm{Cor}(f, g \circ T^n)| \leqslant C e^{-c' n^{\eta(\varepsilon^0)}}$ dans le cas 2.*

En fait, ε^0 peut être choisi *a priori*, ne dépendant que de λ et de T. Il suffirait alors d'avoir (3.5) pour ε_1^0 et ε_2^0 pour obtenir le théorème. Cependant, dans les cas concrets, il est souvent équivalent d'avoir (3.5) pour une valeur particulière de ε ou pour tout ε. C'est pourquoi, comme dans [ABV00] et [ALP05], nous avons préféré cette formulation du théorème, plus parlante.

Dans le premier cas, en prenant $u_n(\varepsilon) = 1/n^\gamma$, on retrouve le résultat de [ALP05].

Le principal problème de ce résultat est que (3.6) est souvent difficile à estimer dans la pratique puisque $h_\varepsilon^1(x)$ exprime une condition sur tous les itérés de x, et pas seulement un nombre fini.

3.1.2 L'application de Viana

Ces résultats s'appliquent à l'application de Viana, donnée par

$$T : \left\{ \begin{array}{ccc} S^1 \times I & \to & S^1 \times I \\ (\omega, x) & \mapsto & (16\omega, a_0 + \varepsilon \sin(2\pi\omega) - x^2) \end{array} \right. \tag{3.7}$$

où $a_0 \in]1,2[$ est un point de Misiurewicz (i.e. le point critique 0 est prépériodique pour l'application $x \mapsto a_0 - x^2$), ε est assez petit, et I est un segment de $]-2,2[$ choisi de telle sorte que T envoie $S^1 \times I$ dans son intérieur.

Cette application a été introduite par Viana dans [Via97]. Il montre que T (et en fait toute application assez proche de T) a presque partout deux exposants de Lyapunov strictement positifs, malgré la présence des points critiques dans les fibres. Plus précisément, Viana montre que les points qui ne voient pas la dilatation dans les fibres au temps n ont une mesure qui décroît en $O(e^{-c\sqrt{n}})$. [AA03] en déduit que, pour tout $\varepsilon = (\varepsilon_1, \varepsilon_2)$, pour tout $c < 1/4$,

$$\mathrm{Leb}\{x \mid h_\varepsilon^1(x) > n\} = O(e^{-c\sqrt{n}}). \tag{3.8}$$

De plus, [AV02] montre que tous les itérés de T sont topologiquement transitifs sur Λ.

Les résultats de [ALP05] leur permettent d'estimer la vitesse de décorrélation de l'application de Viana. Leur méthode de preuve, qui ne permet de contrôler que des vitesses polynomiales (voir le paragraphe 3.1.4) impliquait que les décorrélations étaient plus rapides que tous les polynômes, mais ils ne pouvaient atteindre la majoration conjecturale de $e^{-c'\sqrt{n}}$. Le théorème 3.1.1 permet d'atteindre cet objectif :

Théorème 3.1.2. *Les décorrélations des fonctions hölderiennes pour les applications assez proches (en topologie C^3) de l'application de Viana sont au moins aussi rapides que $e^{-c'\sqrt{n}}$ pour un certain $c' > 0$.*

Ce résultat s'applique encore si on remplace le coefficient de dilatation 16 par 2, d'après [BST03]. Notons que la méthode *ad hoc* de [BG03], qui démontre le théorème 3.1.2 pour les applications de Viana, ne s'applique pas directement quand on remplace 16 par 2, puisqu'elle utilise entre autres la forme des courbes admissibles. En revanche, la méthode abstraite du présent chapitre s'applique immédiatement, puisque [BST03] montre essentiellement (3.8).

3.1.3 Décorrélations et dilatation en temps fini

La fonction $h^1_\varepsilon(x)$ fait intervenir la dilatation de x à partir d'un certain rang, et est donc difficile à estimer en général. Il est plus naturel de considérer le premier instant où on a de la dilatation. Pour des raisons techniques, on aura besoin de trois paramètres pour obtenir des résultats dans ce contexte (voir la preuve du lemme 3.2.1). Soit

$$h^2_{(\varepsilon_1,\varepsilon_2,\varepsilon_3)}(x) = \inf\Bigg\{ n \in \mathbb{N}^* \mid \frac{1}{n}\sum_{k=0}^{n-1} \log \left\| DT(T^k x)^{-1} \right\|^{-1} \geqslant \frac{\lambda}{2} \\ \text{et pour } i = 1,2,3, \frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i \Bigg\}.$$

Cette définition fait intervenir uniquement les n premiers itérés de x, et est donc vérifiable en temps fini. On notera $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$. Il faut rapprocher h^2_ε de la notion de *premier temps hyperbolique*, étudiée par exemple dans [AA04].

Si on avait uniquement deux paramètres dans la définition de h^2, on aurait $h^2 \leqslant h^1$. Cependant, comme on a trois paramètres, h^1 et h^2 ne sont pas comparables en toute rigueur.

On va estimer la vitesse de décorrélation en fonction de la décroissance de $\mathrm{Leb}\{x \mid h^2_\varepsilon(x) > n\}$. On obtient le théorème suivant (complètement nouveau par rapport à [ALP05], et non accessible par leur technique) :

Théorème 3.1.3. *Supposons que tous les itérés de T sont topologiquement transitifs sur Λ et que, pour tout $\varepsilon = (\varepsilon_1, \varepsilon_2, \varepsilon_3)$, il existe une suite $u_n(\varepsilon)$ avec $\sum (\log n) u_n(\varepsilon) < +\infty$ et $\mathrm{Leb}\,\{x \mid h^2_\varepsilon(x) > n\} = O(u_n(\varepsilon))$. On suppose que $u_n(\varepsilon)$ vérifie une des propriétés suivantes :*

1. *$u_n(\varepsilon)$ est à décroissance polynomiale.*
2. *Il existe $c(\varepsilon) > 0$ et $\eta(\varepsilon) \in]0,1]$ tels que $u_n(\varepsilon) = e^{-c(\varepsilon)n^{\eta(\varepsilon)}}$.*

Alors T préserve une unique mesure μ absolument continue par rapport à la mesure de Lebesgue. De plus, cette mesure est une mesure de probabilité, mélangeante, et son bassin contient Lebesgue-presque tout point de M.

Finalement, il existe $\varepsilon^0 = (\varepsilon_1^0, \varepsilon_2^0, \varepsilon_3^0)$ tel que, si $f, g : M \to \mathbb{R}$ sont deux fonctions avec f hölderienne et g bornée, leurs corrélations $\operatorname{Cor}(f, g \circ T^n) = \int f \cdot g \circ T^n \,\mathrm{d}\mu - \int f \,\mathrm{d}\mu \int g \,\mathrm{d}\mu$ décroissent à la vitesse suivante :

1. *$|\operatorname{Cor}(f, g \circ T^n)| \leqslant C \sum_{p=n}^{\infty} (\log p) u_p(\varepsilon^0)$ dans le cas 1.*
2. *Il existe $c' > 0$ tel que $|\operatorname{Cor}(f, g \circ T^n)| \leqslant C e^{-c' n^{\eta(\varepsilon^0)}}$ dans le cas 2.*

Par exemple, lorsque $\operatorname{Leb}\{x \mid h_\varepsilon^2(x) > n\} = O(1/n^\gamma)$ pour $\gamma > 1$, on obtient une décroissance des corrélations en $O(\log n / n^{\gamma-1})$. Il faut noter la perte $\log n$ (peut-être due à la technique de preuve) entre le théorème 3.1.1 et le théorème 3.1.3.

Les commentaires sur le choix de ε^0 qui suivent le théorème 3.1.1 sont encore valides ici. On peut d'ailleurs prendre les mêmes valeurs pour ε_1^0 et ε_2^0 dans les deux théorèmes.

On reviendra plus loin sur l'existence de mesures invariantes dans le théorème 3.3.2. Sans hypothèse de transitivité, on obtiendra une décomposition spectrale : T admet un nombre fini de mesures invariantes ergodiques et chacune de ces mesures a un nombre fini de composantes qui sont mélangeantes pour un itéré de T et ont une décroissance des corrélations avec les mêmes bornes que dans les théorèmes 3.1.1 et 3.1.3 : ces deux théorèmes correspondent au cas où la décomposition spectrale est triviale.

Remarque 3.1.4. Si u_n est à décroissance polynomiale et $u_n = O(1/n^\gamma)$ pour un certain $\gamma > 1$, alors $\sum_{p=n}^{\infty} (\log p) u_p = O\left((\log n) \sum_{p=n}^{\infty} u_p\right)$, ce qui permet de simplifier l'expression de la borne sur la vitesse de décorrélation.

Démonstration. Pour $\alpha > 1$, on a

$$\sum_{p=n}^{\infty} (\log p) u_p \leqslant \sum_{p=n}^{n^\alpha} (\log p) u_p + \sum_{n^\alpha}^{\infty} \frac{\log p}{p^\gamma} \leqslant \sum_{p=n}^{n^\alpha} (\log n^\alpha) u_p + \frac{\log n}{n^{\alpha(\gamma-1)}}.$$

On choisit α assez grand pour que $\frac{\log n}{n^{\alpha(\gamma-1)}} = o(u_n)$, et on obtient la conclusion. □

Remarque 3.1.5. Dans le cas exponentiel gauche (i.e. $0 < \eta < 1$), on peut en fait obtenir la conclusion des théorèmes 3.1.1 et 3.1.3 pour tout $c' < c(\varepsilon^0)$. Cela se vérifie aisément dans toutes les preuves qui vont suivre.

3.1.4 Méthode de preuve

Dans la suite de ce chapitre, on travaillera sans hypothèse de transitivité topologique sur T, et on obtiendra un nombre fini de mesures invariantes absolument continues par rapport à la mesure de Lebesgue. De plus, les estimées sur la vitesse de décorrélation décrites dans les théorèmes 3.1.1 et 3.1.3 seront valables pour chacune de ces mesures pour un itéré de T.

La méthode de preuve va consister à construire une tour localement Gibbs-Markov : on va trouver une partition finie $U_1, \dots, U_N$ de M, et sous-partitionner chacun des U_i en un certain nombre d'ensembles W_j qui seront envoyés bijectivement sur l'un des U_j par un itéré T^{R_j} de T. D'après le chapitre 2, la vitesse de décorrélation sera alors reliée à $\mathrm{Leb}\left(\bigcup_{R_j>n} W_j\right)$.

Pour partitionner U_i, on va utiliser les *temps hyperboliques* (voir la partie 3.2). Le problème principal est que, comme on veut obtenir une partition, on devra prendre des ensembles deux à deux disjoints, ce qui impose des restrictions et empêche en particulier de prendre au temps n tous les points ayant n pour temps hyperbolique. Pour résoudre ce problème, [ALP05] utilise le lemme de Pliss, qui montre que les points ont beaucoup de temps hyperboliques, pour assurer que la plupart des points seront choisis assez rapidement. Cependant, le temps d'attente pour les points qui n'ont pas encore été traités peut être de l'ordre de n, ce qui empêche de montrer des vitesses de décorrélation plus rapides que $e^{-c(\log n)^2}$. On va développer des techniques combinatoires plus précises, qui ne reposent pas de façon essentielle sur le lemme de Pliss, pour obtenir des vitesses plus rapides.

3.2 Temps hyperboliques

On rappelle dans cette section la notion de temps hyperbolique de [Alv00] et [ABV00], et on décrit différents ensembles qu'on construit aux temps hyperboliques. Ces ensembles seront les briques de base pour construire la partition dans la partie 3.3.

Soit b une constante satisfaisant $0 < b < \min(1/2, 1/(4\beta))$. Pour $\sigma < 1$ et $\delta > 0$, on dit que $n \in \mathbb{N}^*$ est un temps (σ, δ)-hyperbolique pour x si, pour tout $1 \leqslant k \leqslant n$,

$$\prod_{j=n-k}^{n-1} \left\| DT(T^j x)^{-1} \right\| \leqslant \sigma^k \text{ et } \mathrm{dist}_\delta(T^{n-k}x, S) \geqslant \sigma^{bk}. \tag{3.9}$$

On notera $H_n = H_n(\sigma, \delta)$ l'ensemble des points ayant n pour temps (σ, δ) hyperbolique.

Dans le paragraphe 3.2.1, on va choisir judicieusement les constantes σ et δ (ainsi que ε^0, qui intervient dans les énoncés des théorèmes 3.1.1 et 3.1.3). Cependant, les raisons de ce choix ne deviendront apparentes que dans le paragraphe 3.3.3, et le lecteur peut donc avantageusement admettre l'existence de σ, δ et ε_0, et ne revenir au paragraphe 3.2.1 qu'au moment d'aborder le paragraphe 3.3.3.

3.2.1 Fréquence des temps hyperboliques

Le lemme qui suit est une légère généralisation de [ABV00, lemme 5.4] :

Lemme 3.2.1. *Soient $T : M \to M$ et $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ tels que (3.1) et (3.5) soient vérifiés.*
Il existe alors $\varepsilon_3 > 0$ et $\kappa > 0$ tels que, pour tous $\varepsilon_1, \varepsilon_2 < \kappa$, il existe $\theta(\varepsilon_1, \varepsilon_2) > 0$ tel que, si $x \in M$ et $n \in \mathbb{N}^*$ vérifient

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \left\|DT(T^k x)^{-1}\right\|^{-1} \geqslant \frac{\lambda}{2} \text{ et pour } i = 1,2,3, \frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\varepsilon_i,$$

alors il existe des instants $1 \leqslant p_1 < \ldots < p_l \leqslant n$ avec $l \geqslant \theta(\varepsilon_1, \varepsilon_2)n$ tels que, pour tout $j \leqslant l$,

$$\begin{aligned} \forall 1 \leqslant k \leqslant p_j, \sum_{s=p_j-k}^{p_j-1} \log \left\|DT(T^s x)^{-1}\right\|^{-1} &\geqslant \frac{\lambda}{4}k \\ \text{et pour } i = 1,2, \sum_{s=p_j-k}^{p_j-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^s x, S) &\leqslant 2\sqrt{\varepsilon_i}k. \end{aligned} \tag{3.10}$$

Autrement dit, on a une densité supérieure à $\theta(\varepsilon_1, \varepsilon_2)$ de temps p entre 1 et n vérifiant (3.10).
Avant de passer à la preuve, mentionnons un autre lemme analogue :

Lemme 3.2.2. *Soient $T : M \to M$ et $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ tels que (3.1) et (3.5) soient vérifiés.*
Soit $\kappa > 0$.
Il existe $\varepsilon_1, \varepsilon_2 < \kappa$ et $\theta > 0$ tel que, si $x \in M$ et $n \in \mathbb{N}^*$ vérifient

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \left\|DT(T^k x)^{-1}\right\|^{-1} \geqslant \frac{\lambda}{4} \text{ et pour } i = 1,2, \frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_i)}(T^k x, S) \leqslant 2\sqrt{\varepsilon_i},$$

alors il existe des instants $1 \leqslant p_1 < \ldots < p_l \leqslant n$ avec $l \geqslant \theta n$ tels que, pour tout $j \leqslant l$,

$$\forall 1 \leqslant k \leqslant p_j, \sum_{s=p_j-k}^{p_j-1} \log \left\|DT(T^s x)^{-1}\right\|^{-1} \geqslant \frac{\lambda}{8}k \text{ et } \sum_{s=p_j-k}^{p_j-1} -\log \operatorname{dist}_{\delta(\varepsilon_1)}(T^s x, S) \leqslant b\frac{\lambda}{8}k.$$

Dans toute la suite de cet article, on notera ε_3^0 pour la valeur de ε_3 donnée par le lemme 3.2.1, et $\varepsilon_1^0, \varepsilon_2^0$ pour les valeurs de ε_1 et ε_2 données par le lemme 3.2.2. On posera aussi $\sigma = e^{-\lambda/8} < 1$. Finalement, on notera $\delta = \delta(\varepsilon_1^0)$.

Ainsi, les temps p_j donnés par la conclusion du lemme 3.2.2 sont (σ, δ)-hyperboliques. De même, les temps p_j vérifiant (3.10) sont des temps (σ, δ)-hyperboliques (si κ a été choisi assez petit), mais ils sont plus que ça, puisqu'ils garantissent un contrôle à la fois pour ε_1^0 et pour

ε_2^0 (si bien qu'on pourra leur appliquer le lemme 3.2.2) : on dira qu'un temps vérifiant (3.10) pour ε_1^0 et ε_2^0 est un *super temps hyperbolique*. On notera SH_n l'ensemble des points ayant n pour super temps hyperbolique, et $H_n = H_n(\sigma, \delta)$ l'ensemble des points ayant n pour temps (σ, δ)-hyperbolique. En particulier, $SH_n \subset H_n$.

Dans la preuve qui va suivre, on verra pourquoi on a perte d'un indice ε_i à chaque fois. Cet indice est utilisé pour obtenir la conclusion sur $\sum_{s=p_j-k}^{p_j-1} \log \|DT(T^s x)^{-1}\|^{-1}$: comme la suite $\log \|DT(T^s x)^{-1}\|^{-1}$ n'est pas bornée, on ne peut pas lui appliquer directement le lemme de Pliss, et il faut donc utiliser un contrôle supplémentaire.

Démonstration du lemme 3.2.1. La preuve est essentiellement donnée par la preuve du lemme 5.4 de [ABV00] : ils montrent tout d'abord qu'il existe $\varepsilon_3 > 0$ (qu'on peut prendre arbitrairement petit) et $\theta_1 > 0$ tels que, si

$$\frac{1}{n}\sum_{k=0}^{n-1} \log \left\|DT(T^k x)^{-1}\right\|^{-1} \geqslant \frac{\lambda}{2} \text{ et } \frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon_3)}(T^k x, S) \leqslant 2\varepsilon_3,$$

alors il y a une proportion au moins $\theta_1 > 0$ de temps p entre 1 et n satisfaisant

$$\forall 1 \leqslant k \leqslant p, \sum_{s=p-k}^{p-1} \log \left\|DT(T^s x)^{-1}\right\|^{-1} \geqslant \frac{\lambda}{4} k.$$

De plus, [ABV00] montre également que, pour $\varepsilon > 0$, si x satisfait

$$\frac{1}{n}\sum_{k=0}^{n-1} -\log \operatorname{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant 2\varepsilon,$$

alors il existe au moins une proportion $\theta(\varepsilon) = 1 - \sqrt{\varepsilon}$ de temps p entre 1 et n vérifiant

$$\forall 1 \leqslant k \leqslant p, \sum_{s=p-k}^{p-1} -\log \operatorname{dist}_{\delta(\varepsilon)}(T^k x, S) \leqslant 2\sqrt{\varepsilon} k.$$

Quand $\varepsilon \to 0$, $\theta(\varepsilon) \to 1$. Ainsi, si κ est assez petit, pour tous $\varepsilon_1, \varepsilon_2 < \kappa$ on aura $\theta(\varepsilon_1, \varepsilon_2) := \theta_1 + \theta(\varepsilon_1) + \theta(\varepsilon_2) - 2 > 0$, ce qui donne la conclusion du lemme. □

La preuve du lemme 3.2.2 est similaire.

3.2.2 Constructions aux temps hyperboliques

Le lemme suivant précise [ABV00, lemme 5.2] et [ALP05, lemme 4.1] :

Lemme 3.2.3. *Il existe $\delta_2, D_1, \lambda_1 < 1$ tels que, pour tout $x \in M$ et tout entier n qui est un temps (σ, δ)-hyperbolique pour x, il existe un unique voisinage $V_n(x)$ de x avec les propriétés suivantes :*

1. *T^n est un difféomorphisme entre $V_n(x)$ et la boule $B(T^n x, \delta_2)$.*
2. *Pour $1 \leqslant k \leqslant n$ et $y, z \in V_n(x)$, $\operatorname{dist}(T^{n-k}y, T^{n-k}z) \leqslant \sigma^{k/2} \operatorname{dist}(T^n y, T^n z)$.*
3. *Pour tous $y, z \in V_n(x)$,*

$$\left| \frac{\det DT^n(y)}{\det DT^n(z)} - 1 \right| \leqslant D_1 \operatorname{dist}(T^n y, T^n z).$$

4. *$V_n(x) \subset B(x, \lambda_1^n)$.*
5. *Si $n \leqslant m$, $y \in H_m$ et $V_n(x) \cap V_m(y) \neq \emptyset$, alors T^n est injective sur $V_n(x) \cup V_m(y)$.*

Notons que le troisième point du lemme implique que T^n a une distorsion en volume bornée par $D_2 := 2\delta_2 D_1 + 1$, i.e. pour tous $U, V \subset V_n(x)$,

$$D_2^{-1} \frac{\operatorname{Leb}(T^n(U))}{\operatorname{Leb}(T^n(V))} \leqslant \frac{\operatorname{Leb}(U)}{\operatorname{Leb}(V)} \leqslant D_2 \frac{\operatorname{Leb}(T^n(U))}{\operatorname{Leb}(T^n(V))}. \tag{3.11}$$

Démonstration. Le lemme 5.2 de [ABV00] montre qu'il existe $\delta_1 > 0$ tel que, si $x \in H_n(\sigma, \delta)$, alors il existe un voisinage $V_n'(x)$ envoyé difféomorphiquement par T^n sur $B(T^n x, \delta_1)$. On pose alors $V_n(x) = V_n'(x) \cap T^{-n}(B(T^n x, \delta_1/4))$, et $\delta_2 = \delta_1/4$. Comme $V_n(x) \subset V_n'(x)$, les deux premières assertions du lemme découlent du lemme 5.2 de [ABV00] et la troisième du lemme 4.1 de [ALP05]. La quatrième est une conséquence de la seconde pour $\lambda_1 = \sigma^{1/2}$.

Pour l'unicité, notons que deux voisinages distincts $V_n^1(x)$ et $V_n^2(x)$ donneraient deux relèvements différents par T^n d'un chemin joignant $T^n(x)$ à un point fixé de $B(T^n x, \delta_1/4)$, ce qui est absurde.

Finalement, supposons que $V_n(x) \cap V_m(y)$ contienne un point z. On a $\operatorname{diam}(T^n(V_m(y))) \leqslant \operatorname{diam}(T^m(V_m(y))) = \delta_1/2$, si bien que $T^n(V_m(y)) \subset B(T^n x, \delta_1)$. On construit un ensemble $W_m(y) = T^{-n}(T^n(V_m(y))) \cap V_n'(x)$: par définition de $V_n'(x)$, T^n est un isomorphisme entre $W_m(y)$ et $T^n(V_m(y))$. Mais T^n est aussi un isomorphisme entre $V_m(y)$ et $T^n(V_m(y))$. Comme $V_m(y)$ et $W_m(y)$ contiennent le point z, l'argument d'unicité qui précède montre que $V_m(y) = W_m(y)$. En particulier, $V_m(y) \subset V_n'(x)$. Comme T^n est injective sur $V_n'(x)$, cela montre que T^n est injective sur $V_n(x) \cup V_m(y)$. □

Soit $\mathcal{U} = \{U_1, \dots, U_N\}$ une partition finie de M par des ensembles de diamètre au plus $\delta_2/10$, d'intérieur non vide et de frontière lisse par morceaux (par exemple une triangulation de M). Ainsi, il existe des constantes $C_2 > 0$ et $\lambda_2 < 1$ telles que

$$\forall 1 \leqslant i \leqslant N, \forall n \in \mathbb{N}, \operatorname{Leb}\{x \in U_i \mid \operatorname{dist}(x, \partial U_i) \leqslant \lambda_1^n\} \leqslant C_2 \lambda_2^n. \tag{3.12}$$

On notera $U_i' = \{x \mid \mathrm{dist}(x, U_i) \leqslant \delta_2/10\}$. Quitte à augmenter C_2 et λ_2, on peut aussi supposer que

$$\forall 1 \leqslant i \leqslant N, \forall n \in \mathbb{N}, \mathrm{Leb}\left\{x \in M \mid \mathrm{dist}(x, \partial U_i') \leqslant \frac{\delta_2}{2}\sigma^{n/2}\right\} \leqslant C_2\lambda_2^n \mathrm{Leb}(U_i). \tag{3.13}$$

On supposera finalement que, pour toute boule $B(x, \delta_2)$ de rayon δ_2 et pour tout $1 \leqslant i \leqslant N$,

$$\mathrm{Leb}\, B(x, \delta_2) \leqslant C_2 \mathrm{Leb}(U_i). \tag{3.14}$$

Soit $x \in H_n$. Alors $T^n x$ est dans un unique $U_i =: U(x, n)$, inclus dans $B(T^n x, \delta_2) = T^n(V_n(x))$. On notera $I_\infty^n(x) = T^{-n}(U_i) \cap V_n(x)$. Dans la construction de la partition de la partie 3.3, les éléments de partition seront des $I_\infty^n(x)$. Si, au cours de cette construction, on sélectionne $I_\infty^n(x)$ puis $I_\infty^{n+1}(y)$ alors que $y \notin I_\infty^n(x)$ et que y est très proche de la frontière de $I_\infty^n(x)$, les deux ensembles $I_\infty^n(x)$ et $I_\infty^{n+1}(y)$ risquent de ne pas être disjoints, ce qu'on veut éviter puisqu'on cherche à construire une partition. Comme dans [You98], on va donc devoir fixer un certain temps d'attente disant à partir de quel moment on aura le droit de sélectionner y, garantissant que $I_\infty^n(x) \cap I_\infty^m(y) = \emptyset$. On pose ainsi, pour $m > n$,

$$I_m^n(x) = \left\{y \in V_n(x) \mid \frac{\delta_2}{10}\sigma^{\frac{m-n}{2}} < \mathrm{dist}(T^n y, U(x, n)) \leqslant \frac{\delta_2}{10}\sigma^{\frac{m-n-1}{2}}\right\}$$

et $I_{\geqslant m}^n(x) = \bigcup_{m \leqslant t < \infty} I_t^n(x)$: ce sont les points qui n'ont pas le droit d'être choisis au temps m, pour ne pas interférer avec x au temps n (ce choix sera justifié par le lemme 3.2.5 et (3.15)). On dira qu'un point de $I_{\geqslant m}^n(x)$ est interdit par le temps n, pour le temps m. On écrira aussi $\widetilde{I}_{\geqslant m}^n(x) = \bigcup_{m \leqslant t \leqslant \infty} I_t^n(x)$, i.e. on rajoute le "centre" $I_\infty^n(x)$. La différence principale avec [You99] ou [ALP05] est que, dans ces articles, les estimées combinatoires requises sont moins précises, si bien qu'ils n'ont pas besoin de garder la trace de l'instant auquel un point est interdit (le n dans $I_{\geqslant m}^n$).

Lemme 3.2.4. *Si $0 < n \leqslant m$ et $\widetilde{I}_{\geqslant n+1}^n(x) \cap \widetilde{I}_{\geqslant m+1}^m(y) \neq \emptyset$, alors $\widetilde{I}_{\geqslant n+1}^n(x) \cup \widetilde{I}_{\geqslant m+1}^m(y) \subset V_n(x)$.*

Notons que, lorsqu'on écrit $\widetilde{I}_{\geqslant n+1}^n(x)$ (par exemple), il est implicite que cet ensemble est bien défini et donc que $x \in H_n$.

Démonstration. Soit $z \in \widetilde{I}_{\geqslant n+1}^n(x) \cap \widetilde{I}_{\geqslant m+1}^m(y)$. La dilatation du lemme 3.2.3 assure que $T^n(\widetilde{I}_{\geqslant m+1}^m(y)) \subset B(T^n z, \delta_2/2) \subset B(T^n x, \delta_2)$. En particulier, chaque point u de $T^n(\widetilde{I}_{\geqslant m+1}^m(y))$ a un antécédent u' par T^n dans $V_n(x)$. Il reste à voir que u' est dans $\widetilde{I}_{\geqslant m+1}^m(y)$. Sinon, u aurait un autre antécédent u'' dans $\widetilde{I}_{\geqslant m+1}^m(y)$. Comme $V_n(x) \cap V_m(y)$ contient z, la cinquième assertion du lemme 3.2.3 donne que T^n est injective sur $V_n(x) \cup V_m(y)$. C'est absurde, puisque $u' \neq u''$ mais $T^n(u') = T^n(u'')$. □

Lemme 3.2.5. *Il existe $P > 0$ tel que, pour tous $0 < n < m$, $x \in H_n$ et $y \in H_m \setminus \widetilde{I}^n_{\geqslant m}(x)$, on a*

$$\widetilde{I}^n_{\geqslant m+P}(x) \cap \widetilde{I}^m_{\geqslant m+P}(y) = \emptyset.$$

Autrement dit, si x n'interdit pas de choisir y au temps m, alors il n'y a plus d'interaction entre x et y après le temps $m+P$. Le délai P permet donc de disjoindre tout ce qui se passe entre les deux points (ce qui sera utilisé dans le lemme 3.3.6). En particulier, on en déduit

$$I^n_\infty(x) \cap I^m_\infty(y) = \emptyset, \tag{3.15}$$

ce qui permettra d'assurer que les ensembles que l'on sélectionnera dans la construction de la partition seront deux à deux disjoints.

Démonstration. Soit $U_i = T^n(I^n_\infty(x))$. Supposons que $z \in \widetilde{I}^n_{\geqslant m+P}(x) \cap \widetilde{I}^m_{\geqslant m+P}(y)$. Alors $\operatorname{dist}(T^n z, U_i) \leqslant \frac{\delta_2}{10}\sigma^{\frac{m+P-n-1}{2}}$ et $\operatorname{dist}(T^m z, T^m y) \leqslant \frac{\delta_2}{10}\left(1+\sigma^{\frac{P-1}{2}}\right)$. Notons aussi que, comme $y, z \in V_m(y)$, le lemme 3.2.3 assure que $\operatorname{dist}(T^n y, T^n z) \leqslant \sigma^{\frac{m-n}{2}} \operatorname{dist}(T^m y, T^m z)$. Ainsi,

$$\begin{aligned}\operatorname{dist}(T^n y, U_i) &\leqslant \operatorname{dist}(T^n y, T^n z) + \operatorname{dist}(T^n z, U_i) \leqslant \sigma^{\frac{m-n}{2}} \operatorname{dist}(T^m y, T^m z) + \operatorname{dist}(T^n z, U_i)\\ &\leqslant \sigma^{\frac{m-n}{2}} \frac{\delta_2}{10}\left(1+\sigma^{\frac{P-1}{2}}\right) + \frac{\delta_2}{10}\sigma^{\frac{m+P-n-1}{2}} = \frac{\delta_2}{10}\sigma^{\frac{m-n}{2}}\left(1+2\sigma^{\frac{P-1}{2}}\right).\end{aligned}$$

Si P est assez grand pour que $1+2\sigma^{\frac{P-1}{2}} \leqslant \sigma^{-1/2}$, on obtient donc $\operatorname{dist}(T^n y, U_i) \leqslant \frac{\delta_2}{10}\sigma^{\frac{m-n-1}{2}}$. Comme $y \in V_n(x)$ par le lemme 3.2.4, on obtient donc $y \in \widetilde{I}^n_{\geqslant m}(x)$. □

Lemme 3.2.6. *Il existe une suite $c_n > 0$ telle que, pour tout $n \in \mathbb{N}^*$, pour tout $x \in H_n$, $\operatorname{Leb} I^n_\infty(x) \geqslant c_n$.*

Démonstration. La condition $x \in H_n$ assure que, pour $k \leqslant n$, $\operatorname{dist}(T^k x, S) \geqslant \alpha_n > 0$. Comme T est C^1 sur $\{y \mid \operatorname{dist}(y, S) \geqslant \alpha_n\}$, on en déduit l'existence d'une constante C_n qui borne $\det DT^n(x)$ pour $x \in H_n$. Comme la distorsion est bornée par D_2 sur $V_n(x)$, on en déduit que, pour tout $y \in V_n(x)$, $|\det DT^n(y)| \leqslant D_2 C_n$. En particulier, $\operatorname{Leb} I^n_\infty(x) \geqslant \operatorname{Leb}(T^n(I^n_\infty(x)))/(D_2 C_n)$. Mais $T^n(I^n_\infty(x)))$ est l'un des U_i, et a donc une mesure uniformément minorée. □

Lemme 3.2.7. *Il existe une constante $C > 0$ telle que, pour tout A mesurable, pour tout $n \in \mathbb{N}^*$, $\operatorname{Leb}(H_n \cap T^{-n}(A)) \leqslant C \operatorname{Leb}(A)$.*

Démonstration. Les ensembles $I^n_\infty(x)$ pour $x \in H_n$ recouvrent H_n, sont deux à deux disjoints ou confondus, et en nombre fini par le lemme 3.2.6. On peut donc en prendre un système fini de représentants $I^n_\infty(x_1), \dots, I^n_\infty(x_k)$ (où k dépend de n).

Pour $1 \leqslant j \leqslant k$, la distorsion est bornée par D_2 sur $I_\infty^n(x_j)$, si bien que

$$\frac{\mathrm{Leb}(I_\infty^n(x_j) \cap T^{-n}A)}{\mathrm{Leb}(I_\infty^n(x_j))} \leqslant D_2 \frac{\mathrm{Leb}(A)}{\mathrm{Leb}(T^n(I_\infty^n(x_j)))}.$$

Mais $T^n(I_\infty^n(x_j))$ est l'un des U_i, de mesure minorée par une constante $c > 0$. En sommant sur j, et comme tous les $I_\infty^n(x_j)$ sont disjoints, on trouve

$$\mathrm{Leb}(H_n \cap T^{-n}(A)) \leqslant \frac{D_2}{c} \mathrm{Leb}(A) \mathrm{Leb}(M). \qquad \square$$

3.3 Description de la partition

Le résultat principal de ce chapitre est le suivant (sans hypothèse de transitivité sur T) :

Théorème 3.3.1. *Soient T une application sur une variété compacte M et $\delta : \mathbb{R}_+^* \to \mathbb{R}_+^*$ tels que (3.1) et (3.5) soient vérifiés. Soit ε^0 donné par les lemmes 3.2.1 et 3.2.2. On suppose que T vérifie une des conditions suivantes :*

1. $\mathrm{Leb}\{x \mid h_{\varepsilon^0}^1(x) > n\} = O(u_n)$ *où u_n est à décroissance polynomiale et tend vers 0.*
2. $\mathrm{Leb}\{x \mid h_{\varepsilon^0}^1(x) > n\} = O(u_n)$ *où $u_n = e^{-cn^\eta}$ avec $\eta \in]0,1]$.*
3. $\mathrm{Leb}\{x \mid h_{\varepsilon^0}^2(x) > n\} = O(u_n)$ *où u_n est à décroissance polynomiale et $(\log n)u_n \to 0$.*
4. $\mathrm{Leb}\{x \mid h_{\varepsilon^0}^2(x) > n\} = O(u_n)$ *où $u_n = e^{-cn^\eta}$ avec $\eta \in]0,1]$.*

Alors il existe une partition finie $U_1, \dots, U_N$ de M, une autre partition (modulo un ensemble de mesure de Lebesgue nulle) $W_1, W_2, \dots$ plus fine, et des temps $R_1, R_2, \dots$ tels que, pour tout j,

- *T^{R_j} est un difféomorphisme entre W_j et l'un des U_i.*
- *$T_{|W_j}^{R_j}$ dilate les distances d'au moins $\sigma^{-1/2} > 1$.*
- *La distorsion en volume de $T_{|W_j}^{R_j}$ est lipschitzienne, i.e. il existe une constante C indépendante de j telle que, pour tous $x, y \in W_j$,*

$$\left|1 - \frac{\det DT^{R_j}(x)}{\det DT^{R_j}(y)}\right| \leqslant C \,\mathrm{dist}(T^{R_j}x, T^{R_j}y).$$

- *Pour $x, y \in W_j$ et $n \leqslant R_j$, $\mathrm{dist}(T^n x, T^n y) \leqslant \mathrm{dist}(T^{R_j}x, T^{R_j}y)$.*

De plus, il existe $c' > 0$ tel qu'on ait sous les différentes hypothèses le contrôle suivant sur la taille des queues :

$$\mathrm{Leb}\left(\bigcup_{R_j > n} W_j\right) = \begin{cases} O(u_n) & \text{dans le premier cas,} \\ O((\log n)u_n) & \text{dans le troisième cas,} \\ O(e^{-c'n^\eta}) & \text{dans les deuxième et quatrième cas.} \end{cases}$$

Pour la preuve de ce théorème, il suffira en fait de travailler sur U_1 puisqu'on pourra ensuite refaire la même construction sur chacun des autres U_i, qui sont en nombre fini.

Notons que ce sont les estimées sur la taille des queues (qui tendent vers 0 par hypothèse) qui vont garantir que presque tout point est sélectionné, i.e. que les W_j forment bien une partition de M modulo un ensemble de mesure de Lebesgue nulle.

On en déduit le résultat suivant sur les mesures invariantes :

Théorème 3.3.2. *Sous les hypothèses du théorème 3.3.1, supposons en plus que $\sum u_n < \infty$ dans le premier cas, $\sum (\log n) u_n < \infty$ dans le troisième cas.*

Il existe alors un nombre fini non nul de mesures de probabilités $\mu_1, \dots, \mu_k$ invariantes absolument continues par rapport à la mesure de Lebesgue et ergodiques. De plus, leurs bassins recouvrent presque tout M. Finalement, il existe des ouverts deux à deux disjoints $O_1, \dots, O_k$ tels que μ_i soit équivalente à Leb *sur O_i et nulle sur $M \backslash O_i$.*

Soient μ l'une des mesures μ_i et O l'ouvert O_i correspondant. Il existe alors une partition (modulo 0) finie $\Omega_0, \dots, \Omega_{d-1}$ de O en un nombre fini d'ouverts, telle que $T(\Omega_i) = \Omega_{i+1}$ (modulo un ensemble de mesure de Lebesgue nulle) pour $0 \leqslant i \leqslant d-1$ (Ω_d étant identifié avec Ω_0). De plus, sur chacun des Ω_i, l'application T^d est mélangeante (et même exacte) pour la mesure μ.

Finalement, pour toutes fonctions $f, g : M \to \mathbb{R}$ avec f hölderienne et g bornée, il existe une constante C telle que, pour $0 \leqslant i \leqslant d-1$, pour tout $n \in \mathbb{N}$, les corrélations $\mathrm{Cor}_{\Omega_i}(f, g \circ T^{dn}) := \int_{\Omega_i} f \cdot g \circ T^{dn} \, \mathrm{d}\mu - \left(\int_{\Omega_i} f \, \mathrm{d}\mu \right) \left(\int_{\Omega_i} g \, \mathrm{d}\mu \right)$ vérifient

$$\left| \mathrm{Cor}_{\Omega_i}(f, g \circ T^{dn}) \right| \leqslant \begin{cases} C \sum_{p=dn}^{\infty} u_p & \text{dans le premier cas,} \\ C \sum_{p=dn}^{\infty} (\log p) u_p & \text{dans le troisième cas,} \\ C e^{-c' n^{\eta}} & \text{dans les deuxième et quatrième cas.} \end{cases} \tag{3.16}$$

En particulier, si T est topologiquement transitive sur Λ, il existe une unique mesure invariante absolument continue. Dans le cas où tous les itérés de T sont topologiquement transitifs, il y a un même un unique ensemble Ω. Ce théorème implique donc les théorèmes 3.1.1 et 3.1.3.

Démonstration. Le théorème 3.3.1 montre que T admet un modèle localement markovien $T' : X' \to X'$, au sens du paragraphe 1.3.4. On en déduit l'existence des mesures invariantes et leur décomposition en composantes mélangeantes par la proposition 1.3.18 : ces mesures sont obtenues en projetant sur M les mesures invariantes $\nu_1, \dots, \nu_s$ de T'.

Soit μ l'une des mesures μ_j. Par le théorème 1.3.18, elle est équivalente à Leb sur l'un des U_i. Construisons l'ouvert $O(\mu)$ de l'énoncé du théorème. Soit A_0 l'intérieur de U_i, non vide. Par récurrence, si A_n est défini et ouvert, on pose $A_{n+1} = T(A_n \backslash S) \cup A_n$. Comme S est fermé et comme T est un difféomorphisme local hors de S, on obtient que A_{n+1} est ouvert. Soit $O(\mu) = \bigcup A_n$. Comme μ est invariante, on vérifie par récurrence que μ est équivalente à Leb

sur A_n, et donc sur $O(\mu)$. Montrons maintenant que, si $B \subset M \backslash O(\mu)$, alors $\mu(B) = 0$. Sinon, par ergodicité, il existerait n tel que $\mu(T^{-n}(B) \cap A_0) > 0$. Comme $\mu(S) = 0$ car $\mathrm{Leb}(S) = 0$, on a donc $\mu(T^{-n}(B) \cap (A_0 \backslash S)) > 0$, puis $\mu(T^{-(n-1)}(B) \cap A_1) > 0$. Par récurrence, on obtient finalement $\mu(B \cap A_n) > 0$, ce qui est absurde.

Par la même construction, on montre que les composantes mélangeantes de μ, que l'on notera $\rho_0, \dots, \rho_{d-1}$, sont supportées par des ouverts, deux à deux disjoints.

Montrons finalement (3.16). Soit ν invariante ergodique sur X' telle que $\pi_*(\nu) = \mu$. Elle admet des composantes mélangeantes $\lambda_0, \dots, \lambda_{d_1-1}$, avec d_1 multiple de d (mettons $d_1 = sd$). De plus, $(X', (T')^{d_1}, \lambda_i)$ est une tour localement Gibbs-Markov mélangeante. Ainsi, lorsque f' est hölderienne sur X' et g' est bornée, les corrélations $\int f' \cdot g' \circ (T')^{d_1 n} \, \mathrm{d}\lambda_i - \left(\int f' \, \mathrm{d}\lambda_i\right) \cdot \left(\int g' \, \mathrm{d}\lambda_i\right)$ décroissent à la vitesse indiquée dans (3.16), par le théorème 2.3.6 (ou la remarque qui le suit, pour le cas exponentiel). Il reste à transférer ce résultat sur M.

Soit ρ_j fixée, et k tel que $\pi_*(\lambda_k) = \rho_j$. Soient f hölderienne sur M et g bornée. Notons $f' = f \circ \pi$ et $g' = g \circ \pi$: la fonction f' est hölderienne sur X' et g' y est bornée. Pour $n \in \mathbb{N}$, on écrit $n = ps + r$ avec $0 \leqslant r < s$. Alors

$$\int_{\Omega_j} f \cdot g \circ T^{dn} = \int_{X'} f' \cdot (g' \circ T'^{dr}) \circ T'^{pds} \, \mathrm{d}\lambda_k = \int_{X'} f' \cdot (g' \circ T'^{dr}) \circ T'^{pd_1} \, \mathrm{d}\lambda_k.$$

La fonction $g' \circ T'^{dr}$ est hölderienne sur X', donc le résultat de décroissance des corrélations pour λ_k sur X' donne la décroissance des corrélations de f et g sur M. □

Ce résultat est une décomposition spectrale pour T. L'existence d'un nombre fini de mesures invariantes était déjà connu sous des hypothèses plus faibles (voir dans [ABV00] la remarque suivant le corollaire D), mais les autres propriétés sont nouvelles.

La suite de cette partie sera consacrée à la démonstration du théorème 3.3.1.

3.3.1 Description de la construction

On va construire une partition de U_1 par des ensembles $W_1, W_2, \dots$ tels que, pour tout n, il existe un entier R_n tel que T^{R_n} est un isomorphisme entre W_n et l'un des U_i, avec dilatation de $\sigma^{-n/2}$ et distorsion en volume D_1-lipschitzienne. En fait, W_n sera un $I_\infty^{R_n}(x)$. On notera $H_n(U_1) = H_n \cap \{y \in U_1 \mid \mathrm{dist}(y, \partial U_1) \geqslant \lambda_1^n\}$. Ainsi, si $x \in H_n(U_1)$, on a $V_n(x) \subset U_1$ par le point 4 du lemme 3.2.3.

On construit en fait des points $x_1^1, \dots, x_{l(1)}^1$ au temps 1, puis $x_1^2, \dots, x_{l(2)}^2$ au temps 2, et ainsi de suite. Ils vérifieront les propriétés suivantes :

- $x_1^n, \dots, x_{l(n)}^n$ appartiennent tous à $H_n(U_1) \backslash \bigcup_{i<n, j \leqslant l(i)} \widetilde{I}_{\geqslant n}^i(x_j^i)$, et cet ensemble est recouvert par $\bigcup_j I_\infty^n(x_j^n)$.
- Les ensembles $I_\infty^n(x_j^n)$ (n et j variables) sont deux à deux disjoints, et inclus dans U_1.

On prendra alors pour W_i les ensembles $I^n_\infty(x^n_j)$, et le temps de retour R_i correspondant sera n.

Construction des x^n_i. La construction procède par récurrence sur n : au rang n, on note que, si $x, y \in H_n(U_1)$, alors $I^n_\infty(x)$ et $I^n_\infty(y)$ sont soit disjoints soit confondus. Il est donc possible de trouver un système de représentants $I^n_\infty(x^n_1), \dots, I^n_\infty(x^n_{l(n)})$ des $I^n_\infty(x)$ pour $x \in H_n(U_1) \setminus \bigcup_{i<n, j\leqslant l(i)} \widetilde{I}^i_{\geqslant n}(x^i_j)$, et ce système est fini d'après le lemme 3.2.6.

Par construction, deux $I^n_\infty(x^n_i)$ construits au même instant sont disjoints. Soient $m > n$, et $x^m_k \in H_m(U_1) \setminus \bigcup_{i<m, j\leqslant l(i)} \widetilde{I}^i_{\geqslant n}(x^i_j)$. Alors $x^m_k \in H_m \setminus \widetilde{I}^n_{\geqslant m}(x^n_i)$, si bien que le lemme 3.2.5 assure que $I^m_\infty(x^m_k)$ est disjoint de $I^n_\infty(x^n_i)$. Ainsi, les ensembles $I^\infty_n(x^n_j)$ (n et j variables) sont bien deux à deux disjoints.

Finalement, pour voir que $I^n_\infty(x^n_i) \subset U_1$, on utilise que $x^n_i \in H_n(U_1)$, donc $\mathrm{dist}(x^n_i, \partial U_1) \geqslant \lambda_1^n$. Comme $V_n(x^n_i) \subset B(x^n_i, \lambda_1^n)$, cela montre que $I^n_\infty(x^n_i) \subset U_1$. □

Les propriétés des temps hyperboliques énoncées dans le lemme 3.2.3 assurent alors que les propriétés de dilatation et de distorsion du théorème 3.3.1 sont vérifiées. Il ne reste plus qu'à estimer $\mathrm{Leb}\{x \mid \exists j \text{ tel que } x \in W_j \text{ et } R_j \geqslant n\}$.

3.3.2 Mesure des points interdits de nombreuses fois

On notera I_n l'ensemble des points interdits pour le temps n, i.e.

$$I_n = \bigcup_{i<n, j\leqslant l(i)} \widetilde{I}^i_{\geqslant n}(x^i_j),$$

et I^n l'ensemble des points interdits par le temps n, i.e.

$$I^n = \bigcup_{j\leqslant l(n)} \widetilde{I}^n_{\geqslant n+1}(x^n_j).$$

En particulier, $I^n \subset I_{n+1}$. Finalement, soit

$$S_n = \bigcup_{i\leqslant n, j\leqslant l(i)} I^i_\infty(x^i_j) \tag{3.17}$$

l'ensemble des points sélectionnés avant le temps n.

L'objectif de ce paragraphe est de démontrer le lemme 3.3.7 qui affirme que l'ensemble des points interdits pour k instants sans être sélectionnés a une mesure qui décroît exponentiellement vite. L'argument est combinatoire : si un point est interdit par peu d'instants, il sera alors interdit pour longtemps à ces instants, et on voit facilement que cela donne une mesure faible (lemme 3.3.6). Sinon, le point est interdit par beaucoup d'instants, et il faut voir que

chacun de ces instants permet de gagner un facteur multiplicatif $\lambda < 1$. On traitera pour cela encore deux cas : soit les ensembles interdits sont imbriqués les uns dans les autres, auquel cas on n'en garde qu'une proportion < 1 à chaque étape, et cela conclut (lemme 3.3.4), soit ils s'intersectent près de leurs frontières respectives, et il suffira de s'assurer que ces frontières sont assez petites (lemme 3.3.3).

On notera B pour une "boule interdite", c'est-à-dire un ensemble de la forme $\widetilde{I}^n_{\geqslant n+1}(x^n_i)$ où x^n_i est un des points définis dans la construction de la partie 3.3.1. On notera également $t(B)$ le temps n par lequel cette boule est interdite et $C(B) = I^n_\infty(x^n_i)$ son "centre", i.e. la partie interne correspondant aux points vraiment sélectionnés. Si $T^{t(B)}(C(B)) = U_i$, alors $T^{t(B)}(B) = \{x \mid \operatorname{dist}(x, U_i) \leqslant \frac{\delta_2}{10}\}$, donc $\operatorname{diam} T^{t(B)}(B) \leqslant \frac{3\delta_2}{10} \leqslant \frac{\delta_2}{2}$. Dans tous les énoncés et preuves de cette partie, les ensembles notés B_i ou B'_i seront implicitement de telles boules interdites. On va définir dans les lemmes qui suivent des ensembles $Z^1, \dots, Z^6$ de points "interdits de nombreuses fois", et montrer qu'ils ont chacun une mesure exponentiellement petite.

Lemme 3.3.3. *Soit $Q \in \mathbb{N}^*$. Soit*

$$Z^1(k, B_0) = \Bigg\{ x \mid \exists B'_1, B_1, \dots, B'_r, B_r \text{ avec } \forall 1 \leqslant i \leqslant r, t(B_{i-1}) \leqslant t(B'_i) \leqslant t(B_i) - Q, \\ B_i \not\subset B'_i, \sum_{i=1}^{r} \left\lfloor \frac{t(B_i) - t(B'_i)}{Q} \right\rfloor \geqslant k, \text{ et } x \in \bigcap_{i=0}^{r} B_i \cap \bigcap_{i=1}^{r} B'_i \Bigg\}.$$

Alors il existe une constante C_3 indépendante de Q telle que, pour tous k et B_0,

$$\operatorname{Leb}(Z^1(k, B_0)) \leqslant C_3 (C_3 \lambda_2^Q)^k \operatorname{Leb}(C(B_0)).$$

Rappelons que λ_2 est une constante qui vérifie (3.12) et (3.13).

Démonstration. Soit C_3 une constante telle que $\operatorname{Leb}\{x \mid \operatorname{dist}(x, U_i) \leqslant \frac{\delta_2}{10}\} \leqslant \frac{C_3}{D_2} \operatorname{Leb}(U_i)$ pour $1 \leqslant i \leqslant N$ et telle que $\frac{C_3^{-1}}{1 - C_3^{-1}} (C_2 D_2)^2 \leqslant 1$. On va montrer par récurrence sur k que C_3 convient.

Soit d'abord $k = 0$. Soient $n = t(B_0)$ et i tel que $T^n(C(B_0)) = U_i$. Alors $Z^1(0, B_0) = B_0$, donc $T^n(Z^1(0, B_0)) = \{x \mid \operatorname{dist}(x, U_i) \leqslant \frac{\delta_2}{10}\}$. Ainsi, $\operatorname{Leb}(T^n(Z^1(0, B_0))) \leqslant \frac{C_3}{D_2} \operatorname{Leb}(T^n(C(B_0)))$. Comme la distorsion de T^n est bornée par D_2 d'après (3.11), on trouve bien $\operatorname{Leb}(Z^1(0, B_0)) \leqslant C_3 \operatorname{Leb}(C(B_0))$.

Soit maintenant $k \geqslant 1$. Alors, en décomposant suivant la valeur de B'_1, on trouve

$$Z^1(k, B_0) \subset \bigcup_{t=1}^{k} \bigcup_{B'_1 \cap B_0 \neq \emptyset} \bigcup_{\substack{B_1 \cap B'_1 \neq \emptyset, B_1 \not\subset B'_1 \\ \left\lfloor \frac{t(B_1) - t(B'_1)}{Q} \right\rfloor \geqslant t}} Z^1(k - t, B_1).$$

Montrons que, si $t(B_1) - t(B_1') = n$, alors B_1 est inclus dans une couronne de taille $\sigma^{n/2}$ autour de B_1'. Plus précisément, soient $p = t(B_1')$, $U_i' = T^p(B_1')$, et montrons que

$$T^p(B_1) \subset \left\{ y \mid \operatorname{dist}(y, \partial U_i') \leqslant \frac{\delta_2}{2} \sigma^{n/2} \right\}. \tag{3.18}$$

Notons que B_1 contient un point de $\partial B_1'$ puisqu'il est connexe et rencontre B_1' et son complémentaire. Ainsi, $T^p(B_1)$ contient un point de $\partial U_i'$. De plus,

$$\operatorname{diam} T^p(B_1) \leqslant \sigma^{n/2} \operatorname{diam} T^{n+p}(B_1) \leqslant \sigma^{n/2} \frac{\delta_2}{2}.$$

Cela montre (3.18). Notons que (3.13) fournit une majoration de la mesure de (3.18).

En utilisant que la distorsion est bornée par D_2 pour les temps hyperboliques et que les centres $C(B_1)$ sont deux à deux disjoints par construction, on obtient par (3.18) et (3.13) que

$$\sum_{\substack{B_1 \cap B_1' \neq \emptyset, B_1 \not\subset B_1' \\ \left\lfloor \frac{t(B_1) - t(B_1')}{Q} \right\rfloor \geqslant t}} \operatorname{Leb}(C(B_1)) \leqslant C_2 \lambda_2^{Qt} D_2 \operatorname{Leb}(C(B_1')). \tag{3.19}$$

Finalement, si $q = t(B_0)$, les $T^q(C(B_1'))$ sont deux à deux disjoints et inclus dans un $T^q(V_q(x)) = B(T^q x, \delta_2)$ par le lemme 3.2.4. Ainsi, d'après (3.14) et la distorsion bornée,

$$\sum_{B_1' \cap B_0 \neq \emptyset} \operatorname{Leb}(C(B_1')) \leqslant C_2 D_2 \operatorname{Leb}(C(B_0)). \tag{3.20}$$

On obtient donc, grâce à l'hypothèse de récurrence,

$$\begin{aligned}
\operatorname{Leb} Z^1(k, B_0) &\leqslant \sum_{t=1}^{k} \sum_{B_1' \cap B_0 \neq \emptyset} \sum_{\substack{B_1 \cap B_1' \neq \emptyset, B_1 \not\subset B_1' \\ \left\lfloor \frac{t(B_1) - t(B_1')}{Q} \right\rfloor \geqslant t}} \operatorname{Leb} Z^1(k - t, B_1) \\
&\leqslant \sum_{t=1}^{k} \sum_{B_1' \cap B_0 \neq \emptyset} \sum_{\substack{B_1 \cap B_1' \neq \emptyset, B_1 \not\subset B_1' \\ \left\lfloor \frac{t(B_1) - t(B_1')}{Q} \right\rfloor \geqslant t}} C_3 (C_3 \lambda_2^Q)^{k-t} \operatorname{Leb}(C(B_1)) \\
&\leqslant \sum_{t=1}^{k} \sum_{B_1' \cap B_0 \neq \emptyset} C_3 (C_3 \lambda_2^Q)^{k-t} C_2 \lambda_2^{Qt} D_2 \operatorname{Leb}(C(B_1')) \\
&\leqslant C_3 \lambda_2^{Qk} C_3^k (C_2 D_2)^2 \left(\sum_{t=1}^{k} C_3^{-t} \right) \operatorname{Leb}(C(B_0)).
\end{aligned}$$

Par définition de C_3, on a $(C_2 D_2)^2 \left(\sum_{t=1}^{k} C_3^{-t} \right) \leqslant 1$. Cela conclut la récurrence. □

Lemme 3.3.4. *Soit*

$$Z^2_{k,N} = \{x \mid \exists B_1 \supsetneq B_2 \cdots \supsetneq B_k \text{ avec } t(B_k) \leqslant N \text{ et } x \in (B_1 \cap \cdots \cap B_k)\backslash S_N\}.$$

Il existe alors une constante $\lambda_3 < 1$ *telle que* $\mathrm{Leb}(Z^2_{k,N}) \leqslant \lambda_3^k \mathrm{Leb}(M)$.

Démonstration. On fixe N une fois pour toutes dans cette preuve, si bien qu'on omettra de préciser la dépendance en N. On va montrer que $\lambda_3 = \frac{C_2D_2}{C_2D_2+1}$ convient. Notons que, pour tout ensemble B,

$$\mathrm{Leb}(B) \leqslant C_2D_2 \,\mathrm{Leb}(C(B)) \tag{3.21}$$

par (3.14) et la distorsion bornée des temps hyperboliques.

On notera $\mathcal{B}_1$ l'ensemble des B avec $t(B) \leqslant N$ qui ne sont inclus dans aucun B', puis $\mathcal{B}_2$ l'ensemble des boules $B \notin \mathcal{B}_1$ avec $t(B) \leqslant N$ qui ne sont incluses que dans des boules de $\mathcal{B}_1$, et ainsi de suite. On dira qu'un élément de $\mathcal{B}_i$ est une boule de rang i. Toutes les boules B sont de rang fini puisqu'une boule construite au temps n est au maximum de rang n.

On notera $S'_k = \bigcup_{i=1}^k \bigcup_{B\in\mathcal{B}_i} C(B)$: ce sont les points sélectionnés dans des boules de rang au plus k.

Soit

$$Z^3_k = \left(\bigcup_{B\in\mathcal{B}_k} B\right)\backslash S'_k.$$

Montrons que $Z^2_k \subset Z^3_k$.

Soit $x \in Z^2_k$, il est donc dans un ensemble $(B_1 \cap \cdots \cap B_k)\backslash S_N$ avec $B_1 \supsetneq B_2 \cdots \supsetneq B_k$ et $t(B_k) \leqslant N$. En particulier, B_k est de rang $r \geqslant k$. Soit $B'_1 \supsetneq B'_2 \supsetneq \cdots \supsetneq B'_{r-1} \supsetneq B'_r$ une suite avec $B'_i \in \mathcal{B}_i$ et $B'_r = B_k$. En particulier, $x \in B'_k$. De plus, $S'_k \subset S_N$. Comme $x \notin S_N$, on en déduit $x \notin S'_k$. Cela montre donc que $x \in Z^3_k$.

Estimons $\mathrm{Leb}(Z^3_{k+1})$ en fonction de $\mathrm{Leb}(Z^3_k)$. Un point x de Z^3_{k+1} provient d'une boule B_{k+1} de rang $k+1$. Ainsi,

$$Z^3_{k+1} \subset \bigcup_{B_{k+1}\in\mathcal{B}_{k+1}} B_{k+1}.$$

Soit B_k une boule de rang k contenant B_{k+1}. Comme les centres des différentes boules sont deux à deux disjoints, on a $C(B_{k+1}) \cap S'_k = \emptyset$. Ainsi, $C(B_{k+1}) \subset B_k\backslash S'_k \subset Z^3_k$. Cependant, $C(B_{k+1}) \subset S'_{k+1}$ par définition, si bien que $C(B_{k+1}) \cap Z^3_{k+1} = \emptyset$. Ainsi, $C(B_{k+1}) \subset Z^3_k\backslash Z^3_{k+1}$.

Par conséquent, on obtient en utilisant (3.21)

$$\mathrm{Leb}(Z^3_{k+1}) \leqslant \sum_{B_{k+1}\in\mathcal{B}_{k+1}} \mathrm{Leb}(B_{k+1}) \leqslant C_2D_2 \sum_{B_{k+1}\in\mathcal{B}_{k+1}} \mathrm{Leb}\, C(B_{k+1}) \leqslant C_2D_2 \,\mathrm{Leb}(Z^3_k\backslash Z^3_{k+1})$$

puisque les $C(B_{k+1})$ sont deux à deux disjoints. Ainsi,

$$(C_2D_2+1)\,\mathrm{Leb}(Z^3_{k+1}) \leqslant C_2D_2\,\mathrm{Leb}(Z^3_{k+1}) + C_2D_2\,\mathrm{Leb}(Z^3_k\backslash Z^3_{k+1}) = C_2D_2\,\mathrm{Leb}(Z^3_k).$$

On en déduit par récurrence que $\mathrm{Leb}(Z_k^3) \leqslant \left(\frac{C_2D_2}{C_2D_2+1}\right)^k \mathrm{Leb}(M)$, ce qui donne la même inégalité pour $\mathrm{Leb}(Z_k^2)$ puisque $Z_k^2 \subset Z_k^3$. □

Lemme 3.3.5. *Soit*

$$Z^4(k,N) = \{x \mid \exists t_1 < \ldots < t_k \leqslant N, x \in I^{t_1} \cap \cdots \cap I^{t_k}\} \backslash S_N.$$

Il existe des constantes $C_4 > 0$ et $\lambda_4 < 1$ telles que, pour tous $1 \leqslant k \leqslant N$, $\mathrm{Leb}(Z^4(k,N)) \leqslant C_4\lambda_4^k$.

Autrement dit, les points interdits par au moins k instants ont une mesure exponentiellement petite.

Démonstration. Soit Q assez grand pour que $C_3\lambda_2^Q < 1$ dans le lemme 3.3.3. Écrivons $N = rQ + s$ avec $s < Q$.

Soit $x \in Z^4(k,N)$, interdit par les temps $t_1 < \ldots < t_k$. Pour $0 \leqslant a < r$, on choisit dans chaque intervalle $[aQ,(a+1)Q[$ le premier temps t_i s'il en existe un, ce qui donne une suite $t'_1 < \ldots < t'_{k'}$, avec $Qk' + s \geqslant k$. On garde ensuite les temps d'indices impairs, d'où une suite d'instants $u_1 < \ldots < u_l$ avec $2l \geqslant k'$, d'où $l \geqslant k/(2Q) - s/2$. De plus, $u_{i+1} - u_i \geqslant Q$ pour tout i. Soient $B_1, \ldots, B_l$ des boules construites aux temps u_i et qui interdisent x.

Soient $I = \{1 \leqslant i \leqslant l, B_i \subset B_1 \cap \ldots \cap B_{i-1}\}$ et $J = [1,l]\backslash I$. Si $\mathrm{Card}\, I \geqslant l/2$, on ne garde que les boules correspondant aux indices dans I. Il y en a au moins $l/2$, donc $x \in Z^2_{l/2,N}$ (où Z^2 est défini dans le lemme 3.3.4). Ce lemme assure que les points obtenus de cette manière ont une mesure exponentiellement petite (en l, et donc en k).

Sinon, $\mathrm{Card}\, J \geqslant l/2$. Soient $j_0 = \sup J$, puis $i_0 = \inf\{i < j_0, B_{j_0} \not\subset B_i\}$. Soient $j_1 = \sup\{j \leqslant i_0, j \in J\}$, puis $i_1 = \inf\{i < j_1, B_{j_1} \not\subset B_i\}$, et ainsi de suite, la construction se terminant à un rang i_n. Alors $J \subset \bigcup]i_s, j_s]$ par construction, donc $\sum(j_s - i_s) \geqslant \mathrm{Card}\, J \geqslant l/2$. Cela implique que $\sum \left\lfloor \frac{t(B_{j_s}) - t(B_{i_s})}{Q} \right\rfloor = \sum \left\lfloor \frac{u_{j_s} - u_{i_s}}{Q} \right\rfloor \geqslant l/2$ puisque deux instants u_j et u_i sont séparés d'au moins $Q(j-i)$ par construction. La suite $B_{i_n}, B_{i_n}, B_{j_n}, \ldots, B_{i_0}, B_{j_0}$ montre donc que $x \in Z^1(l/2, B_{i_n})$. En sommant les estimées données par le lemme 3.3.3 sur toutes les boules B_{i_n} possibles, qui ont des centres deux à deux disjoints, on trouve finalement une mesure exponentiellement petite en l et donc en k. □

Lemme 3.3.6. *Pour une boule $B_1 = \widetilde{I}^{t_1}_{\geqslant t_1+1}(x_1)$, soit*

$$Z^5(n_1,\ldots,n_k,B_1) = \{x \mid \exists t_2,\ldots,t_k \text{ avec } t_1 < \ldots < t_k \text{ et } x_2,\ldots,x_k \text{ tels que } \forall 1 \leqslant i \leqslant k, x \in I^{t_i}_{\geqslant t_i+n_i}(x_i)\}.$$

Il existe une constante C_5 indépendante de $B_1, n_1, \ldots, n_k$ telle que, dès que $n_1, \ldots, n_k > P$ donné par le lemme 3.2.5,

$$\mathrm{Leb}(Z^5(n_1,\ldots,n_k,B_1)) \leqslant C_5(C_5\lambda_2^{n_1}) \cdots (C_5\lambda_2^{n_k})\, \mathrm{Leb}\, C(B_1).$$

En fait, $Z^5(n_1,\dots,n_k,B_1)$ désigne l'ensemble des points qui sont interdits pendant un temps au moins n_1 par B_1, puis pendant un temps au moins n_2 par une autre boule B_2, et ainsi de suite.

Démonstration. On raisonne par récurrence sur k.

Soit $x \in Z^5(n_1,\dots,n_k,B_1)$. Il existe par définition une boule $B_2 = \widetilde{I}^{t_2}_{\geqslant t_2+1}(x_2)$ construite à un temps $t_2 > t_1$ telle que $x \in Z^5(n_2,\dots,n_k,B_2)$. Comme x_2 n'est pas interdit pour le temps t_2 (puisqu'il est sélectionné au cours de la construction du paragraphe 3.3.1), le lemme 3.2.5 assure que $\widetilde{I}^{t_1}_{\geqslant t_2+P}(x_1) \cap \widetilde{I}^{t_2}_{\geqslant t_2+P}(x_2) = \emptyset$. Mais x est interdit par le temps t_2 pour une durée au moins $n_2 > P$, donc $x \in \widetilde{I}^{t_2}_{\geqslant t_2+P}(x_2)$, d'où $x \notin \widetilde{I}^{t_1}_{\geqslant t_2+P}(x_1)$. Comme $x \in \widetilde{I}^{t_1}_{\geqslant t_1+n_1}(x_1)$, on obtient $t_1+n_1 < t_2+P$, i.e. $t_2-t_1 > n_1-P$.

Soit $U_i = T^{t_1}(C(B_1))$. La dilatation des temps hyperboliques donne que

$$\operatorname{diam}(T^{t_1}(B_2)) \leqslant \sigma^{\frac{t_2-t_1}{2}} \operatorname{diam}(T^{t_2}(B_2)) \leqslant \sigma^{\frac{n_1-P}{2}} \frac{\delta_2}{2}.$$

De plus, $\operatorname{dist}(T^{t_1}(x), \partial U_i) \leqslant \frac{\delta_2}{10}\sigma^{\frac{n_1-1}{2}}$ puisque x est interdit pour une durée au moins n_1. On a donc prouvé qu'il existe une constante C_6 telle que

$$T^{t_1}(B_2) \subset \mathcal{C} := \left\{ y \mid \operatorname{dist}(y, \partial U_i) \leqslant C_6 \sigma^{\frac{n_1}{2}} \right\}.$$

Par hypothèse de récurrence, $\operatorname{Leb}(Z^5(n_2,\dots,n_k,B_2)) \leqslant C_5(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \operatorname{Leb} C(B_2)$. On en déduit que $\operatorname{Leb}(T^{t_1}(Z^5(n_2,\dots,n_k,B_2))) \leqslant D_2C_5(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \operatorname{Leb}(T^{t_1}(C(B_2)))$ car la distorsion est bornée. On somme sur toutes les boules B_2, en notant que tous les $T^{t_1}(C(B_2))$ sont inclus dans la couronne $\mathcal{C}$ et deux à deux disjoints. On obtient donc

$$\begin{aligned}
\operatorname{Leb}(T^{t_1}(Z^5(n_1,\dots,n_k,B_1))) &\leqslant \sum_{B_2} \operatorname{Leb}(T^{t_1}(Z^5(n_2,\dots,n_k,B_2))) \\
&\leqslant C_5D_2(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \sum_{B_2} \operatorname{Leb}(T^{t_1}(C(B_2))) \\
&\leqslant C_5D_2(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \operatorname{Leb}(\mathcal{C}).
\end{aligned}$$

Par (3.13), il existe une constante C_7 telle que $\operatorname{Leb}(\mathcal{C}) \leqslant C_7\lambda_2^{n_1} \operatorname{Leb}(U_i)$. Ainsi,

$$\operatorname{Leb}(T^{t_1}(Z^5(n_1,\dots,n_k,B_1))) \leqslant C_5C_7D_2\lambda_2^{n_1}(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \operatorname{Leb}(U_i).$$

Comme $U_i = T^{t_1}(C(B_1))$, la borne de D_2 sur la distorsion donne finalement

$$\operatorname{Leb}(Z^5(n_1,\dots,n_k,B_1)) \leqslant C_5C_7D_2^2\lambda_2^{n_1}(C_5\lambda_2^{n_2})\cdots(C_5\lambda_2^{n_k}) \operatorname{Leb}(C(B_1)).$$

Cela conclut la preuve si on prend $C_5 \geqslant C_7D_2^2$ assez grand pour que le résultat soit vrai au rang $k=0$. □

Le lemme qui suit est l'aboutissement de ce qui précède : il montre que les points interdits pour k instants ont une mesure exponentiellement petite en k.

Lemme 3.3.7. *Soit*

$$Z^6(k,N) = \{x \mid \exists t_1 < \ldots < t_k \leqslant N, x \in I_{t_1} \cap \ldots \cap I_{t_k}\} \backslash S_N.$$

Il existe des constantes $C_8 > 0$ et $\lambda_5 < 1$ telles que, pour tous $k \leqslant N$,

$$\mathrm{Leb}(Z^6(k,N)) \leqslant C_8 \lambda_5^k.$$

Démonstration. Soit $R > P$ (donné par le lemme 3.2.5) tel que $\lambda_2 + C_5\lambda_2^R < 1$. Soit $x \in Z^6(k,N)$, et soient $u_1 < \ldots < u_p$ les instants par lesquels il est interdit pour une durée $n_i \geqslant R$. Ainsi, $x \in Z^5(n_1,\ldots,n_p,B_1)$ pour une certaine boule B_1. Si $\sum n_i \geqslant k/2$, on ne fait rien de plus. Sinon, soient $v_1 < \ldots < v_q$ les instants par lesquels x est interdit pour une durée $m_j < R$. Alors $\sum n_i + \sum m_j$ majore le nombre d'instants pour lesquels x est interdit, donc $\sum m_j \geqslant k/2$, si bien que $Rq \geqslant k/2$. On a donc

$$Z^6(k,N) \subset \left(\bigcup_{B_1} \bigcup_{\substack{n_1,\ldots,n_p \geqslant R \\ \sum n_i \geqslant k/2}} Z^5(n_1,\ldots,n_p,B_1) \right) \cup Z^4(k/(2R),N).$$

Les lemmes 3.3.5 et 3.3.6 impliquent donc que

$$\mathrm{Leb}(Z^6(k,N)) \leqslant \sum_{B_1} \sum_{\substack{n_1,\ldots,n_p \geqslant R \\ \sum n_i \geqslant k/2}} C_5(C_5\lambda_2^{n_1}) \cdots (C_5\lambda_2^{n_p}) \,\mathrm{Leb}(C(B_1)) + C_4\lambda_4^{k/(2R)}.$$

Comme les $C(B_1)$ sont deux à deux disjoints, on a $\sum \mathrm{Leb}(C(B_1)) \leqslant \mathrm{Leb}(M) < \infty$. Pour conclure, il suffit donc de vérifier que

$$\sum_{\substack{n_1,\ldots,n_p \geqslant R \\ \sum n_i \geqslant k/2}} (C_5\lambda_2^{n_1}) \cdots (C_5\lambda_2^{n_p})$$

décroît exponentiellement vite.

On utilise des séries génératrices :

$$\sum_n \sum_{\substack{n_1,\ldots,n_p \geqslant R \\ \sum n_i = n}} (C_5\lambda_2^{n_1}) \cdots (C_5\lambda_2^{n_p}) z^n = \sum_{p=0}^{\infty} \left(C_5 \sum_{n=R}^{\infty} \lambda_2^n z^n \right)^p = \frac{1-\lambda_2 z}{1-\lambda_2 z - C_5\lambda_2^R z^R}.$$

Comme $\lambda_2+C_5\lambda_2^R<1$, la fonction ainsi définie n'a pas de pôle sur un voisinage du disque unité de $\mathbb{C}$. Ainsi, ses coefficients décroissent exponentiellement vite, i.e. il existe des constantes $C_9>0$ et $\lambda_6<1$ telles que

$$\sum_{\substack{n_1,\dots,n_p\geqslant R\\ \sum n_i=n}} (C_5\lambda_2^{n_1})\cdots(C_5\lambda_2^{n_p})\leqslant C_9\lambda_6^n.$$

Il suffit ensuite de sommer sur $n\geqslant k/2$ pour conclure. □

3.3.3 Preuve du théorème 3.3.1

On vérifie dans chacun des quatre cas du théorème 3.3.1 la conclusion annoncée sur la mesure des queues. C'est pour cette preuve que les choix de σ,δ et ε^0 dans le paragraphe 3.2.1 sont importants.

Preuve des deux premiers cas. Supposons que $\mathrm{Leb}\{x\mid h^1_{\varepsilon^0}(x)>n\}=O(u_n)$. Rappelons que S_n désigne l'ensemble des points sélectionnés avant le temps n et que θ est défini dans le lemme 3.2.2. Montrons que

$$U_1\backslash S_n\subset\left\{x\in U_1\mid h^1_{\varepsilon^0}(x)>n\right\}\cup\left\{x\in U_1\mid \mathrm{dist}(x,\partial U_1)\leqslant\lambda_1^{\theta n/2}\right\}\cup Z^6(\theta n/2,n).$$

Cela conclura la preuve puisque les deuxième et troisième ensembles ont des mesures exponentiellement petites, par (3.12) et le lemme 3.3.7.

Soit donc x dans $U_1\backslash S_n$, mais ni dans $\{h^1_{\varepsilon^0}(x)>n\}$ ni dans $\left\{\mathrm{dist}(x,\partial U_1)\leqslant\lambda_1^{\theta n/2}\right\}$. Par le lemme 3.2.2, x a au moins θn temps hyperboliques entre 1 et n, donc au moins $\theta n/2$ entre $\theta n/2$ et n, qu'on notera $t_1<\ldots<t_k<n$. Comme $\mathrm{dist}(x,\partial U_1)\geqslant\lambda_1^{\theta n/2}$, on a en fait $x\in H_{t_i}(U_1)$ pour tous ces instants. Si x n'est pas interdit pour le temps t_i, il sera alors séctionné à cet instant, ce qui est absurde. Ainsi, $x\in I_{t_i}$. On obtient donc au moins $\theta n/2$ instants entre $\theta n/2$ et n pour lesquels x sera interdit, donc $x\in Z^6(\theta n/2,n)$. □

Preuve du troisième cas. Supposons que $\mathrm{Leb}\{x\mid h^2_{\varepsilon^0}(x)>n\}=O(u_n)$ où u_n est à décroissance polynomiale. On choisit K assez grand pour que $k(n):=\lfloor K\log n\rfloor$ vérifie $\lambda_5^{k(n)/2}=O(u_n)$ et $\lambda_2^{k(n)}=O(u_n)$. Notons $N(x)$ le nombre de temps hyperboliques de x entre 1 et n, on va montrer que

$$\mathrm{Leb}\{x\mid N(x)\leqslant 2k(n)\}=O((\log n)u_n). \tag{3.22}$$

Comme

$$U_1\backslash S_n\subset\{x\mid N(x)\leqslant 2k(n)\}\cup\{x\mid \mathrm{dist}(x,\partial U_1)\leqslant\lambda_1^{k(n)}\}\cup Z^6(k(n),n) \tag{3.23}$$

par les mêmes argument que ci-dessus, cela suffira à conclure la preuve puisque les deux autres ensembles ont une mesure $O(u_n)$, par (3.12) et le lemme 3.3.7.

On note SH_l^* l'ensemble des points qui ont l pour premier super temps hyperbolique non nul. D'après le lemme 3.2.1, un point x a au moins un super temps hyperbolique entre 1 et $h^2_{\varepsilon^0}(x)$, si bien que $\bigcup_{l>n} SH_l^* \subset \{x \mid h^2_{\varepsilon^0}(x) > n\}$. En particulier,

$$\mathrm{Leb}\left(\bigcup_{l>n} SH_l^*\right) = O(u_n). \tag{3.24}$$

Soit $K' > 2K/\theta > 2K$. Estimons la mesure des points qui n'ont pas de super temps hyperbolique entre $K' \log n$ et n. Si $k \geqslant 0$ désigne le dernier super temps hyperbolique précédant $K' \log n$, on a

$$\begin{aligned}\mathrm{Leb}\left(M \setminus \bigcup_{K'\log n \leqslant j \leqslant n} SH_j\right) &\leqslant \sum_{0\leqslant k<K'\log n} \mathrm{Leb}\left(SH_k \cap T^{-k}\left(\bigcup_{l>n-k} SH_l^*\right)\right)\\ &\leqslant \sum_{0\leqslant k<K'\log n} C\,\mathrm{Leb}\left(\bigcup_{l>n-k} SH_l^*\right)\end{aligned}$$

en utilisant le fait que $SH_k \subset H_k$ et le lemme 3.2.7. D'après (3.24), cette mesure est donc majorée par $CK'(\log n)u_{n-K'\log n} = O((\log n)u_n)$.

De plus, si un point x a un super temps hyperbolique j entre $K' \log n$ et n, alors il a par le lemme 3.2.2 au moins θj temps hyperboliques entre 1 et j, d'où $N(x) \geqslant \theta K' \log n > 2k(n)$. Ainsi,

$$\{x \mid N(x) \leqslant 2k(n)\} \subset M \setminus \bigcup_{K'\log n \leqslant j \leqslant n} SH_j.$$

Ce dernier ensemble a une mesure $O((\log n)u_n)$, ce qui démontre (3.22) et conclut donc la preuve. □

Preuve du quatrième cas. La preuve est analogue, en prenant $k(n) = \lfloor \varepsilon n^\eta \rfloor$, où ε est assez petit (on n'a besoin de ε que pour le cas $\eta = 1$, on peut prendre $k(n) = \lfloor n^\eta \rfloor$ lorsque $\eta < 1$). Dans (3.23), les deux termes les plus à droite ont une mesure en $O(e^{-c'n^\eta})$ pour un certain $c' > 0$, et il suffit donc de vérifier que $\mathrm{Leb}\{x \mid N(x) \leqslant 2k(n)\} = O(e^{-c'n^\eta})$. On procède comme ci-dessus, et on obtient une majoration en $\frac{2k(n)}{\theta} e^{-c(n-2k(n)/\theta)^\eta}$, ce qui suffit à conclure. □

3.4 Tour de Young

Dans cette partie, on montre comment obtenir une vraie tour de Young : on ne veut plus qu'un seul ensemble de la partition $\{U_1, \dots, U_N\}$, et des retours qui reviennent au même ensemble. Cela n'apporte rien puisqu'on a établi dans le chapitre 2 des résultats de décorrélation pour des tours localement Gibbs-Markov générales. Cependant, certains lecteurs sont

probablement plus à l'aise avec le cadre des tours de Young, dans lequel sont par exemple formulés les résultats de couplage de [You99]. On va voir qu'ici, il est en fait assez rapide de se ramener à une tour de Young.

Théorème 3.4.1. *Sous les hypothèses du théorème 3.3.2, soit μ une des mesures de probabilité invariantes ergodiques absolument continues données par ce théorème.*

Il existe alors un ouvert B d'intérieur non vide sur lequel μ est équivalente à Leb*, une partition $Z_1, Z_2, \dots$ de B et des temps $R'_1, R'_2, \dots$ tels que, pour tout j*

1. *$T^{R'_j}$ est un difféomorphisme entre Z_j et B.*
2. *$T^{R'_j}_{|Z_j}$ dilate les distances d'au moins $\sigma^{-1/2} > 1$.*
3. *La distorsion en volume de $T^{R'_j}_{|Z_j}$ est lipschitzienne.*
4. *Pour $x, y \in Z_j$ et $n \leqslant R'_j$, $\mathrm{dist}(T^n x, T^n y) \leqslant \mathrm{dist}(T^{R'_j} x, T^{R'_j} y)$.*

De plus, on a le même contrôle sur la taille des queues que dans le théorème 3.3.1.

Démonstration. Soient X' l'extension de M construite dans la preuve du théorème 3.3.2 en utilisant la partition (W_j), et m la mesure induite par Leb sur X'. Soit ν une des mesures invariantes ergodiques sur X' telle que $\pi_*(\nu) = \mu$. On identifie U_i avec $U_i \times \{0\}$ dans X'.

La mesure ν est équivalente à m sur l'un des U_i, mettons U_1. La base B de la tour sera U_1. Notons $U_2, \dots, U_s$ les autres ensembles U_i sur lesquels ν est équivalente à m. Soit T'_Y l'application induite par T' sur $Y = \{(x, 0)\} \subset X'$: elle est définie par $T'_Y(x, 0) = (T^{R_j}(x), 0)$ sur un élément W_j de la partition $\mathcal{B}$ du théorème 3.3.1 de temps de retour R_j. On définit une partition $\mathcal{B}^n$ de Y par $\mathcal{B}^n = \bigcap_0^{n-1} (T'_Y)^{-i}(\mathcal{B})$: ainsi, un élément de $\mathcal{B}^n$ est envoyé par $T'_Y, \dots, (T'_Y)^{n-1}$ sur des parties d'éléments de $\mathcal{B}$, et par $(T'_Y)^n$ sur l'un des U_i. Comme ν est ergodique, il existe $L > 0$ tel que, pour tout $i \leqslant s$, U_i contient un élément de $\mathcal{B}^n$, pour un certain $n < L$, dont l'image par $(T'_Y)^n$ est U_1.

Pour $x \in \bigcup_1^s U_i$, on définit une suite de temps $t_0(x) = 0, t_1(x), t_2(x), \dots$ et un entier $k(x)$ (correspondant au nombre d'itérations nécessaires avant de sélectionner x) de la façon suivante : soient $B_0 \in \mathcal{B}$ contenant x, et R_1 son temps de retour. On pose $t_1(x) = R_1$. Si $T'_Y(B_0) = U_1$, on pose $k(x) = 1$, et on s'arrête là. Sinon, $T^{R_1}(B_0)$ est l'un des ensembles U_i avec $2 \leqslant i \leqslant s$. On regarde dans quel ensemble B_1 de la partition $\mathcal{B}$ se trouve $T^{R_1}(x)$, et on appelle R_2 son temps de retour. Soit $t_2(x) = t_1(x) + R_2$. Si $T^{t_2(x)}(x)$ est dans U_1, on pose $k(x) = 2$ et on s'arrête là. Sinon, on regarde le prochain itéré de $T^{t_2(x)}$, qu'on note $T^{t_3(x)}$, et on continue ainsi. Plus formellement, cela revient à poser $k(x) = k(T^{R_1}x) + 1$ et $t_j(x) = t_{j-1}(T^{R_1}(x)) + t_1(x)$ pour tout $j \leqslant k(x)$. Autrement dit, $k(x)$ est le plus petit entier n tel que l'élément de $\mathcal{B}^n$ contenant x ait U_1 pour image par $(T'_Y)^n$.

Les éléments de la partition finale seront donnés par les ensembles Z_j ainsi construits contenus dans U_1, et les temps de retour correspondants seront donnés par $t_{k(x)}(x)$ pour $x \in Z_j$: ce temps de retour est indépendant du choix de x dans Z_j. Par construction, $T^{t_{k(x)}(x)}(Z_j) = U_1$, donc on a bien une tour de Young.

Au final, presque tout point sera sélectionné : on verra plus loin que la taille des queues tend vers 0. Notons $R(x) \in [1, \infty]$ le temps final de retour de x, i.e. $R(x) = \infty$ si x n'est jamais sélectionné et $R(x) = t_{k(x)}(x)$ sinon. Les propriétés de dilatation et de distorsion de la partition $\mathcal{B}$ restent vérifiées pour la tour de Young, il reste à obtenir le contrôle des queues.

Soit $\tau(x) = t_{k(x)}(x)$. Par définition de L, un élément de chaque U_i est sélectionné pour revenir à U_1 en au plus L étapes. Comme la distorsion est bornée, il existe donc $\varepsilon > 0$ tel que

$$\mathrm{Leb}(\tau = t_j \text{ ou } \dots \text{ ou } \tau = t_{j+L-1} \mid t_1, \dots, t_{j-1}, \tau > t_{j-1}) \geqslant \varepsilon. \tag{3.25}$$

De plus, encore par distorsion bornée,

$$\mathrm{Leb}\{t_{j+1} - t_j > n \mid t_1, \dots, t_j\} \leqslant C \sum_{W_k \in \mathcal{B}, R_k > n} \mathrm{Leb}(W_k), \tag{3.26}$$

ce dernier terme étant estimé par le théorème 3.3.1. On veut obtenir des estimées sur la taille des queues, i.e. sur $\mathrm{Leb}\{x \mid \tau(x) > n\}$, et on va pour cela utiliser (3.25) et (3.26).

Le problème considéré ici n'est pas le même que celui considéré dans [You99] : dans notre cas, on veut estimer la mesure de certains ensembles, alors que dans [You99] on estime une vitesse de mélange. Cependant, il y a dans les deux cas des suites de temps $t_1(x), t_2(x), \dots$, et ces temps vérifient les mêmes propriétés combinatoires : (3.25) et (3.26) sont analogues aux hypothèses que Young utilise dans la partie 4 de [You99] pour estimer la mesure $\mathrm{Leb}\{\tau(x) > n\}$ (l'hypothèse (3.25) est en fait un peu plus faible : Young peut prendre $L = 1$). Comme les preuves de Young ne dépendent que de la combinatoire de ces temps t_i, elles s'appliquent donc encore dans notre contexte. Dans le cas où (3.26) décroît exponentiellement vite, Young montre dans [You99, partie 4.2] que $\mathrm{Leb}\{\tau(x) > n\}$ décroît également exponentiellement vite, ce qui conclut la preuve (ses estimées marchent encore dans le cas où $L > 1$). Elle traite également le cas des suites à décroissance polynomiale : seul le cas $1/n^\gamma$ est écrit mais la même preuve marche directement dans le cas général, en notant que $u_{n/i} \leqslant u_n i^\gamma$ pour un certain $\gamma > 0$. Cependant, dans le cas exponentiel gauche (i.e. $0 < \eta < 1$), elle obtient des estimées un peu plus faibles que ce qui est annoncé dans le théorème 3.4.1, et on donne donc une preuve différente qui ne repose en fait que sur (3.25) et (3.26).

On suppose que (3.25) et (3.26) sont vérifiées, avec $C \sum_{W_k \in \mathcal{B},\ R_k \geqslant n} \mathrm{Leb}(W_k) \leqslant C' e^{-c' n^\eta} =: v_n$, pour $0 < \eta < 1$, et on montre que $\mathrm{Leb}\{x \mid \tau(x) > n\} = O(e^{-c'' n^\eta})$ pour un certain $c'' > 0$.

Si w^1 et w^2 sont deux suites réelles, on notera $w^1 \star w^2$ leur convolution, donnée par $(w^1 \star w^2)_n = \sum_{a+b=n} w^1_a w^2_b$. Si w est une suite, on notera aussi $w^{\star l}$ la suite obtenue en convolant l fois w avec elle-même.

D'après la proposition 2.2.13, pour K assez grand, la suite $w_n = 1_{n \geqslant K} v_n$ satisfait

$$\forall p \in \mathbb{N}, (w \star w)_p \leqslant w_p. \tag{3.27}$$

Soient k un entier et $A \subset \{1, \dots, k\}$. Pour $j \in A$, soit $n_j \geqslant 1$ un entier. Posons $Y(A, n_j) = \{x \mid k(x) \geqslant \sup A \text{ et } \forall j \in A, t_j(x) - t_{j-1}(x) = n_j\}$. En conditionnant successivement par

rapport aux différents instants, on trouve

$$\operatorname{Leb} Y(A,n_j) \leqslant \prod_{j\in A} \operatorname{Leb}\{t_j - t_{j-1} = n_j \mid t_1,\dots,t_{j-1}\} \leqslant \prod_{j\in A} v_{n_j}$$

par (3.26) et la définition de v_n.

Soit $q(n) = \lfloor \alpha n^\eta \rfloor$ où α sera choisi par la suite. Soit maintenant x tel que $\tau(x) > n$. Si $k(x) > q(n)$, i.e. x met plus de $q(n)$ étapes à être sélectionné, on ne fait rien. Autrement, soit $l = k(x) \leqslant q(n)$, et soient $n_j = t_j(x) - t_{j-1}(x)$ pour $j \leqslant l$. Notons $A = \{j \mid n_j \geqslant K\}$. Ainsi, $x \in Y(A, n_j)$. De plus, comme $\sum n_j = \tau(x) > n$, on a $\sum_{j\in A} n_j \geqslant n - Kq(n) \geqslant n/2$ si n est assez grand. On a donc montré

$$\{x \mid \tau(x) > n\} \subset \{k(x) > q(n)\} \cup \bigcup_{A\subset\{1,\dots,q(n)\}} \bigcup_{\substack{n_j\geqslant K \\ \sum_A n_j \geqslant n/2}} Y(A,n_j). \tag{3.28}$$

Par (3.25), on a $\operatorname{Leb}\{k(x) > q(n)\} \leqslant (1-\varepsilon)^{q(n)/L} \leqslant e^{-c''n^\eta}$ pour un certain c''. De plus, en notant $l = \operatorname{Card} A$ et en utilisant (3.27), on a

$$\begin{aligned}
\operatorname{Leb}\left(\bigcup_{A\subset\{1,\dots,q(n)\}} \bigcup_{\substack{n_j\geqslant K \\ \sum_A n_j \geqslant n/2}} Y(A,n_j) \right) &\leqslant \sum_{A\subset\{1,\dots,q(n)\}} \sum_{\substack{n_j\geqslant K \\ \sum_A n_j \geqslant n/2}} \prod_{j\in A} v_{n_j} \\
&\leqslant \sum_{0\leqslant l\leqslant q(n)} \binom{q(n)}{l} \sum_{\substack{n_1,\dots,n_l\geqslant K \\ \sum n_j \geqslant n/2}} v_{n_1}\cdots v_{n_l} \\
&= \sum_{0\leqslant l\leqslant q(n)} \binom{q(n)}{l} \sum_{n/2}^{\infty} \left(w^{\star l}\right)_p \\
&\leqslant \sum_{0\leqslant l\leqslant q(n)} \binom{q(n)}{l} \sum_{n/2}^{\infty} w_p = 2^{q(n)} \sum_{n/2}^{\infty} w_p.
\end{aligned}$$

Comme $w_n = O(e^{-c'n^\eta})$, on obtient $\sum_{n/2}^\infty w_p = O(n^{1-\eta} e^{-c'(n/2)^\eta})$ en comparant avec une intégrale. Ainsi, si on choisit α assez petit, on aura $2^{q(n)} \sum_{n/2}^\infty w_p = O(e^{-c''n^\eta})$ pour un certain $c'' > 0$. D'après (3.28), on a montré que $\operatorname{Leb}\{\tau(x) > n\} = O(e^{-c''n^\eta})$. □

On peut déduire du théorème 3.4.1 presque tous les résultats du théorème 3.3.1 sur la vitesse de décorrélation en utilisant les résultats de [You99], qui sont montrés par une méthode de couplage. Le seul cas où ces résultats ne sont pas suffisants est le cas où $u_n = e^{-cn^\eta}$, puisque Young obtient alors une vitesse en $O(e^{-n^{\eta'}})$ pour tout $\eta' < \eta$, ce qui est plus faible que les résultats du théorème 3.3.1. Cependant, en modifiant un petit peu les estimées combinatoires de Young, on peut démontrer le lemme suivant, et donc obtenir tous les résultats du théorème 3.3.1 par couplage.

Lemme 3.4.2. *Soit (X, T') une tour de Young mélangeante, et supposons que le temps de retour à la base R satisfait $m(R > n) = O(e^{-cn^\eta})$ pour un certain $0 < \eta < 1$. Alors, si f et g sont deux fonctions respectivement hölderienne et bornée, leur corrélations sont majorées par $e^{-c'n^\eta}$ pour un certain $c' > 0$.*

Démonstration. On reprend les arguments de couplage de la partie 3 de [You99], avec la construction des temps d'arrêt. Pour obtenir les estimées recherchées, on remplace les majorations directes de la partie 4.2 de [You99] par les arguments qu'on a utilisés dans la preuve du cas exponentiel gauche du théorème 3.4.1 : la combinatoire des deux problèmes, *a priori* différents, est exactement la même ! □

Chapitre 4

Applications markoviennes et théorèmes limites

Sommaire

4.1 Théorèmes limites des applications Gibbs-Markov 142
4.1.1 Estimées préliminaires 142
4.1.2 Théorème central limite 146
4.1.3 Lois stables 151
4.2 Tours localement Gibbs-Markov : approche élémentaire 153
4.2.1 Théorèmes limites et induction 153
4.2.2 Application aux tours localement Gibbs-Markov 160
4.2.3 Exemples 163
Applications de Liverani-Saussol-Vaienti 163
Produit gauche en dimension 2 165
4.3 Un résultat spectral perturbatif 168
4.3.1 Inversibilité de $I - R(z,t)$ 169
4.3.2 Convergence de $\widetilde{R}(z,t)$ 171
4.3.3 Conclusion de la preuve du théorème 4.3.1 175
4.4 Tours localement Gibbs-Markov : méthode spectrale 177
4.4.1 Réduction au cas mélangeant 178
4.4.2 Le résultat local sur Y 180
4.4.3 Le résultat local sur $\bigcup_{i=0}^{N-1} T^i(Y)$ 182
4.4.4 Démonstration du théorème central limite 184
4.4.5 Preuve des autres cas du théorème 4.4.1 186

Dans ce chapitre, on s'intéresse aux théorèmes limites de type probabiliste pour les tours localement Gibbs-Markov. D'après la partie 1.3, cela impliquera le même type de résultats

pour les applications localement Gibbs-Markov générales, et même pour les applications qui admettent des modèles localement Gibbs-Markov. Plus précisément, étant données une tour localement Gibbs-Markov transitive $T : X \to X$ préservant une mesure de probabilité m et une fonction $f : X \to \mathbb{R}$ de moyenne nulle, on s'intéresse à la convergence en loi de la suite $\frac{1}{B_n} S_n f$ où B_n est une normalisation convenablement choisie. On dit que cette suite converge en loi vers une variable aléatoire Z si, pour toute partie mesurable I de $\mathbb{R}$ avec $P(Z \in \partial I) = 0$,

$$m\left\{x \in X \mid \frac{1}{B_n} S_n f(x) \in I\right\} \to P(Z \in I).$$

La version la plus classique de ce type de résultat est le théorème central limite, lorsque Z est gaussienne et $B_n = \sqrt{n}$, mais d'autres lois vont également apparaître.

Dans la première partie, on étudie les théorèmes limites pour les applications Gibbs-Markov, en utilisant une méthode de perturbation spectrale classique due à Nagaev et Guivarc'h (voir par exemple [Bro96]). Dans la deuxième partie, qui reprend une partie de [Gou07], on montre comment en déduire des théorèmes limites pour des tours localement Gibbs-Markov par une méthode élémentaire, mais sous des hypothèses assez fortes sur le temps de retour φ. Dans les parties qui suivent, on décrit une méthode plus sophistiquée développée dans [Gou04a], utilisant les opérateurs de transfert de premier retour du chapitre 2, qui permet de s'affranchir de ces hypothèses restrictives.

4.1 Théorèmes limites des applications Gibbs-Markov

Dans toute cette partie, on considère une application Gibbs-Markov transitive $T : X \to X$ de partition α, X étant muni de la distance $d_\tau(x, y) = \tau^{s(x,y)}$ où $s(x, y)$ est le temps de séparation de x et de y et $\tau < 1$ est choisi assez proche de 1 pour que la distorsion de T soit τ-hölderienne.

Soit aussi $f : X \to \mathbb{R}$ une fonction, pour laquelle on va essayer de démontrer un théorème limite.

4.1.1 Estimées préliminaires

Soit $\underline{a} = [a_0, \dots, a_{n-1}]$ un cylindre de longueur n. Pour tout $t \in \mathbb{R}$, on définit un opérateur $M_{\underline{a},t}$ par $M_{\underline{a},t} u(x) = e^{itS_n f(x')} g^{(n)}(x') u(x')$ si la préimage x' de x dans $\underline{a}$ par T^n existe, et $M_{\underline{a},t} u(x) = 0$ sinon. C'est une perturbation de l'opérateur $M_{\underline{a}}$ qui a été estimé dans le lemme 1.1.13.

On va obtenir des estimées similaires sur $M_{\underline{a},t}$. Pour des raisons techniques, on aura besoin de varier les espaces utilisés. On notera $\mathcal{L}_\tau$ l'espace des fonctions bornées, et lipschitziennes pour la distance d_τ sur chaque élément de la partition α. Pour $u : X \to \mathbb{C}$ et $a \subset X$, on

notera

$$D_\tau u(a) = \inf\{C > 0 \mid \text{pour presque tous } x, y \in a, |u(x) - u(y)| \leqslant C d_\tau(x, y)\},$$

comme dans l'équation (1.10) mais en précisant la dépendance en la distance.

Lemme 4.1.1. *Soit $0 < \eta \leqslant 1$. Il existe une constante $C > 0$ telle que, pour tout $n \in \mathbb{N}^*$, pour tout cylindre $\underline{a} = [a_0, \dots, a_{n-1}]$ de longueur n, pour tout $t \in \mathbb{R}$, pour toute fonction $u : X \to \mathbb{C}$,*

$$\|M_{\underline{a},t} u\|_{\mathcal{L}_{\tau^\eta}} \leqslant C\big(1 + |t|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)}\big) \left(m[\underline{a}] \tau^{\eta n} \|u\|_{\mathcal{L}_{\tau^\eta}} + \int_{\underline{a}} |u| \right).$$

Démonstration. Notons $\psi_t(x) = e^{itS_n f(x')}$ si x a une préimage x' dans $\underline{a}$, et 0 sinon. Alors $M_{\underline{a},t} u(x) = \psi_t(x) M_{\underline{a}} u(x)$. De plus, pour toute fonction v, $\|\psi_t v\|_{\mathcal{L}_{\tau^\eta}} \leqslant \|\psi_t\|_{\mathcal{L}_{\tau^\eta}} \|v\|_{\mathcal{L}_{\tau^\eta}}$, par (1.11). Compte tenu de la majoration sur $\|M_{\underline{a}} u\|_{\mathcal{L}_{\tau^\eta}}$ donnée par le lemme 1.1.13 appliqué avec la distance d_{τ^η} et en utilisant la propriété de grande image, il suffit de démontrer que $\|\psi_t\|_{\mathcal{L}_{\tau^\eta}} \leqslant C(1 + |t|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)})$.
On a $|\psi_t| \leqslant 1$, il faut donc majorer la norme Lipschitz de ψ_t. Fixons deux points x et y dans le même élément de partition, qui ont deux antécédents x' et y' dans $\underline{a}$. Soit C tel que, pour tout $w \in \mathbb{R}$, $|e^{iw} - 1| \leqslant C|w|^\eta$: une telle constante existe car cette condition est facile à garantir et sur un voisinage de 0 et pour $|w|$ assez grand. Alors

$$\begin{aligned} |e^{itS_n f(x')} - e^{itS_n f(y')}| &= |e^{it\sum_{i=0}^{n-1}(f(T^i x') - f(T^i y'))} - 1| \leqslant \sum_{i=0}^{n-1} |e^{it(f(T^i x') - f(T^i y'))} - 1| \\ &\leqslant C|t|^\eta \sum_{i=0}^{n-1} |f(T^i x') - f(T^i y')|^\eta \leqslant C|t|^\eta \sum_{i=0}^{n-1} \big[D_\tau f(a_i) \tau^{n-i} d_\tau(x, y)\big]^\eta. \end{aligned}$$

□

On aura aussi besoin d'un résultat de continuité en fonction de t.

Lemme 4.1.2. *Soit $0 < \eta \leqslant 1$. Il existe une constante $C > 0$ telle que, pour tout $n \in \mathbb{N}^*$, pour tout cylindre $\underline{a} = [a_0, \dots, a_{n-1}]$ de longueur n, pour tout $t \in \mathbb{R}$, pour tout $s \in \mathbb{R}$, pour toute fonction $u : X \to \mathbb{C}$,*

$$\begin{aligned} \|M_{\underline{a},t+s} u - M_{\underline{a},t} u\|_{\mathcal{L}_{\tau^\eta}} \leqslant C &\left(|s|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)} + \frac{1}{m[\underline{a}]} \int_{\underline{a}} |e^{isS_n f} - 1| \right) \cdot \\ &\cdot \big(1 + |t|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)}\big) \cdot \left(m[\underline{a}] \tau^{\eta n} \|u\|_{\mathcal{L}_{\tau^\eta}} + \int_{\underline{a}} |u| \right). \end{aligned}$$

Démonstration. On réutilise la fonction ψ_t de la preuve précédente. Posons $\chi(x) = \psi_{t+s}(x) - \psi_t(x)$, si bien que $M_{\underline{a},t+s}u - M_{\underline{a},t}u = \chi \cdot M_{\underline{a}}u$. Compte tenu des estimées sur $M_{\underline{a}}u$ données par le lemme 1.1.13, il suffit donc d'estimer $\|\chi\|_{\mathcal{L}_{\tau^\eta}}$.

Notons que $\chi = (\psi_s - 1)\psi_t$, si bien que $\|\chi\|_{\mathcal{L}_{\tau^\eta}} \leqslant \|\psi_s - 1\|_{\mathcal{L}_{\tau^\eta}} \|\psi_t\|_{\mathcal{L}_{\tau^\eta}}$. La norme de ψ_t a été majorée par $C\big(1 + |t|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)}\big)$ dans la preuve du lemme 4.1.1. De même, la norme de Lipschitz de $\psi_s - 1$ est majorée par $C|s|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)}$. Il reste donc à estimer $|\psi_s(x) - 1|$. Pour $y' \in \underline{a}$, on a

$$\begin{aligned}|e^{isS_nf(x')} - 1| &\leqslant |e^{isS_nf(x')} - e^{isS_nf(y')}| + |e^{isS_nf(y')} - 1| \\ &\leqslant C|s|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)} + |e^{isS_nf(y')} - 1|.\end{aligned}$$

En intégrant, on obtient

$$|e^{isS_nf(x')} - 1| \leqslant C|s|^\eta \sum_{i=0}^{n-1} D_\tau f(a_i)^\eta \tau^{\eta(n-i)} + \frac{1}{m[\underline{a}]} \int_{\underline{a}} |e^{isS_nf} - 1|,$$

ce qui est la majoration recherchée sur $\|\psi_s - 1\|_\infty$. □

Corollaire 4.1.3. *Soit $\widehat{T}_t(u) = \widehat{T}(e^{itf}u)$. Supposons que $\sum_{a\in\alpha} m[a]D_\tau f(a) < \infty$.*

Alors $\widehat{T}_t$ agit continûment sur $\mathcal{L}_\tau$ pour tout $t \in \mathbb{R}$, et $t \mapsto \widehat{T}_t$ est continu en 0. Plus précisément,

$$\big\|\widehat{T}_t - \widehat{T}\big\|_{\mathcal{L}_\tau} = O(|t| + E(|e^{itf} - 1|)).$$

En particulier, si f est intégrable, $\big\|\widehat{T}_t - \widehat{T}\big\|_{\mathcal{L}_\tau} = O(|t|)$.

De plus, pour tout $t \in \mathbb{R}$, le rayon spectral de $\widehat{T}_t$ agissant sur $\mathcal{L}_\tau$ est majoré par 1, et son rayon spectral essentiel est majoré par τ. Comme $\widehat{T}$ a une valeur propre simple isolée en 1, $\widehat{T}_t$ a donc également une valeur propre simple $\lambda(t)$ proche de 1 pour t assez petit. Elle vérifie $|\lambda(t) - 1| = O(|t| + E(|e^{itf} - 1|))$.

Démonstration. On va utiliser les lemmes 4.1.1 et 4.1.2 pour $\eta = 1$ et l'égalité

$$\widehat{T}_t^n = \sum_{\underline{a}=[a_0,\dots,a_{n-1}]} M_{\underline{a},t}.$$

Pour $n = 1$, le lemme 4.1.1 donne $\|M_{\underline{a},t}\| \leqslant C(1 + |t|D_\tau f(a_0))m[\underline{a}]$, donc en sommant $\big\|\widehat{T}_t\big\| \leqslant C + |t| \sum_{a\in\alpha} m[a]D_\tau f(a) < \infty$. Ainsi, $\widehat{T}_t$ agit continûment sur $\mathcal{L}_\tau$.

Pour n arbitraire, on obtient en sommant les inégalités du lemme 4.1.1 que

$$\big\|\widehat{T}_t^n u\big\|_{\mathcal{L}_\tau} \leqslant C \sum_{\underline{a}} (1 + |t| \sum_{i=0}^{n-1} D_\tau f(a_i)\tau^{n-i}) \left[m[\underline{a}]\tau^n \|u\|_{\mathcal{L}_\tau} + \min(m[\underline{a}] \|u\|_\infty, \|u\|_1)\right] \quad (4.1)$$

Remarquons que

$$\sum_{\underline{a}} m[\underline{a}] \sum_{i=0}^{n-1} D_\tau f(a_i)\tau^{n-i}$$
$$= \sum_a \sum_{i=0}^{n-1} D_\tau f(a)\tau^{n-i} \sum_{a_0,\dots,a_{i-1},a_{i+1},\dots,a_{n-1}} m[a_0,\dots,a_{i-1},a,a_{i+1},\dots,a_{n-1}].$$

La somme sur $a_{i+1},\dots,a_{n-1}$ donne simplement la mesure de $[a_0,\dots,a_{i-1},a]$. Quand on somme sur $a_0,\dots,a_{i-1}$, on obtient la mesure de $T^{-i}a$, qui coïncide avec celle de a par invariance. Ainsi, la somme de l'équation précédente est bornée par

$$\sum_a D_\tau f(a) m(a) \frac{1}{1-\tau} < \infty.$$

Cette inégalité montre que, dans (4.1), la contribution du terme $\tau^n \|u\|_{\mathcal{L}_\tau}$ est bornée par $C(1+|t|)\tau^n \|u\|_{\mathcal{L}_\tau}$.

Pour traiter le terme $\min(m[\underline{a}] \|u\|_\infty, \|u\|_1)$, on découpe l'ensemble des cylindres de longueur n en deux sous-ensembles, U infini et V fini, sur lesquels on utilisera respectivement le premier et le deuxième terme du minimum. On choisit V assez grand pour que

$$\sum_{\underline{a}\in U} m[\underline{a}] \sum_{i=0}^{n-1} D_\tau f(a_i)\tau^{n-i} \leqslant \tau^n.$$

La contribution des éléments de U à (4.1) est alors bornée par $C(1+|t|)\tau^n \|u\|_{\mathcal{L}_\tau}$, tandis que celle de V est au plus $(1+|t|)C(n)\|u\|_1$ pour une certaine constante $C(n)$.

On a montré une inégalité du type

$$\|\widehat{T}_t^n u\|_{\mathcal{L}_\tau} \leqslant C(1+|t|)\tau^n \|u\|_{\mathcal{L}_\tau} + C(n)(1+|t|)\|u\|_1.$$

Comme l'injection de $\mathcal{L}_\tau$ dans L^1 est compacte d'après la proposition 1.1.16, le théorème 1.1.11 assure donc que le rayon spectral essentiel de $\widehat{T}_t$ sur $\mathcal{L}_\tau$ est majoré par τ. Si son rayon spectral était strictement plus grand que 1, il admettrait donc une fonction propre u telle que $\widehat{T}_t u = \lambda u$ avec $|\lambda| > 1$. Ceci est absurde puisque $\widehat{T}_t$ est une contraction sur L^1. Ainsi, le rayon spectral de $\widehat{T}_t$ est au plus 1.

Montrons maintenant la continuité de $s \mapsto \widehat{T}_s$ en 0, en sommant les égalités données par le lemme 4.1.2 sur les cylindres de longueur 1 (pour $t = 0$) :

$$\|\widehat{T}_s - \widehat{T}\|_{\mathcal{L}_\tau} \leqslant C \sum_{a\in\alpha} \left(|s| D_\tau f(a) + \frac{1}{m[a]} \int_a |e^{isf} - 1| \right) m[a] \leqslant C|s| + C \int_X |e^{isf} - 1|.$$

Par convergence dominée, cette estimée tend vers 0 quand $s \to 0$, ce qui montre la continuité.

Finalement, comme T est Markov-transitive, $\widehat{T}$ a une valeur propre simple isolée en 1 d'après la proposition 1.1.17. Par continuité, la proposition 1.1.10 assure que, pour t assez petit, $\widehat{T}_t$ a également une valeur propre simple isolée proche de 1, qui vérifie $|\lambda(t) - \lambda(0)| = O\left(\|\widehat{T}_t - \widehat{T}\|\right)$. □

Notons que, pour obtenir la continuité de $t \mapsto \widehat{T}_t$ en $t \neq 0$, il faudrait avoir $\sum m[a] D_\tau f(a)^2 < \infty$ pour que les estimées du lemme 4.1.2 soient sommables, ce qui n'est pas impliqué par les hypothèses du corollaire. Cependant, en prenant $\eta = 1/2$, on a $\sum m[a] \left(D_\tau f(a)^\eta\right)^2 < \infty$, ce qui peut suffire pour montrer la continuité de $t \mapsto \widehat{T}_t$ sur $\mathcal{L}_{\tau^{1/2}}$. Ce type de considérations sera utile dans la suite mais n'est pas important ici.

Notons que le lemme s'applique encore pour tout $\tau' \geqslant \tau$. En particulier, on obtient une valeur propre $\lambda'(t)$ de $\widehat{T}_t$ agissant sur $\mathcal{L}_{\tau'}$, pour t assez petit. Mais $\lambda(t)$ est déjà valeur propre puisque la fonction propre correspondante est dans $\mathcal{L}_\tau$ et donc dans $\mathcal{L}_{\tau'}$. Par unicité, on obtient $\lambda(t) = \lambda'(t)$. Pour obtenir un développement précis des valeurs propres, on pourra donc se placer sur n'importe quel espace $\mathcal{L}_{\tau'}$ avec $\tau' \geqslant \tau$. On utilisera parfois un espace $\mathcal{L}_{\tau^\eta}$ avec η petit, pour le même genre de raison technique que ci-dessus.

Par définition de $\widehat{T}_t$, on a $\int_X e^{itS_n f} = \int_X \widehat{T}_t^n(1)$, qui est donc comparable à $\lambda(t)^n$ (pour t petit et n grand). Comme la compréhension de $\int_X e^{itS_n f}$ est l'ingrédient essentiel pour démontrer des théorèmes limites, il va donc souvent suffire de comprendre précisément $\lambda(t)$ pour obtenir ce type de résultat.

4.1.2 Théorème central limite

Dans ce paragraphe, on décrit le comportement de $\lambda(t)$ lorsque la fonction f est dans L^2, ce qui va nous permettre de démontrer un théorème central limite.

Théorème 4.1.4. *Soient $T : X \to X$ une application Gibbs-Markov transitive préservant une mesure de probabilité m, et $0 < \tau < 1$ assez grand pour que la distorsion de T soit τ-hölderienne. Soit $f : X \to \mathbb{R}$ telle que $\sum_{a\in\alpha} m[a] D_\tau f(a) < \infty$, $f \in L^2$ et $\int f = 0$.*

Soient $\widehat{T}_t(u) = \widehat{T}(e^{itf}u)$, et $\lambda(t)$ la valeur propre de $\widehat{T}_t$ pour t petit donnée par le corollaire 4.1.3. Il existe alors une fonction $\Lambda(t)$ avec $\Lambda(t) \sim t^2$ en 0 et une constante $\sigma^2 \geqslant 0$ telles que $\lambda(t) = 1 - \frac{\sigma^2}{2}\Lambda(t)$.

Soient $a = (I - \widehat{T})^{-1}(\widehat{T}f)$ (où on considère $I - \widehat{T}$ agissant sur l'ensemble des fonctions de $\mathcal{L}_\tau$ d'intégrale nulle, sur lequel il est inversible), et $u = (I - \widehat{T})^{-1}(f)$. Alors

$$\sigma^2 = \int f^2 \, \mathrm{d}m + 2\int af \, \mathrm{d}m = \int \widehat{T}(u^2) - (\widehat{T}u)^2 \, \mathrm{d}m.$$

Finalement, $\sigma^2 = 0$ si et seulement si il existe une fonction $\chi \in \mathcal{L}_\tau$ avec $f = \chi - \chi \circ T$.

La proposition 1.1.15 assure que $\widehat{T}f \in \mathcal{L}_\tau$. Comme $I-\widehat{T}$ est inversible sur les fonctions de $\mathcal{L}_\tau$ d'intégrale nulle d'après la proposition 1.1.17, on peut effectivement poser $a = (I-\widehat{T})^{-1}(\widehat{T}f)$. Cependant, $u = (I-\widehat{T})^{-1}(f)$ n'est pas défini *a priori*. En fait, on pose simplement $u = f+a$, et on a bien $(I-\widehat{T})u = f - \widehat{T}f + (I-\widehat{T})a = f$.

Démonstration. Soient P_t la projection spectrale de $\widehat{T}_t$ correspondant à la valeur propre $\lambda(t)$, et ξ_t la fonction propre correspondante normalisée de telle sorte que $\int \xi_t = 1$. Comme les valeurs propres et les fonctions propres dépendent de façon lipschitzienne de l'opérateur, on a $\|\xi_t - 1\| = O\big(\|\widehat{T}_t - \widehat{T}\|\big) = O(|t|)$ par le corollaire 4.1.3. De la même façon, $\lambda(t) = 1+O(t)$. Notons déjà que

$$E(e^{itf}) = 1 - \frac{t^2}{2}\int f^2 \,\mathrm{d}m + o(t^2). \tag{4.2}$$

En effet, si on considère $\Phi(t) = E(e^{itf})$ comme la fonction caractéristique de la variable aléatoire f, il est bien connu que, si $f \in L^2$, alors Φ est C^2 avec $\Phi(0) = 1$, $\Phi'(0) = iE(f) = 0$ et $\Phi''(0) = -E(f^2)$ ([Fel66, corollaire du lemme XV.4.2]).

Comme $\lambda(t)\xi_t = \widehat{T}_t\xi_t$, on obtient en intégrant que

$$\begin{aligned}\lambda(t) &= \int \widehat{T}_t\xi_t = \int (\widehat{T}_t - \widehat{T})(\xi_t - 1) + \int (\widehat{T}_t - \widehat{T})(1) + \int \widehat{T}\xi_t \\ &= \int (\widehat{T}_t - \widehat{T})(\xi_t - 1) + E(e^{itf} - 1) + 1.\end{aligned} \tag{4.3}$$

Comme $\widehat{T}_t - \widehat{T} = O(t)$ et $\xi_t - 1 = O(t)$, on obtient $\lambda(t) = 1 + O(t^2)$, en utilisant (4.2) pour estimer $E(e^{itf} - 1)$.

Ainsi,

$$\begin{aligned}\frac{\xi_t - \xi_0}{t} &= \frac{\lambda(t)\xi_t - \xi_0}{t} + O(t) = \frac{\widehat{T}_t\xi_t - \widehat{T}_0\xi_0}{t} + O(t) \\ &= (\widehat{T}_t - \widehat{T}_0)\frac{\xi_t - \xi_0}{t} + \widehat{T}_0\frac{\xi_t - \xi_0}{t} + \frac{\widehat{T}_t - \widehat{T}_0}{t}\xi_0 + O(t).\end{aligned} \tag{4.4}$$

Mais $(\widehat{T}_t - \widehat{T}_0)(\xi_0) = \widehat{T}(e^{itf} - 1)$. De plus, $\frac{\xi_t - \xi_0}{t}$ est borné et $\widehat{T}_t - \widehat{T}_0 = O(t)$, donc $(\widehat{T}_t - \widehat{T}_0)\frac{\xi_t - \xi_0}{t} = O(t)$. Ainsi,

$$(I - \widehat{T})\frac{\xi_t - \xi_0}{t} = \widehat{T}\left(\frac{e^{itf} - 1}{t}\right) + O(t). \tag{4.5}$$

La suite $(\xi_t - \xi_0)/t$ est bornée dans $\mathcal{L}_\tau$. Par compacité, on peut en extraire une sous-suite qui converge dans L^2 vers une fonction ia, qui appartient encore à $\mathcal{L}_\tau$; on a multiplié par i pour des raisons qui vont apparaître plus tard. Dans L^2, $\frac{e^{itf}-1}{t} \to if$ par le théorème de convergence dominée car $\left|\frac{e^{itf}-1}{t}\right|^2 \leqslant |f|^2$ intégrable. Par conséquent, $(I - \widehat{T})a = \widehat{T}(f)$. De

plus, par la proposition 1.1.15, $\widehat{T}f \in \mathcal{L}_\tau$, et $\int \widehat{T}f = \int f = 0$. Comme $I - \widehat{T}$ est inversible sur les fonctions de $\mathcal{L}_\tau$ d'intégrale nulle, on obtient $a = (I-\widehat{T})^{-1}(\widehat{T}f)$. Ainsi, $(\xi_t - \xi_0)/t$ a une unique valeur d'adhérence, donc elle converge (dans L^2) vers $ia = i(I-\widehat{T})^{-1}(\widehat{T}f)$.

Par continuité du produit $L^2 \times L^2 \to L^1$, on obtient la convergence de $\int \frac{e^{itf}-1}{t}\frac{\xi_t-\xi_0}{t}$ vers $\int (if)\cdot(ia)$. Ainsi,

$$\int (\widehat{T}_t - \widehat{T})(\xi_t - 1) = \int \widehat{T}((e^{itf}-1)(\xi_t-1)) = \int (e^{itf}-1)(\xi_t-1) = -t^2 \int fa + o(t^2).$$

L'équation (4.3) donne alors

$$\lambda(t) = 1 + E(e^{itf}-1) + \int (\widehat{T}_t - \widehat{T})(\xi_t - 1) = 1 - \frac{t^2}{2}\int f^2 - t^2 \int fa + o(t^2).$$

On transforme un petit peu ce résultat pour obtenir l'autre expression annoncée pour $\lambda(t)$. Notons que $f = u - \widehat{T}u$ et $a = u - f$.

$$\begin{aligned} -\lambda''_{|t=0} &= \int f^2 + 2\int af = \int f^2 + 2(u-f)f = \int (u-\widehat{T}u)^2 + \int 2\widehat{T}u\cdot(u-\widehat{T}u) \\ &= \int (u+\widehat{T}u)(u-\widehat{T}u) = \int u^2 - \int (\widehat{T}u)^2 = \int \widehat{T}(u^2) - (\widehat{T}u)^2. \end{aligned}$$

On a $\widehat{T}(u^2)(x) = \sum_{Ty=x} g(y)u^2(y)$ et $(\widehat{T}u)^2(x) = \left(\sum_{Ty=x} g(y)u(y)\right)^2$ où g désigne l'inverse du jacobien de T. Comme la mesure m est invariante, $\widehat{T}1 = 1$, i.e. $\sum_{Ty=x} g(y) = 1$. La convexité de $w \mapsto w^2$ implique alors que $(\widehat{T}u)^2(x) \leqslant \widehat{T}(u^2)(x)$, ce qui montre que $-\lambda'' \geqslant 0$. On peut donc écrire $\lambda'' = -\sigma^2$.

Finalement, si $\sigma^2 = 0$, alors il y a égalité dans l'inégalité de convexité, si bien que tous les $u(y)$ sont égaux pour presque tout x. En effet, pour presque tout x, tous les termes $g(y)$ pour y préimage de x sont non nuls, puisque T est non-singulière. Soit $\chi = -\widehat{T}u = -a \in \mathcal{L}_\tau$. On obtient $u = -\chi \circ T$ presque partout. Comme $f = u - \widehat{T}u$, cela donne que $f = \chi - \chi\circ T$. Réciproquement, si $f = \chi - \chi\circ T$ avec $\chi \in \mathcal{L}_\tau$, alors $\widehat{T}_t(e^{-it\chi}) = e^{-it\chi}$ pour tout t, donc $\lambda(t) = 1$ pour tout t, puis $\sigma^2 = 0$. □

On en déduit que f vérifie un théorème central limite :

Corollaire 4.1.5. *Sous les hypothèses du théorème 4.1.4, $\frac{1}{\sqrt{n}}S_n f$ converge en loi vers la loi normale $\mathcal{N}(0,\sigma^2)$.*

Démonstration. La convergence en loi équivaut à la convergence simple des fonctions caractéristiques correspondantes, si bien qu'il suffit de voir que $\int e^{itS_nf/\sqrt{n}} \to e^{-\sigma^2t^2/2}$.

On décompose $\widehat{T}_t = \lambda(t)P_t + Q_t$, où P_t est le projecteur spectral correspondant à la valeur propre $\lambda(t)$ de $\widehat{T}_t$. Par continuité de $t \mapsto \widehat{T}_t$ en 0, il existe $C > 0$ et $\theta < 1$ tels que, pour tout t assez petit, pour tout $n \in \mathbb{N}$, $\|Q_t^n\| \leqslant C\theta^n$. Fixons $t \in \mathbb{R}$, et soit $n > 0$ assez grand. Alors

$$\int e^{itS_n f/\sqrt{n}} = \int \widehat{T}^n_{t/\sqrt{n}}(1) = \int \lambda(t/\sqrt{n})^n P_{t/\sqrt{n}}(1) + \int Q^n_{t/\sqrt{n}}(1).$$

Le développement de $\lambda(t)$ donné par le théorème 4.1.4 assure que $\lambda(t/\sqrt{n})^n \to e^{-\sigma^2 t^2/2}$ lorsque $n \to \infty$. De plus, par continuité, $\int P_{t/\sqrt{n}}(1) \to \int P_0(1) = 1$. Finalement, $|Q^n_{t/\sqrt{n}}(1)| \leqslant C\theta^n$ tend uniformément vers 0 quand $n \to \infty$. Ainsi, $\int e^{itS_n f/\sqrt{n}} \to e^{-\sigma^2 t^2/2}$, ce qui est la conclusion recherchée. □

En fait, comme le terme d'erreur est très bien contrôlé (il est exponentiellement petit), on peut raffiner ce théorème, et obtenir par exemple la vitesse de convergence ou le théorème de la limite locale (voir par exemple [Bro96]). Le chapitre 5 sera consacré à de tels raffinements, qu'on démontrera directement dans le cadre plus général des tours localement Gibbs-Markov.

Le lemme qui suit sera utile dans le chapitre 5, pour préciser la vitesse de convergence dans le théorème central limite.

Lemme 4.1.6. *Sous les hypothèses du théorème 4.1.4, supposons de plus que, pour un certain $\eta > 0$,*

$$\sum_{a\in\alpha} m[a]\left(\|f_{|a}\|_\infty D_\tau f(a) + \|f_{|a}\|^2_\infty D_\tau f(a)^\eta\right) < \infty. \tag{4.6}$$

Alors $\lambda(t) = E(e^{itf}) - t^2 \int af\,\mathrm{d}m + O(t^3)$, où $a = (I - \widehat{T})^{-1}(\widehat{T}f)$. En particulier, si $f \in L^p$ avec $2 < p \leqslant 3$, alors $\Lambda(t) = t^2 + O(|t|^p)$.

Démonstration. Soit $\nu = \tau^\eta$. On va travailler dans $\mathcal{L}_\nu$, ce qui ne pose pas de problème puisque les valeurs propres de $\widehat{T}_t$ dans $\mathcal{L}_\tau$ et $\mathcal{L}_\nu$ coïncident pour t assez petit.

Montrons tout d'abord que, dans $\mathcal{L}_\nu$,

$$\widehat{T}\left(\frac{e^{itf} - 1}{t}\right) = i\widehat{T}(f) + O(t). \tag{4.7}$$

La formule de Taylor donne

$$\frac{e^{itf} - 1}{t} - if = -tf^2 \int_0^1 (1 - u)e^{ituf}\,\mathrm{d}u.$$

On utilise cette formule pour borner $D_\nu\left(\frac{e^{itf}-1}{t} - if\right)(a)$ pour $a \in \alpha$. Comme $D_\nu(h_1h_2) \leqslant$

$D_\nu(h_1)\,\|h_2\|_\infty + \|h_1\|_\infty\, D_\nu(h_2)$, on a

$$D_\nu\left(\frac{e^{itf}-1}{t} - if\right)(a) \leqslant |t|\,\|f_{|a}\|_\infty\, D_\nu f(a) + |t|\,\|f_{|a}\|_\infty^2 \int_0^1 D_\nu(e^{ituf})(a)\,\mathrm{d}u.$$

Soit C tel que $|e^{iw}-1| \leqslant C|w|^\eta$ pour tout $w \in \mathbb{R}$. Alors, pour $x, y \in a$ et $|t| \leqslant 1$, $u \leqslant 1$, on a

$$\begin{aligned}\left|e^{ituf(x)} - e^{ituf(y)}\right| &= \left|e^{itu(f(x)-f(y))} - 1\right| \leqslant C|tu(f(x)-f(y))|^\eta \\ &\leqslant C|D_\tau f(a) d_\tau(x,y)|^\eta = C D_\tau f(a)^\eta d_{\tau^\eta}(x,y).\end{aligned}$$

Ainsi, $D_\nu(e^{ituf})(a) \leqslant C D_\tau f(a)^\eta$. En intégrant sur u, on trouve

$$D_\nu\left(\frac{e^{itf}-1}{t} - if\right)(a) \leqslant C|t|\left(\|f_{|a}\|_\infty\, D_\tau f(a) + \|f_{|a}\|_\infty^2\, D_\tau f(a)^\eta\right).$$

Finalement, (4.6) assure que

$$\sum_{a\in\alpha} m[a] D_\nu\left(\frac{e^{itf}-1}{t} - if\right)(a) = O(t). \tag{4.8}$$

La proposition 1.1.15 (ou plutôt sa preuve) montre que, pour toute fonction u,

$$\|\widehat{T}u\|_{\mathcal{L}_\nu} \leqslant C\left(\sum_{a\in\alpha} m[a] D_\nu u(a) + \int |u|\right).$$

En appliquant cette estimée à $u = \frac{e^{itf}-1}{t} - if$, on obtient donc (4.7) grâce à (4.8) et au fait que

$$\int \left|\frac{e^{itf}-1}{t} - if\right| \leqslant |t| \int f^2 = O(|t|).$$

Reprenons ensuite les différentes étapes de la preuve du théorème 4.1.4. Par (4.7), l'équation (4.5) donne

$$(I - \widehat{T})\frac{\xi_t - \xi_0}{t} = i\widehat{T}(f) + O(t).$$

Toutes les fonctions qui apparaissent dans cette équation sont d'intégrale nulle. Comme $I - \widehat{T}$ admet un inverse continu sur l'espace des fonctions de $\mathcal{L}_\nu$ d'intégrale nulle, on obtient donc

$$\frac{\xi_t - \xi_0}{t} = ia + O(t).$$

Cette estimée est vraie dans $\mathcal{L}_\nu$, et donc dans L^∞. De plus, $e^{itf} = 1 + itf + O(t^2)$ dans L^1 puisque f^2 est intégrable. En multipliant ces estimées, on obtient

$$\int (\widehat{T}_t - \widehat{T})(\xi_t - 1) = \int (e^{itf} - 1)(\xi_t - \xi_0) = -t^2 \int fa + O(t^3).$$

La conclusion découle alors de (4.3). □

4.1.3 Lois stables

On dit qu'une variable aléatoire Z non triviale (i.e. non presque sûrement constante) est stable s'il existe des variables aléatoires indépendantes identiquement distribuées $Z_0, Z_1, \dots$ et des réels A_n et $B_n > 0$ tels que

$$\frac{\sum_{i=0}^{n-1} Z_i - A_n}{B_n} \to Z \text{ en loi.}$$

Dans ce cas, on dit que Z_0 est attirée par Z, ou encore que la loi de Z_0 appartient au domaine d'attraction de Z. Les lois stables sont complètement classifiées (voir par exemple [Fel66]) et dépendent de quatre paramètres. Le plus important est certainement l'indice $p \in]0, 2]$. Le cas $p = 2$ correspond aux lois normales tandis que, pour $p \in]1, 2[$, les lois stables sont dans L^1 mais pas dans L^2. Pour $p < 2$, l'indice est caractérisé par le fait suivant : si Z est une loi stable d'indice p, alors il existe $c_1 \geqslant 0$ et $c_2 \geqslant 0$ avec $c_1 + c_2 > 0$ tels que $P[Z > x] = (c_1 + o(1))x^{-p}$ et $P[Z < -x] = (c_2 + o(1))x^{-p}$. Dans la suite, on ne considérera que des lois stables intégrables, c'est-à-dire avec $1 < p \leqslant 2$.

Pour $p \in]1, 2]$, $c > 0$ et $\beta \in [-1, 1]$, on définit $Z_{p,c,\beta}$ comme étant une variable aléatoire de fonction caractéristique

$$E\left(e^{itZ_{p,c,\beta}}\right) = e^{-c|t|^p\left(1 - i\beta \operatorname{sgn}(t) \tan\left(\frac{p\pi}{2}\right)\right)}. \tag{4.9}$$

C'est une loi stable d'indice p, de moyenne nulle. En fait, toutes les lois stables centrées et d'indice $p \in]1, 2]$ sont de cette forme. Notons que, quand $p = 2$, la valeur de β est sans importance.

Rappelons qu'une fonction $L : (0, \infty) \to \mathbb{R}$ est à variation lente si $\forall \lambda > 0, \lim_{x\to\infty} \frac{L(\lambda x)}{L(x)} = 1$. Une variable aléatoire Z est alors dans le domaine d'attraction d'une loi stable d'indice p si et seulement si elle vérifie les conditions suivantes ([Fel66]) :

- Pour $p \in]1, 2[$, $P[Z > x] = (c_1 + o(1))x^{-p}L(x)$ et $P[Z < -x] = (c_2 + o(1))x^{-p}L(x)$ pour certains réels $c_1, c_2 \geqslant 0$ avec $c_1 + c_2 > 0$ et une fonction à variation lente L.
- Pour $p = 2$, soit $Z \in L^2$, soit $P[|Z| > x] \sim x^{-2}l(x)$ pour une fonction l telle que $L(x) := 2\int_1^x \frac{l(u)}{u}\, du$ soit à variation lente et non bornée.

Notons que, dans le deuxième cas, si la fonction l est à variation lente, alors L est automatiquement à variation lente (mais la réciproque est fausse). Lorsque Z vérifie une de ces conditions et n'appartient pas à L^2, on définit des constantes de normalisation B_n et A_n par

$$nL(B_n) = B_n^p, \quad A_n = nE(Z).$$

Soit $Z_0, Z_1, \dots$ une suite de variables aléatoires indépendantes identiquement distribuées de même distribution que Z. On a alors la convergence

$$\frac{\sum_{i=0}^{n-1} Z_i - A_n}{B_n} \to Z_{p,c,\beta} \tag{4.10}$$

où

$$c = (c_1 + c_2)\Gamma(1-p)\cos\left(\frac{p\pi}{2}\right) \text{ et } \beta = \frac{c_1 - c_2}{c_1 + c_2}$$

dans le cas $p \in]1,2[$, et $c = 1/2$ dans le cas $p = 2$ (la loi limite étant alors la gaussienne $\mathcal{N}(0,1)$).

Dans le cas des applications Gibbs-Markov, [AD01b] a montré que, si une fonction $f : X \to \mathbb{R}$ est hölderienne et a une distribution dans le domaine d'attraction d'une loi stable Z, alors ses sommes de Birkhoff (convenablement renormalisées par des constantes A_n et B_n qui sont les mêmes que dans le cas indépendant) convergent en loi vers Z. On va démontrer le même type de résultat, sous des hypothèses de régularité un peu plus faibles.

Théorème 4.1.7. *Soient $T : X \to X$ une application Gibbs-Markov transitive préservant une mesure de probabilité m, et $0 < \tau < 1$ assez grand pour que la distorsion de T soit τ-Hölder. Soit $f : X \to \mathbb{R}$ intégrable telle que $\sum_{a\in\alpha} m[a]D_\tau f(a) < \infty$ et $\int f = 0$. On suppose de plus que $m[f > x] = (c_1 + o(1))x^{-p}L(x)$ et $m[f < -x] = (c_2 + o(1))x^{-p}L(x)$ pour certaines constantes $c_1, c_2 \geqslant 0$ avec $c_1 + c_2 > 0$, une fonction à variation lente L et un réel $p \in]1,2[$.*

Soient $\widehat{T}_t(u) = \widehat{T}(e^{itf}u)$, et $\lambda(t)$ la valeur propre de $\widehat{T}_t$ pour t petit donnée par le corollaire 4.1.3. Alors cette valeur propre satisfait

$$\lambda(t) = 1 - c|t|^p L(1/|t|) + ic|t|^p L(1/|t|)\beta\,\mathrm{sgn}(t)\tan\left(\frac{p\pi}{2}\right) + o(|t|^p L(1/|t|))$$

pour $\beta = \frac{c_1-c_2}{c_1+c_2}$ et $c = (c_1 + c_2)\Gamma(1-p)\cos\left(\frac{p\pi}{2}\right)$.

Démonstration. Ce théorème est essentiellement le théorème 5.1 de [AD01b], avec un petit affaiblissement des hypothèses puisqu'on ne suppose pas $f \in \mathcal{L}_\tau$. Pour faire fonctionner leur preuve dans ce contexte, il suffit de vérifier que la fonction propre ξ_t de $\widehat{T}_t$ (normalisée pour être d'intégrale 1) vérifie $\|\xi_t - 1\| = O(|t|)$. Mais, comme les fonctions propres dépendent de façon lipschitzienne de l'opérateur, $\|\xi_t - 1\| = O(\|\widehat{T}_t - \widehat{T}\|) = O(|t|)$ par le corollaire 4.1.3. □

On en déduit que les sommes de Birkhoff de f convergent vers une loi stable :

Corollaire 4.1.8. *Sous les hypothèses du théorème 4.1.7, soit B_n défini par $nL(B_n) = B_n^p$. Alors $\frac{S_n f}{B_n}$ converge en loi vers $Z_{p,c,\beta}$ où c et β sont comme dans le théorème 4.1.7.*

Les hypothèses du théorème 4.1.7 sur la distribution de f signifient exactement que, si $Z_0, Z_1, \ldots$ sont des variables aléatoires indépendantes de même distribution que f, alors $\frac{Z_0+\ldots+Z_{n-1}}{B_n}$ converge en loi vers $Z_{p,c,\beta}$, d'après (4.10). Ainsi, $f, f\circ T, f\circ T^2, \ldots$ se comportent exactement comme $Z_0, Z_1, Z_2, \ldots$, et les paramètres de la loi limite ne dépendent que de la distribution de f. C'est très différent du cas du théorème central limite, puisque dans ce cas la variance σ^2 de la loi limite ne dépend pas que de la distribution de f mais aussi des interactions entre les différentes itérations de T : lorsque T est Markov-mélangeante, on vérifie que $\sigma^2 = \int f^2 + 2\sum_{k\geqslant 1}\int f\cdot f\circ T^k$.

Démonstration du corollaire 4.1.8. On raisonne comme dans la preuve du corollaire 4.1.5. Il suffit alors de montrer que $\lambda(t/B_n)^n$ tend vers la valeur en t de la fonction caractéristique de $Z_{p,c,\beta}$, donnée par (4.9).

Fixons $t > 0$. Alors, d'après le théorème 4.1.7,

$$n \log \lambda(t/B_n) \sim n \left[-c \left(\frac{t}{B_n} \right)^p L\left(\frac{B_n}{t} \right) \right] \left[1 - i\beta \tan\left(\frac{p\pi}{2} \right) \right].$$

Comme L est à variation lente, $L(B_n/t) \sim L(B_n)$, donc $nL(B_n/t)/B_n^p \to 1$ par définition de B_n. Cela montre que

$$n \log \lambda(t/B_n) \to -ct^p \left[1 - i\beta \tan\left(\frac{p\pi}{2} \right) \right].$$

On obtient le même type de convergence pour $t < 0$, ce qui conclut la preuve. □

Dans le cas où la distribution de f est dans le domaine d'attraction d'une gaussienne, sans être dans L^2, on démontre exactement de la même manière le même type de résultat, en utilisant [AD01a] au lieu de [AD01b] :

Théorème 4.1.9. *Soient $T : X \to X$ une application Gibbs-Markov transitive préservant une mesure de probabilité m, et $0 < \tau < 1$ assez grand pour que la distorsion de T soit τ-Hölder. Soit $f : X \to \mathbb{R}$ intégrable telle que $\sum_{a\in\alpha} m[a] D_\tau f(a) < \infty$ et $\int f = 0$. On suppose de plus que $m[|f| > x] \sim x^{-2} l(x)$ où la fonction l est telle que $L(x) := 2\int_1^x \frac{l(u)}{u}\,du$ soit non bornée et à variation lente.*

Soient $\widehat{T}_t(u) = \widehat{T}(e^{itf}u)$, et $\lambda(t)$ la valeur propre de $\widehat{T}_t$ pour t petit donnée par le corollaire 4.1.3. Alors cette valeur propre satisfait $\lambda(t) = 1 - \frac{t^2}{2} L(1/|t|) + o(t^2 L(1/|t|))$.

En particulier, soit B_n défini par $nL(B_n) = B_n^2$. Alors $\frac{S_n f}{B_n}$ converge en loi vers $\mathcal{N}(0,1)$.

4.2 Tours localement Gibbs-Markov : approche élémentaire

Dans cette partie, on va montrer comment on peut obtenir des théorèmes limites pour des tours localement Gibbs-Markov à partir du même type de résultat pour les applications Gibbs-Markov. On va pour cela établir un résultat abstrait très général reliant les théorèmes limites pour une application et pour une application induite sur une partie plus petite de l'espace. On verra ensuite comment appliquer ce théorème aux tours localement Gibbs-Markov.

4.2.1 Théorèmes limites et induction

Dans ce paragraphe, on établit un lien entre les théorèmes limites pour une application et pour une application induite. Ce type de méthode élémentaire et directe a été initié dans le

cas des flots par Melbourne et Török dans [MT04].

Si Y est une partie de mesure non nulle d'un espace probabilisé (X, m), $T : X \to X$ est une application mesurable préservant m et T_Y est l'application induite par T sur Y, on écrira $S_n^Y u = \sum_{k=0}^{n-1} u \circ T_Y^k$: c'est la somme de Birkhoff de u pour la transformation T_Y. On notera aussi $E_Y(u) = \frac{\int_Y u}{m[Y]}$.

Théorème 4.2.1. *Soient $T : X \to X$ une application ergodique préservant une mesure de probabilité m et $f : X \to \mathbb{R}$ une fonction intégrable d'intégrale nulle. Soit Y une partie de X de mesure non nulle. Pour $y \in Y$, soient $\varphi(y) = \inf\{n > 0 \mid T^n(y) \in Y\}$ et $f_Y(y) = \sum_{k=0}^{\varphi(y)-1} f(T^k y)$.*

On suppose que les propriétés suivantes sont vérifiées :

1. *Il existe une suite $B_n \to +\infty$ avec $\inf_{r \geqslant n} \frac{B_r}{B_n} > 0$ et une variable aléatoire Z telles que $\frac{S^Y_{\lfloor nm(Y) \rfloor} f_Y}{B_n}$ converge en loi vers Z, i.e., pour tout $t \in \mathbb{R}$,*

$$E_Y\left(e^{it\frac{S^Y_{\lfloor nm(Y)\rfloor} f_Y}{B_n}}\right) \to E\left(e^{itZ}\right). \tag{4.11}$$

2. *Il existe $b > 0$ tel que, dans l'extension naturelle de T_Y, $\frac{1}{N^b}\sum_0^{N-1} f_Y(T_Y^k y)$ tend presque partout vers 0 quand $N \to \pm\infty$.*
3. *Pour tout $\varepsilon > 0$, il existe $A > 0$ et N_0 tels que, pour tout $n \geqslant N_0$,*

$$m\left\{y \in Y \mid \left|\frac{S_n^Y \varphi - nE_Y(\varphi)}{B_n^{1/b}}\right| \geqslant A\right\} \leqslant \varepsilon. \tag{4.12}$$

Alors la fonction f satisfait également un théorème limite :

$$E\left(e^{it\frac{S_n f}{B_n}}\right) \to E(e^{itZ}),$$

i.e. $\frac{S_n f}{B_n}$ tend en loi vers Z.

Les hypothèses du théorème sont taillées sur mesure pour que la preuve qui va suivre marche, mais elles sont en fait souvent satisfaites dans des cas naturels :

1. La convergence (4.11) est très souvent satisfaite dès que T_Y est suffisamment dilatante, par exemple localement Gibbs-Markov, comme on l'a vu dans la section précédente.
2. Il est utile de considérer l'extension naturelle pour pouvoir faire tendre N vers $-\infty$ et considérer T_Y^{-1} dans la preuve. En toute généralité, le théorème de Birkhoff implique que cette hypothèse est satisfaire pour $b = 1$, et c'est souvent suffisant. Parfois, on a cependant besoin de meilleures estimées. On peut par exemple utiliser alors [Kac96, théorème 16] : si les corrélations de $f_Y \in L^2$ décroissent au moins en $O(1/n)$, alors l'hypothèse est satisfaite pour tout $b > 1/2$ (pour $N \to -\infty$, on utilise le fait que $\int f_Y \cdot f_Y \circ T_Y^n = \int f_Y \circ T_Y^{-n} \cdot f_Y$, et on applique le théorème à T_Y^{-1}).

3. La troisième hypothèse est plus faible que :

$$\exists B'_n = O(B_n^{1/b}) \text{ tel que } \frac{S_n^Y \varphi - nE_Y(\varphi)}{B'_n} \text{ converge en distribution.} \qquad (4.12')$$

De plus, φ est souvent plus simple que f_Y. Comme f_Y satisfait un théorème limite d'après la première hypothèse, c'est aussi le cas de φ, et on obtient (4.12'). Ainsi, (4.12') – et donc (4.12) – sont très souvent satisfaits.

Pour démontrer le théorème, on aura besoin du résultat abstrait suivant, dû à Eagleson ([Eag76], voir aussi [Aar97, Proposition 3.6.1]) et dont l'importance a été soulignée par Roland Zweimüller.

Théorème 4.2.2. *Soient (X, m) un espace probabilisé et $T : X \to X$ une application ergodique préservant m. Soit $f : X \to \mathbb{R}$ mesurable telle que $\frac{S_n f}{B_n}$ converge en loi vers une loi Z (pour la mesure m), pour une suite B_n qui tend vers $+\infty$.*

Alors, pour toute fonction $g : \mathbb{R} \to \mathbb{R}$ uniformément continue bornée, pour toute fonction $\varphi \in L^1(m)$, on a

$$\int_X \varphi \cdot g\left(\frac{S_n f}{B_n}\right) \to \left(\int_X \varphi\right) E(g(Z)).$$

Autrement dit, pour toute mesure de probabilité μ absolument continue par rapport à m, $\frac{S_n f}{B_n}$ tend encore en loi vers Z pour μ.

La dernière affirmation du théorème s'obtient en prenant $\varphi = \frac{\mathrm{d}\mu}{\mathrm{d}m}$ et $g(x) = e^{itx}$: on obtient que la fonction caractéristique de $\frac{S_n f}{B_n}$ (pour la mesure μ) converge vers la fonction caractéristique de Z.

Ce théorème signifie qu'un système dynamique ergodique vérifiant un théorème limite vérifie automatiquement un théorème limite "mélangeant au sens de Rényi".

Démonstration. On doit montrer que $g_n = g\left(\frac{S_n f}{B_n}\right)$ converge, pour la topologie faible $*$ sur $L^\infty(X)$, vers la fonction constante égale à $E(g(Z))$.

Soit ψ une valeur d'adhérence de g_n, limite d'une suite $g_{j(n)}$. Alors $\int 1 \cdot g_{j(n)} \to \int 1 \cdot \psi$, i.e., $\int \psi = E(g(Z))$. Si on montre que $\psi = \psi \circ T$ presque partout, alors ψ sera presque partout constante par ergodicité, puis $\psi = E(g(Z))$, ce qui conclura.

Soit $\varepsilon > 0$. Comme f est mesurable, $m(|f| \geqslant \varepsilon B_{j(n)}) \to 0$. Quitte à extraire une sous-suite, on peut supposer que $\sum m(|f| \geqslant \varepsilon B_{j(n)}) < \infty$. Par procédé diagonal, on peut même le garantir pour tout $\varepsilon > 0$.

La suite $g_{j(n)}$ étant bornée dans L^1, on peut d'après le théorème de Komlos ([Kom67]) extraire encore plus de telle sorte que $\frac{1}{n}\sum_{k=0}^{n-1} g_{j(k)}$ converge presque partout vers une fonction $\widetilde{\psi}$. Comme $\int_A g_{j(k)} \to \int_A \psi$ d'une part, $\int_A \left(\frac{1}{n}\sum_{k=0}^{n-1} g_{j(k)}\right) \to \int_A \widetilde{\psi}$ par le théorème de convergence dominée d'autre part, on obtient $\psi = \widetilde{\psi}$ presque partout.

Comme m est invariante par T, on a $\sum m(|f \circ T^{j(n)}| \geqslant \varepsilon B_{j(n)}) < \infty$. Ainsi, presque tout x satisfait $\frac{|f(T^{j(n)}x)|}{B_{j(n)}} < \varepsilon$ pour n assez grand. Comme

$$\frac{S_{j(n)}f \circ T(x) - S_{j(n)}f(x)}{B_{j(n)}} = \frac{f \circ T^{j(n)}(x) - f(x)}{B_{j(n)}},$$

on en déduit que $\frac{S_{j(n)}f\circ T(x)-S_{j(n)}f(x)}{B_{j(n)}} \to 0$ presque partout. Comme g est uniformément continue, cela implique que $g_{j(n)}(Tx) - g_{j(n)}(x) \to 0$. Puis

$$\frac{1}{n}\sum_{k=0}^{n-1} g_{j(k)}(Tx) - \frac{1}{n}\sum_{k=0}^{n-1} g_{j(k)}(x) \to 0.$$

Mais ce terme converge simplement vers $\psi(Tx)-\psi(x)$. Ainsi, $\psi(Tx) = \psi(x)$ presque partout, ce qui conclut la preuve. □

Démonstration du théorème 4.2.1. Le résultat est trivial si $m(Y) = 1$, et on peut donc supposer $m(Y) < 1$.

Sans perte de généralité, on peut travailler dans une tour, c'est-à-dire supposer que $X = \{(y, i) \mid y \in Y, i \in \{0, \dots, \varphi(y) - 1\}\}$ et que, pour $i < \varphi(y) - 1$, $T(y, i) = (y, i + 1)$, tandis que $T(y, \varphi(y) - 1) = (T_Y(y), 0)$. En effet, d'après le paragraphe 1.3.2 (et plus précisément la proposition 1.3.7), on peut construire une extension de X satisfaisant ces propriétés, et il est équivalent de démontrer un théorème limite dans X ou dans cette extension. Notons que $E_Y(\varphi) = 1/m(Y)$ par la formule de Kac. Soit π la projection de X dans Y donnée par $\pi(y, i) = y$.

Dans cette preuve, on écrira $S_t f(x)$ en lieu de $S_{\lfloor t \rfloor} f(x)$ lorsque t n'est pas entier. De la même façon, T^t signifiera $T^{\lfloor t \rfloor}$. On étend aussi B_n à $\mathbb{R}_+$ en posant $B_t := B_{\lfloor t \rfloor}$.

Comme T est ergodique, T_Y est également ergodique d'après la proposition 1.3.3. Le théorème de Birkhoff donne

$$S_n^Y \varphi = \frac{n}{m(Y)} + o(n) \tag{4.13}$$

presque partout sur Y. Pour $y \in Y$ et $N \in \mathbb{N}$, soit $n(y, N)$ le plus grand entier tel que $S_n^Y \varphi(y) < N$. Si y satisfait $S_n^Y \varphi(y) = \frac{n}{m(Y)} + o(n)$, ce qui est vrai presque partout, alors $n(y, N)$ est fini pour tout N et $\frac{n(y,N)}{m(Y)} \sim N$, i.e.

$$\frac{n(y, N)}{Nm(Y)} \to 1. \tag{4.14}$$

Grâce à (4.11), on peut appliquer le théorème 4.2.2 dans Y à la transformation ergodique T_Y, et obtenir que $\int_Y \varphi e^{itS^Y_{Nm(Y)} f_Y / B_N} \to E(e^{itZ})$. Comme $\int_X e^{it(S_N^Y f_Y)\circ\pi} = \int_Y \varphi e^{itS_N^Y f_Y}$, on obtient

que

$$\frac{(S^Y_{Nm(Y)} f_Y) \circ \pi}{B_N} \to Z \tag{4.15}$$

en loi sur X. L'idée de la preuve va être de montrer que $(S^Y_{Nm(Y)} f_Y) \circ \pi$ et $S_N f$ sont proches, ce qui n'est pas surprenant puisqu'une itération de T_Y correspond approximativement à $1/m(Y)$ itérations de T. Cela donnera la convergence de $\frac{S_N f}{B_N}$ vers Z.

On écrit

$$\begin{aligned} S_N f(y,i) = (S_N f(y,i) - S_N f(y,0)) + \big(S_N f(y,0) - S^Y_{n(y,N)} f_Y(y)\big) \\ + \big(S^Y_{n(y,N)} f_Y(y) - S^Y_{Nm(Y)} f_Y(y)\big) + S^Y_{Nm(Y)} f_Y(y). \end{aligned}$$

Le dernier terme, égal à $\big(S^Y_{Nm(Y)} f_Y\big) \circ \pi$, satisfait un théorème limite par (4.15). Pour conclure, on va voir que les trois autres termes divisés par B_N tendent vers 0 en probabilité.

Les deuxième et troisième termes ne dépendent que de y ; le lemme suivant va donc être utile pour montrer qu'ils tendent vers 0 en loi sur X.

Lemme 4.2.3. *Soit f_n une suite de fonctions sur Y tendant vers 0 en probabilité sur Y. Alors $f_n \circ \pi$ tend vers 0 en probabilité sur X.*

Démonstration. Soit $\varepsilon > 0$. Comme $f_n \to 0$ en probabilité, la mesure de $E_n := \{y \in Y \mid |f_n(y)| \geqslant \varepsilon\}$ tend vers 0. Comme $\varphi \in L^1$, le théorème de convergence dominée implique que $\int_{E_n} \varphi \to 0$, i.e. la mesure de $\pi^{-1}(E_n)$ tend vers 0. Mais $\pi^{-1}(E_n)$ est exactement l'ensemble sur lequel $|f_n \circ \pi| \geqslant \varepsilon$. □

Fait : $\frac{1}{B_N}(S_N f(y,i) - S_N f(y,0))$ tend en probabilité vers 0 sur X.

Démonstration. Soit $V_N(y) = \sum_{i=0}^{\varphi(y)-1} |f \circ T^N(y,i)|$. Alors $\|V_N\|_{L^1(Y)} = \big\|f \circ T^N\big\|_{L^1(X)} = \|f\|_{L^1(X)}$ comme T préserve la mesure. Ainsi, V_N/B_N tend vers 0 dans $L^1(Y)$, et en probabilité. Le lemme 4.2.3 montre donc que $\frac{1}{B_N} V_N \circ \pi$ tend vers 0 en probabilité sur X.

Comme $S_N f(y,i) - S_N f(y,0) = \sum_N^{N+i-1} f(T^k(y,0)) - \sum_0^{i-1} f(T^k(y,0))$, on a $|S_N f(y,i) - S_N f(y,0)| \leqslant V_N(y) + V_0(y)$. Ainsi, $\frac{1}{B_N}(S_N f(y,i) - S_N f(y,0))$ est borné par une fonction qui tend vers 0 en probabilité. ◊

Fait : $\frac{1}{B_N}\Big(S_N f(y,0) - S^Y_{n(y,N)} f_Y(y)\Big)$ tend vers 0 en probabilité sur X.

Démonstration. D'après le lemme 4.2.3, il suffit de le démontrer sur Y. Soit $H(y,i) = \Big|\sum_{j=0}^{i-1} f(y,j)\Big|$. Alors

$$\big|S_N f(y,0) - S^Y_{n(y,N)} f_Y(y)\big| = H \circ T^N(y,0).$$

Comme T préserve la mesure m, on a pour tout $a > 0$

$$m\left\{y \in Y \mid \frac{1}{B_N} H \circ T^N(y,0) \geqslant a\right\} \leqslant m\left\{x \in X \mid \frac{1}{B_N} H \circ T^N(x) \geqslant a\right\} = m\left\{x \in X \mid \frac{1}{B_N} H(x) \geqslant a\right\}.$$

Comme H est mesurable et $B_N \to \infty$, cette mesure tend vers 0 quand $N \to \infty$. ◊

Fait : $\frac{1}{B_N}\left(S^Y_{n(y,N)} f_Y - S^Y_{Nm(Y)} f_Y\right)$ *tend vers* 0 *en probabilité sur* X.

Démonstration. D'après le lemme 4.2.3, il suffit de le démontrer sur Y. Sans perte de généralité, on peut passer à l'extension naturelle et supposer que T_Y est inversible.
Pour $n < 0$, soit $S^Y_n f_Y = \sum_1^{|n|} f_Y \circ T_Y^{-j}$. En posant $\nu(y,N) = n(y,N) - Nm(Y)$, on obtient

$$S^Y_{n(y,N)} f_Y(y) - S^Y_{Nm(Y)} f_Y(y) = S^Y_{\nu(y,N)} f_Y\left(T_Y^{Nm(Y)}(y)\right). \tag{4.16}$$

Soient $A > 0$ et $N \in \mathbb{N}$, on va estimer la mesure de $\{y \mid \nu(y,N) \geqslant AB_N^{1/b}\}$. Soit $\alpha > 0$ tel que $m(Y) + \alpha < 1$. Supposons tout d'abord que $AB_N^{1/b} > \alpha N$. Alors

$$\{y \mid \nu(y,N) \geqslant AB_N^{1/b}\} \subset \{y \mid n(y,N) \geqslant (m(Y)+\alpha)N\}.$$

D'après l'équation (4.14), la mesure de cet ensemble tend vers 0. Supposons ensuite que $AB_N^{1/b} \leqslant \alpha N$. Alors, comme $E_Y(\varphi) = 1/m(Y)$,

$$\{y \mid \nu(y,N) \geqslant AB_N^{1/b}\} = \{n(y,N) \geqslant AB_N^{1/b} + Nm(Y)\} = \{S^Y_{AB_N^{1/b}+Nm(Y)}\varphi < N\}$$
$$= \left\{\frac{S^Y_{AB_N^{1/b}+Nm(Y)}\varphi - (AB_N^{1/b} + Nm(Y))E_Y(\varphi)}{\left(B_{AB_N^{1/b}+Nm(Y)}\right)^{1/b}} < -\frac{A}{m(Y)}\left(\frac{B_N}{B_{AB_N^{1/b}+Nm(Y)}}\right)^{1/b}\right\}.$$

Mais $AB_N^{1/b} + Nm(Y) \leqslant (m(Y)+\alpha)N \leqslant N$. Par hypothèse, il existe une constante $c > 0$ telle que, pour tous $n \leqslant r$, $\frac{B_r}{B_n} \geqslant c$. En particulier, $\frac{B_N}{B_{AB_N^{1/b}+Nm(Y)}} \geqslant c$. Ainsi,

$$\{y \mid \nu(y,N) \geqslant AB_N^{1/b}\} \subset \left\{\frac{S^Y_{AB_N^{1/b}+Nm(Y)}\varphi - (AB_N^{1/b} + Nm(Y))E_Y(\varphi)}{\left(B_{AB_N^{1/b}+Nm(Y)}\right)^{1/b}} < -\frac{Ac^{1/b}}{m(Y)}\right\}.$$

Par conséquent, si A est assez grand, l'hypothèse 3 implique que $m\{y \mid \nu(y,N) \geqslant AB_N^{1/b}\} \leqslant \varepsilon$ pour N assez grand.

Démontrons le même résultat pour $m\{y \mid \nu(y,N) \leqslant -AB_N^{1/b}\}$. Si A est assez grand, il existe d'après l'hypothèse 3 un entier K tel que, pour tout $n \geqslant K$,

$$m\left\{y \mid \frac{S_n^Y\varphi(y) - nE_Y(\varphi)}{B_n^{1/b}} > \frac{Ac^{1/b}}{m(Y)}\right\} \leqslant \varepsilon.$$

Soit $N \in \mathbb{N}$. On a toujours $\nu(y,N) \geqslant -Nm(Y)$ donc l'ensemble $\{y \mid \nu(y,N) \leqslant -AB_N^{1/b}\}$ ne peut être non vide que si $AB_N^{1/b} \leqslant Nm(Y)$, et alors

$$\begin{aligned}&\{y \mid \nu(y,N) \leqslant -AB_N^{1/b}\}\\ &\quad= \left\{\frac{S^Y_{-AB_N^{1/b}+Nm(Y)}\varphi - (-AB_N^{1/b}+Nm(Y))E_Y(\varphi)}{\left(B_{-AB_N^{1/b}+Nm(Y)}\right)^{1/b}} \geqslant \frac{A}{m(Y)}\left(\frac{B_N}{B_{-AB_N^{1/b}+Nm(Y)}}\right)^{1/b}\right\}.\end{aligned}$$

Supposons tout d'abord que $-AB_N^{1/b} + Nm(Y) \geqslant K$. Comme $\frac{B_N}{B_{-AB_N^{1/b}+Nm(Y)}} \geqslant c > 0$, on obtient dans ce cas

$$\{y \mid \nu(y,N) \leqslant -AB_N^{1/b}\} \subset \left\{\frac{S^Y_{-AB_N^{1/b}+Nm(Y)}\varphi - (-AB_N^{1/b}+Nm(Y))E_Y(\varphi)}{\left(B_{-AB_N^{1/b}+Nm(Y)}\right)^{1/b}} \geqslant \frac{Ac^{1/b}}{m(Y)}\right\},$$

ce dernier ensemble étant de mesure au plus ε par définition de K. Supposons ensuite que $-AB_N^{1/b} + Nm(Y) < K$. Alors

$$\{y \mid \nu(y,N) \leqslant -AB_N^{1/b}\} \subset \bigcup_{i=0}^{K-1}\left\{y \mid \frac{S_i^Y\varphi(y) - iE_Y(\varphi)}{(B_i)^{1/b}} \geqslant \frac{A}{m(Y)}\left(\frac{B_N}{B_i}\right)^{1/b}\right\}.$$

Comme $B_N \to \infty$ tandis que i reste borné, la mesure de cet ensemble tend donc vers 0 quand $N \to \infty$.

On a donc démontré :

$$\forall \varepsilon > 0, \exists A > 0, \exists N_0 > 0, \forall N \geqslant N_0,\ m\{y \mid |\nu(y,N)| \geqslant AB_N^{1/b}\} \leqslant \varepsilon. \tag{4.17}$$

Soit $W_N(y) = \frac{1}{B_N} S_{\nu(y,N)} f_Y\left(T_Y^{Nm(Y)}(y)\right)$, on va montrer qu'il tend vers 0 en distribution. Cela conclura la preuve d'après (4.16). Soit $a > 0$, montrons que $m(|W_N| > a) \to 0$ quand $N \to \infty$.

Soit $\varepsilon > 0$. L'hypothèse 2 implique qu'il existe $\widetilde{Y}$ avec $m(\widetilde{Y}) \geqslant m(Y) - \varepsilon$ et N_1 tels que, pour tout $|N| \geqslant N_1$, $\frac{1}{|N|^b}|S_N^Y f_Y| \leqslant \varepsilon$ sur $\widetilde{Y}$. Soient $Y'_N = \{y \in Y \mid |\nu(y,N)| < N_1\}$ et $Y''_N = \{y \in Y \mid |\nu(y,N)| \geqslant N_1\}$. On estime tout d'abord la contribution de Y'_N.

Soit $\psi(y) = \sum_{-N_1}^{N_1-1} |f_Y \circ T_Y^j(y)|$. Comme ψ est mesurable, il existe une constante C et un ensemble Z de Y avec $m(Z) \geqslant m(Y) - \varepsilon$ et $|\psi| \leqslant C$ sur Z. Alors, pour $y \in Y'_N$, on a $|W_N(y)| \leqslant \frac{1}{B_N}\psi\left(T_Y^{Nm(Y)}y\right)$. Soit $Z_N = Y'_N \cap T_Y^{-Nm(Y)}(Z)$: il satisfait $m(Z_N) \geqslant m(Y'_N) - \varepsilon$. Sur Z_N, on a $|W_N| \leqslant \frac{C}{B_N}$, si bien que, pour N assez grand, $|W_N| < a$ sur Z_N. Ainsi, pour N assez grand,

$$m\{y \in Y'_N \mid |W_N(y)| \geqslant a\} \leqslant m\{y \in Z_N \mid |W_N(y)| \geqslant a\} + \varepsilon = \varepsilon.$$

Estimons maintenant la contribution de Y''_N. Soit $\widetilde{Y}''_N = Y''_N \cap T_Y^{-Nm(Y)}(\widetilde{Y})$, avec $m(\widetilde{Y}''_N) \geqslant m(Y''_N) - \varepsilon$. Alors

$$m(|W_N| \geqslant a) \leqslant m\{y \in \widetilde{Y}''_N \mid |W_N(y)| \geqslant a\} + 2\varepsilon.$$

Sur $\widetilde{Y}''_N$, $|\nu(y,N)| \geqslant N_1$ donc $\frac{1}{|\nu(y,N)|^b}\left|S^Y_{\nu(y,N)}f_Y\left(T_Y^{Nm(Y)}y\right)\right| \leqslant \varepsilon$. Ainsi, $|W_N(y)| \leqslant \varepsilon\frac{|\nu(y,N)|^b}{B_N} = \varepsilon\left(\frac{|\nu(y,N)|}{B_N^{1/b}}\right)^b$. Par conséquent,

$$m(|W_N| \geqslant a) \leqslant m\left(\frac{|\nu(y,N)|}{B_N^{1/b}} \geqslant \left(\frac{a}{\varepsilon}\right)^{1/b}\right) + 2\varepsilon. \tag{4.18}$$

Finalement, (4.17) et (4.18) impliquent que $m(|W_N| \geqslant a)$ tend vers 0. ◊

Les trois faits qu'on a démontrés impliquent que $\frac{S_N f(y,i)}{B_N} - \frac{S^Y_{Nm(Y)}f_Y(y)}{B_N} \to 0$ en loi sur X. Comme $\frac{S^Y_{Nm(Y)}f_Y(y)}{B_N} \to Z$ en loi sur X, par (4.15), cela termine la démonstration du théorème 4.2.1. □

4.2.2 Application aux tours localement Gibbs-Markov

Théorème 4.2.4. *Soit $T : X \to X$ une tour localement Gibbs-Markov transitive de base Y, préservant une mesure de probabilité m. On notera $\varphi : Y \to \mathbb{N}^*$ le temps de retour à Y et α_0 la partition induite sur Y par la partition de Markov de T. Soit $f : X \to \mathbb{R}$ une fonction intégrable de moyenne nulle telle que $f_Y(y) := \sum_{k=0}^{\varphi(y)-1} f(T^k y)$ satisfait*

$$\sum_{a\in\alpha_0} m(a)Df_Y(a) < \infty. \tag{4.19}$$

Alors :

1. *Supposons que $f_Y \in L^2$. Supposons aussi que φ satisfait l'une des propriétés suivantes :*
   - *$\varphi \in L^2$.*
   - *$m(\varphi > x) = x^{-p}L(x)$ où L est à variation lente et $p \in]1,2]$.*

Alors il existe $\sigma^2 \geqslant 0$ *tel que* $\frac{1}{\sqrt{n}} S_n f \to \mathcal{N}(0, \sigma^2)$.

2. *Supposons que* $m(|f_Y| > x) = x^{-2} l(x)$, *avec* $L(x) := 2\int_1^x \frac{l(u)}{u}\,\mathrm{d}u$ *à variation lente et non bornée. Supposons de plus que* $m(\varphi > x) = (c + o(1))x^{-2} l(x)$ *avec* $c > 0$. *Soit* $B_n \to +\infty$ *telle que* $nL(B_n) = B_n^2$. *Alors* $\frac{1}{B_n} S_n f \to \mathcal{N}(0, 1)$.
3. *Supposons que* $m(f_Y > x) = (c_1 + o(1))x^{-p} L(x)$ *et* $m(f_Y < -x) = (c_2 + o(1))x^{-p} L(x)$, *où* L *est une fonction à variation lente,* $p \in]1, 2[$, *et* $c_1, c_2 \geqslant 0$ *satisfont* $c_1 + c_2 > 0$. *Supposons aussi que* $m(\varphi > x) = (c_3 + o(1))x^{-p} L(x)$ *avec* $c_3 > 0$. *Soit* $B_n \to +\infty$ *telle que* $nL(B_n) = B_n^p$. *Alors* $\frac{1}{B_n} S_n f \to Z_{p,c,\beta}$ *où* $c = (c_1 + c_2)\Gamma(1 - p)\cos\left(\frac{p\pi}{2}\right)$ *et* $\beta = \frac{c_1 - c_2}{c_1 + c_2}$.

Notons que les hypothèses du théorème sur f_Y correspondent exactement au fait que la distribution de f_Y est dans le domaine d'attraction d'une loi normale ou stable, d'indice respectivement 2 ou p. Plus précisément, comme Y n'est pas de mesure pleine, la distribution μ sur $\mathbb{R}$ donnée par $\mu(A) = m(f_Y^{-1}(A))$ n'est pas une mesure de probabilité, mais la distribution $\widetilde{\mu} = \mu + (1 - m(Y))\delta_0$ est de masse 1. Alors le théorème affirme que, si cette distribution (qu'on appellera encore, par abus de langage, la distribution de f_Y) est dans le domaine d'attraction d'une loi stable de paramètre $p \in]1, 2]$, et que d'autres hypothèses techniques sont vérifiées, alors les sommes de Birkhoff de f vérifient un théorème limite avec la même normalisation que dans le cas aléatoire.

Démonstration. On va utiliser le théorème 4.2.1 : il faut vérifier ses trois hypothèses. On utilisera les notations de ce théorème, et on écrira en particulier $E_Y(u) = \frac{\int_Y u\,\mathrm{d}m}{m(Y)}$.

On traite tout d'abord le troisième cas, c'est-à-dire le cas des lois stables, en utilisant les résultats du théorème 4.1.7 appliqué à l'application Gibbs-Markov transitive T_Y. Notons qu'on peut sans perte de généralité remplacer d par une distance d_τ pour τ assez proche de 1 pour être dans le cadre de ce théorème. Pour avoir une mesure de probabilité sur Y, il faut aussi remplacer $m_{|Y}$ par $\frac{m_{|Y}}{m(Y)}$, ce qui change les normalisations : on obtient que la valeur propre $\lambda(t)$ de $\widehat{T}_{Y,t}$ pour t petit vérifie

$$\lambda(t) = e^{-\frac{c}{m(Y)}|t|^p\left(1 - i\beta\,\mathrm{sgn}(t)\tan\left(\frac{p\pi}{2}\right)\right)L(|t|^{-1})(1+o(1))}.$$

Le corollaire 4.1.8 implique alors la convergence en loi

$$E_Y\left(e^{i\frac{t}{B_n}S^Y_{\lfloor nm(Y)\rfloor} f_Y}\right) \to E(e^{itZ}) \tag{4.20}$$

où la variable aléatoire $Z = Z_{p,c,\beta}$ est comme dans l'énoncé du théorème. Cela démontre (4.11). De plus, l'égalité $nL(B_n) = B_n^p$ permet de démontrer que $\inf_{r \geqslant n} \frac{B_r}{B_n} > 0$ en utilisant les bornes de Potter sur les fonctions à variation lente ([BGT87, théorème 1.5.6]).

Le théorème de Birkhoff assure que l'hypothèse 2 du théorème 4.2.1 est satisfaite pour $b = 1$, car T_Y est ergodique puisque T l'est. Finalement, l'hypothèse sur la distribution de φ assure,

d'après le corollaire 4.1.8, que $\frac{S^Y_{\lfloor nm(Y)\rfloor}\varphi - nm(Y)E_Y(\varphi)}{B_n}$ converge en loi. Ainsi, (4.12') est satisfait. On peut donc utiliser le théorème 4.2.1, qui donne $\frac{S_n f}{B_n} \to Z$.

La preuve du deuxième cas du théorème 4.2.4 est exactement la même, en utilisant le théorème 4.1.9 au lieu du théorème 4.1.7 pour montrer la convergence en loi de $\frac{S^Y_{\lfloor nm(Y)\rfloor} f_Y}{B_n}$ et $\frac{S^Y_{\lfloor nm(Y)\rfloor}\varphi - nm(Y)E_Y(\varphi)}{B_n}$.

Dans le premier cas, où $f_Y \in L^2$, la preuve est encore identique lorsque $\varphi \in L^2$, avec $B_n = \sqrt{n}$, en utilisant le théorème 4.1.4. Cependant, lorsque $m(\varphi > x) = x^{-p}L(x)$, on doit vérifier différemment les hypothèses 2 et 3 du théorème 4.2.1. Si B'_n est donné par

$$nL(B'_n) = (B'_n)^p, \tag{4.21}$$

alors $\frac{S^Y_n\varphi - nE_Y(\varphi)}{B'_n}$ converge en loi, d'après le théorème 4.1.7. De plus, $\widehat{T}_Y f_Y \in \mathcal{L}$ et est d'intégrale nulle. Si T_Y est Markov-mélangeante, alors $\widehat{T}^n_Y f_Y \to 0$ exponentiellement vite d'après le corollaire 1.1.21. En particulier, il existe $0 < \rho < 1$ tel que $\int f_Y \circ T^n_Y \cdot f_Y = \int (\widehat{T}^n_Y f_Y)\cdot f_Y = O(\rho^n)$. Comme $f_Y \in L^2$, [Kac96, théorème 16] implique que, pour tout $b > 1/2$, $\frac{1}{N^b}\sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0$ presque partout quand $N \to \infty$. Dans l'extension naturelle, $\int f_Y \circ T^{-n}_Y \cdot f_Y = \int f_Y \cdot f_Y \circ T^n_Y$ décroît aussi exponentiellement vite, donc le même argument donne que $\frac{1}{|N|^b}\sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0$ quand $N \to -\infty$. Ainsi, l'hypothèse 2 du théorème 4.2.1 est satisfaite pour tout $b > 1/2$ si T_Y est Markov-mélangeante. Dans le cas général, il existe d'après le corollaire 1.1.19 une décomposition $Y = Y_1 \cup \ldots \cup Y_d$ telle que T_Y envoie Y_i sur Y_{i+1} pour $1 \leqslant i \leqslant d-1$, et Y_d sur Y_1, et telle que T^d_Y soit mélangeante sur chaque Y_i. Par invariance de la mesure, $m(Y_i) = 1/d$ pour $1 \leqslant i \leqslant d$. Posons $u = d\sum_{i=1}^d \left(\int_{Y_i} f_Y\right) 1_{Y_i}$ et $v = f_Y - g$. Comme $\int_{Y_i} v = 0$ pour tout i, on peut utiliser les mêmes arguments que ci-dessus sur chaque Y_i pour obtenir $\widehat{T}^{dn}_Y v = O(\rho^{dn})$. On en déduit que $\widehat{T}^n_Y v = O(\rho^n)$. En particulier, $\frac{1}{|N|^b}\sum_{k=0}^{N-1} v(T^k_Y x) \to 0$ presque partout quand $N \to \pm\infty$, comme ci-dessus. De plus, $\sum_{i=1}^d \left(\int_{Y_i} f_Y\right) = \int f_Y = 0$, donc $\sum_{k=0}^{N-1} u(T^k_Y x)$ est uniformément borné. Cela implique que $\frac{1}{|N|^b}\sum_{k=0}^{N-1} f_Y(T^k_Y x) \to 0$ presque partout quand $N \to \pm\infty$, et l'hypothèse 2 du théorème 4.2.1 est encore satisfaite pour tout $b > 1/2$.

Soit $\kappa > 0$ très petit. Comme L est à variation lente, $L(B'_n) = O((B'_n)^\kappa)$, donc (4.21) donne $B'_n = O(n^{1/(p-\kappa)})$. Ainsi, si $b < \frac{p}{2}$, on a $B'_n = O(B_n^{1/b})$, ce qui implique (4.12'). □

4.2.3 Exemples

Applications de Liverani-Saussol-Vaienti

On considère tout d'abord l'application LSV de paramètre $\alpha \in]0,1[$ définie sur $[0,1]$ par

$$T(x) = \begin{cases} x(1+2^\alpha x^\alpha) & \text{si } 0 \leqslant x \leqslant 1/2, \\ 2x-1 & \text{si } 1/2 < x \leqslant 1. \end{cases} \tag{4.22}$$

Elle a été étudiée précisément dans le paragraphe 1.4.1. En particulier, elle est localement Gibbs-Markov de base $Y =]1/2,1]$ et admet une unique mesure de probabilité m invariante absolument continue, dont la densité h est lipschitzienne sur tout compact de $]0,1]$. De plus, si φ_Y désigne le temps de retour à Y, on a $m(\varphi_Y > n) \sim \frac{h(1/2)}{4(\alpha n)^{1/\alpha}}$ (théorème 1.4.1).

En estimant la vitesse de décorrélation des fonctions hölderiennes, dans le chapitre 2, on a montré que les corrélations de certaines fonctions hölderiennes étaient sommables, ce qui impliquait un théorème central limite. Plus précisément, si $\alpha < 1/2$, toute fonction hölderienne f de moyenne nulle satisfait un théorème central limite (proposition 2.3.9), tandis que si $\alpha > 1/2$ cela reste vrai si $f(0) = 0$ et $|f(x)| \leqslant Cx^\gamma$ pour un certain $\gamma > 2\alpha - 1$ (théorème 2.4.14).

Ces résultats ne sont pas satisfaisants car ils ne disent rien sur les théorèmes limites possibles lorsque $\alpha > 1/2$ et $f(0) \neq 0$. En utilisant le paragraphe précédent, on va pouvoir répondre à cette question :

Théorème 4.2.5. *Soient $\alpha \in]0,1[$ et T donnée par (4.22). Notons h la densité de sa mesure de probabilité invariante absolument continue. Soit $f : [0,1] \to \mathbb{R}$ hölderienne et d'intégrale nulle. Alors*

- *Si $\alpha < 1/2$, alors il existe $\sigma^2 \geqslant 0$ tel que $\frac{S_n f}{\sqrt{n}} \to \mathcal{N}(0,\sigma^2)$.*
- *Si $\alpha = 1/2$, alors*
  1. *si $f(0) = 0$, il existe $\sigma^2 \geqslant 0$ tel que $\frac{S_n f}{\sqrt{n}} \to \mathcal{N}(0,\sigma^2)$.*
  2. *si $f(0) \neq 0$, $\frac{S_n f}{\sqrt{n \log n}} \to \mathcal{N}(0, h(1/2) f(0)^2)$.*
- *Si $1/2 < \alpha < 1$, alors*
  1. *si $f(0) = 0$ et $|f(x)| \leqslant Cx^\gamma$ pour un certain $\gamma > \alpha - 1/2$, il existe $\sigma^2 \geqslant 0$ tel que $\frac{S_n f}{\sqrt{n}} \to \mathcal{N}(0,\sigma^2)$.*
  2. *si $f(0) \neq 0$, $\frac{S_n f}{n^\alpha} \to Z_{1/\alpha, c, \mathrm{sgn}(f(0))}$ où $c = \frac{h(1/2)}{4(\alpha/|f(0)|)^{1/\alpha}} \Gamma(1-1/\alpha) \cos\left(\frac{\pi}{2\alpha}\right)$.*

Démonstration. Soit $J_n =]y_{n+1}, y_n]$ l'ensemble des points de $Y =]1/2,1]$ dont le temps de retour à Y est égal à n ; cet intervalle a été introduit dans la preuve du théorème 1.4.1. Alors T est localement Gibbs-Markov pour la base Y et la partition J_n. De plus, $m(J_n) \sim \frac{c}{n^{1/\alpha+1}}$ pour une certaine constante $c > 0$.

Si $f : [0,1] \to \mathbb{R}$ est hölderienne, elle vérifie $Df_Y(J_n) \leqslant Cn$, donc $\sum m(J_n) Df_Y(J_n) \leqslant C \sum \frac{n}{n^{1/\alpha+1}} < +\infty$.

Supposons $\alpha < 1/2$. Alors la fonction f_Y est majorée sur J_n par Cn, si bien que

$$\int_Y f_Y(y)^2 \leqslant C^2 \sum n^2 m(J_n) < \infty.$$

Ainsi, $f_Y \in L^2(Y)$. De même, $\varphi_Y \in L^2(Y)$. Ainsi, le premier cas du théorème 4.2.4 montre que f satisfait un théorème central limite.

Comme les cas $\alpha = 1/2$ et $\alpha > 1/2$ sont semblables, on ne va traiter que le cas $\alpha > 1/2$. Supposons tout d'abord que $f(0) = 0$. Comme f est hölderienne, il existe $\gamma > 0$ tel que $|f(x)| \leqslant Cx^\gamma$. Quitte à diminuer γ, on peut supposer que $\gamma < \alpha$. Soit $x_n = f(y_{n+1}) \sim \frac{1}{2}(\alpha n)^{-1/\alpha}$ d'après la preuve du théorème 1.4.1. Pour $y \in J_n$, on a $|f_Y(y)| \leqslant C \sum_{k=0}^{n-1} x_k^\gamma$ donc $|f_Y(y)| \leqslant Cn^{1-\gamma/\alpha}$. En particulier,

$$\int |f_Y|^2 \leqslant C \sum \frac{n^{2(1-\gamma/\alpha)}}{n^{1+1/\alpha}},$$

qui est fini dès lors que $\gamma > \alpha - 1/2$. Ainsi, f_Y appartient à L^2 si $\gamma > \alpha - 1/2$. De plus, $m(\varphi_Y > n) \sim \frac{h(1/2)}{4(\alpha n)^{1/\alpha}}$ d'après le théorème 1.4.1. Ainsi, $m(\varphi_Y > x) \sim x^{-p} L(x)$ pour une certaine fonction à variation lente L, et $p = 1/\alpha \in]1,2[$. Le premier cas du théorème 4.2.4 s'applique donc encore, et il montre que f satisfait un théorème central limite.

Supposons finalement que $f(0) \neq 0$, par exemple $f(0) > 0$. Alors $f = f(0) + g$ où g est hölderienne et satisfait $g(0) = 0$. Ainsi $f_Y = f(0)\varphi_Y + g_Y$. Comme ci-dessus, il existe $\gamma > 0$ tel que $|g_Y(y)| \leqslant Cn^{1-\gamma/\alpha}$ sur J_n, si bien que $m(|g_Y| > x) = o(x^{-1/\alpha})$. De plus, $m(f(0)\varphi_Y > x) \sim \frac{h(1/2)}{4(\alpha x/f(0))^{1/\alpha}}$. Ajouter g_Y ne modifie pas cette asymptotique, et on obtient donc la même estimée pour $m(f_Y > x)$. Plus précisément, on peut écrire

$$m(f(0)\varphi_Y > x(1+\varepsilon)) - m(|g_Y| > x\varepsilon) \leqslant m(f_Y > x) \leqslant m(f(0)\varphi_Y > x(1-\varepsilon)) + m(|g_Y| > x\varepsilon).$$

Le fait que $m(|g_Y| > x) = o(x^{-1/\alpha})$ implique donc que $m(f_Y > x) \sim m(f(0)\varphi_Y > x) \sim \frac{h(1/2)}{4(\alpha x/f(0))^{1/\alpha}}$, qui s'écrit sous la forme $x^{-p}L(x)$ pour $p = 1/\alpha$ et une certaine fonction à variation lente L. De la même façon, on vérifie que $m(f_Y < -x) = o(x^{-p}L(x))$. Toutes les hypothèses du troisième cas du théorème 4.2.4 sont donc vérifiées, et ce théorème donne la conclusion recherchée. □

On peut aussi appliquer le résultat à d'autre exemples. Par exemple, considérons $f(x) = x^{-\beta} + \kappa$, avec $0 < \beta < 1 - \alpha$ de telle sorte que f est intégrable puisque la densité $h(x)$ diverge en 0 comme $x^{-\alpha}$, et où κ est choisi de telle sorte que $\int f = 0$. Un petit calcul montre que $Df_Y(J_n) \asymp n^{\beta/\alpha}$, donc $\sum m[J_n] Df_Y(J_n) < \infty$ puisque $\alpha + \beta < 1$.

De plus, sur J_n, $f_Y \sim Cn^{\beta/\alpha+1}$, donc $m[f_Y > x] \sim Cx^{-\frac{1}{\alpha+\beta}}$. Ainsi, quand $\alpha + \beta < 1/2$, on obtient un théorème central limite usuel, alors que quand $\alpha + \beta = 1/2$ on a un théorème

central limite avec normalisation $\frac{1}{\sqrt{n\log n}}$, et quand $1/2<\alpha+\beta<1$ on obtient une convergence vers une loi stable pour la normalisation $\frac{1}{n^{\alpha+\beta}}$. Ainsi, même quand $\alpha<1/2$, on peut avoir convergence vers une loi stable pour des observables non bornées.

Produit gauche en dimension 2

On va maintenant démontrer que le théorème 4.2.4 s'applique également au produit gauche étudié dans la partie 1.5. Dans tout ce paragraphe, T désignera donc une application satisfaisant toutes les hypothèses de cette partie, et on utilisera librement les notations et les résultats s'y rapportant.

Soit

$$A=\frac{1}{4\left(\alpha_{\min}{}^{3/2}\sqrt{\frac{\pi}{2\alpha''(x_0)}}\right)^{1/\alpha_{\min}}}\int_{S^1\times\{1/2\}}h\,\mathrm{dLeb},$$

où h est la densité de la mesure de probabilité invariante m par rapport à la mesure de Lebesgue. D'après la proposition 1.5.18, on a $m(\varphi_Y>n)\sim\left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}}A$. Alors :

Théorème 4.2.6. *Soit f une fonction hölderienne sur $S^1\times[0,1]$ avec $\int f\,\mathrm{d}m=0$. Notons $e=\int_{S^1\times\{0\}}f\,\mathrm{dLeb}$. Alors*

- *Si $\alpha_{min}<1/2$, il existe $\sigma^2\geqslant 0$ tel que $\frac{1}{\sqrt{n}}S_nf\to\mathcal{N}(0,\sigma^2)$.*
- *Si $\alpha_{min}=1/2$ et $e\neq 0$, alors $\frac{S_nf}{\sqrt{\frac{e^2A}{4}n(\ln n)^2}}\to\mathcal{N}(0,1)$.*
- *Si $1/2<\alpha_{min}<1$ et $e\neq 0$, alors $\frac{S_nf}{n^{\alpha_{min}}\sqrt{\alpha_{min}\ln n}}\to Z_{1/\alpha_{min},c,\mathrm{sgn}(e)}$ où $c=A|e|^{1/\alpha_{min}}\Gamma(1-1/\alpha_{min})\cos\left(\frac{\pi}{2\alpha_{min}}\right)$.*
- *Si $1/2\leqslant\alpha_{min}<1$ et $e=0$, supposons aussi qu'il existe $\gamma>\frac{\alpha_{max}}{\alpha_{min}}\left(\alpha_{min}-\frac{1}{2}\right)$ tel que $|f(\omega,x)-f(\omega,0)|\leqslant Cx^\gamma$. Alors il existe $\sigma^2\geqslant 0$ tel que $\frac{1}{\sqrt{n}}S_nf\to\mathcal{N}(0,\sigma^2)$.*

Les étapes de la preuve vont être les mêmes que celles du théorème 4.2.5. Fixons f hölderienne sur $S^1\times[0,1]$.

Lemme 4.2.7. *Si $\alpha_{min}<1/2$, alors $f_Y\in L^2(Y,\mathrm{d}m)$.*

Démonstration. On a

$$\begin{aligned}\int f_Y^2\,\mathrm{d}m&\leqslant C\sum m(\varphi_Y=n)n^2=C\sum\big(m(\varphi_Y>n-1)-m(\varphi_Y>n)\big)n^2\\&\leqslant C\sum m(\varphi_Y>n)n.\end{aligned}$$

Cette dernière estimée est sommable puisque $m(\varphi_Y>n)\sim A\left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}}$ d'après la proposition 1.5.18, et $1/\alpha_{\min}>2$. □

Lemme 4.2.8. *Supposons que* $\int_{S^1\times\{0\}} f = 0$. *Soit* $\alpha_{max} > \gamma > 0$ *tel que* $|f(\omega,x) - f(\omega,0)| \leqslant Cx^\gamma$. *Si* $1 < p < \min\left(\frac{2}{\alpha_{min}}, \frac{1}{\alpha_{min}(1-\gamma/\alpha_{max})}\right)$, *alors* $f_Y \in L^p(Y, \mathrm{d}m)$.

Démonstration. Comme h est bornée sur Y, il suffit de montrer que $f_Y \in L^p(Y, \mathrm{dLeb})$.

Supposons tout d'abord que $f \equiv 0$ sur $S^1 \times \{0\}$. Alors, si $\mathbf{x} = (\omega, x)$ satisfait $\varphi_Y(\mathbf{x}) = n$, on a $f_Y(\mathbf{x}) = \sum_0^{n-1} f(T^k\mathbf{x})$. Si $k \geqslant 1$, $T^k_\omega(x) \leqslant X_{n-k}(F^k\omega) \leqslant \frac{C}{(n-k)^{1/\alpha_{\max}}}$ par la proposition 1.5.2 donc $|f(T^k\mathbf{x})| \leqslant \frac{C}{(n-k)^{\gamma/\alpha_{\max}}}$, et on obtient en sommant $|f_Y(\mathbf{x})| \leqslant Cn^{1-\gamma/\alpha_{\max}}$. Ainsi,

$$\begin{aligned}\int |f_Y|^p &\leqslant C\sum m(\varphi_Y = n)n^{p(1-\gamma/\alpha_{\max})}\\ &= C\sum \big(m(\varphi_Y > n-1) - m(\varphi_Y > n)\big)n^{p(1-\gamma/\alpha_{\max})}\\ &\leqslant C\sum m(\varphi_Y > n)n^{p(1-\gamma/\alpha_{\max})-1}.\end{aligned}$$

Comme $m(\varphi_Y > n) \sim A\left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}}$, cette dernière série est sommable dès que

$$-\frac{1}{\alpha_{\min}} + p\left(1 - \frac{\gamma}{\alpha_{\max}}\right) - 1 < -1,$$

ce qui est le cas par hypothèse sur p.

Supposons maintenant seulement que l'intégrale de f sur $S^1 \times \{0\}$ s'annule. Soit $g(\omega, x) = f(\omega, 0)$. La fonction $f - g$ s'annule sur $S^1 \times \{0\}$ donc $f_Y - g_Y \in L^p$ d'après la première partie de la preuve. Il suffit donc de montrer que $g_Y \in L^p$. Posons $\chi(\omega) = f(\omega, 0)$ et $S_n\chi(\omega) = \sum_{k=0}^{n-1} \chi(F^k\omega)$. Ainsi, $g_Y(\omega, x) = S_{\varphi_Y(\omega,x)}\chi(\omega)$.

Soit $M_n\chi(\omega) = \max_{k\leqslant n} |S_k\chi(\omega)|$. Soient $\delta > 0$ et $l = \frac{1+\delta}{\delta}$, si bien que $\frac{1}{l} + \frac{1}{1+\delta} = 1$. Alors

$$\begin{aligned}\int_{\varphi_Y\geqslant 2} |g_Y|^p &= \sum_{n=2}^\infty \int_{S^1}\int_{1/2+X_n(F\omega)/2}^{1/2+X_{n-1}(F\omega)/2} \big|S_n\chi(\omega)\big|^p \,\mathrm{d}u\,\mathrm{d}\omega\\ &\leqslant \sum_{k=1}^\infty \int_{S^1}\int_{1/2+X_{2^k}(F\omega)/2}^{1/2+X_{2^{k-1}}(F\omega)/2} |M_{2^k}\chi(\omega)|^p \,\mathrm{d}u\,\mathrm{d}\omega\\ &\leqslant \sum_{k=1}^\infty \int_{S^1} X_{2^{k-1}}(F\omega)|M_{2^k}\chi(\omega)|^p \,\mathrm{d}\omega \leqslant \sum_{k=1}^\infty \|X_{2^{k-1}}\circ F\|_{1+\delta}\,\|M_{2^k}\chi\|_{lp}^p\,,\end{aligned}$$

où on a utilisé l'inégalité de Hölder à la dernière ligne. Si δ est assez petit, $lp > 2$, donc le corollaire 1.5.16 implique que $\|M_{2^k}\chi\|_{lp} \leqslant Ck^{\frac{lp-1}{lp}}\sqrt{2^k}$. De plus,

$$\|X_{2^{k-1}}\circ F\|_{1+\delta} = \|X_{2^{k-1}}\|_{1+\delta} \leqslant \left(\int X_{2^{k-1}}\right)^{1/(1+\delta)} \sim C\left(\frac{\sqrt{\ln(2^{k-1})}}{2^{k-1}}\right)^{\frac{1}{(1+\delta)\alpha_{\min}}}$$

par le théorème 1.5.17. Ainsi, $\int |g_Y|^p < \infty$ si $\frac{1}{(1+\delta)\,\alpha_{\min}} > \frac{p}{2}$, et il est possible de choisir $\delta > 0$ de sorte que cette inégalité soit satisfaite puisque $\frac{1}{\alpha_{\min}} > \frac{p}{2}$ par hypothèse. □

Démonstration du théorème 4.2.6. Soit f une fonction hölderienne. Le lemme 1.5.7 assure que $Df_Y(A_{s,n}) \leqslant Cn$ pour tout ensemble $A_{s,n}$ de la partition de Markov de T_Y, donc $\sum m[A_{s,n}]Df_Y(A_{s,n}) \leqslant C\sum nm[A_{s,n}] = C$ par la formule de Kac. Ainsi, la condition (4.19) est satisfaite.

Dans le cas $\alpha_{\min} < 1/2$, le lemme 4.2.7 montre que $f_Y \in L^2$. En appliquant ce lemme à la fonction 1, on obtient de même que $\varphi_Y \in L^2$. On a déjà vérifié (4.19), donc on peut appliquer le premier cas du théorème 4.2.4. Cela démontre que f satisfait un théorème central limite.

Les deuxième et troisième cas étant similaires, on ne va s'intéresser qu'au troisième, i.e. $1/2 < \alpha_{\min} < 1$ et $e \neq 0$. Supposons par exemple $e > 0$. Estimons $m(f_Y > x)$.

Fait : $m(f_Y > x) \sim \left(\frac{e\sqrt{\ln x}}{x}\right)^{1/\alpha_{min}} A$ *et* $m(f_Y < -x) = o\left(\frac{\sqrt{\ln x}}{x}\right)^{1/\alpha_{min}}$.

Démonstration. On démontre l'estimée sur $m(f_Y > x)$, l'autre étant analogue.

Soit $g \equiv e$ sur $S^1 \times [0,1]$. Alors $g_Y = ne$ sur $[\varphi_Y = n]$, ce qui implique que $m(g_Y > ne) = m(\varphi_Y > n) \sim \left(\frac{\sqrt{\ln n}}{n}\right)^{1/\alpha_{\min}} A$ d'après la proposition 1.5.18.

Dans le cas général, considérons $j = f - g$, et prouvons que $m(|j_Y| > x) = o\left(\frac{\sqrt{\ln x}}{x}\right)^{1/\alpha_{\min}}$. Comme $f_Y = g_Y + j_Y$, l'équation

$$m(g_Y > x(1+\varepsilon)) - m(|j_Y| > x\varepsilon) \leqslant m(f_Y > x) \leqslant m(g_Y > x(1-\varepsilon)) + m(|j_Y| > x\varepsilon)$$

permettra d'en déduire que $m(f_Y > x) \sim m(g_Y > x)$ et donc de conclure.

Soit $\gamma > 0$ avec $\gamma < \min(\theta, \alpha_{\max})$ où θ est le coefficient de Hölder de f. Le lemme 4.2.8 montre que $j_Y \in L^p$ pour tout $p < \min\left(\frac{2}{\alpha_{\min}}, \frac{1}{\alpha_{\min}(1-\gamma/\alpha_{\max})}\right)$. On peut en particulier choisir $p > 1/\alpha_{\min}$. Alors $m(|j_Y| > x) \leqslant \int \left(\frac{|j_Y|}{x}\right)^p = O(x^{-p})$, ce qui termine la preuve. ◊

Les mêmes propriétés sont satisfaites par φ_Y, avec la même preuve. On est donc dans le troisième cas du théorème 4.2.4, ce qui permet de conclure.

Supposons finalement que $\frac{1}{2} \leqslant \alpha_{\min} < 1$ et que $e = 0$. Sous les hypothèses du théorème, on peut appliquer le lemme 4.2.8 avec $p = 2$ et obtenir que $f_Y \in L^2$. Finalement, la proposition 1.5.18 montre que $m[\varphi_Y > x] \sim \left(\frac{\sqrt{\ln x}}{x}\right)^{1/\alpha_{\min}} A$. On a vérifié toutes les hypothèses du premier cas du théorème 4.2.4. □

4.3 Un résultat spectral perturbatif

Dans la suite de ce chapitre, on va développer une autre méthode pour démontrer le même type de résultats que dans le théorème 4.2.4 en s'affranchissant de toute hypothèse sur φ ou M, au prix d'un léger renforcement de (4.19) toujours vérifié dans les exemples. Cette méthode sera spectrale : on a vu dans la partie 4.1 que, en perturbant les opérateurs de transfert, on pouvait obtenir des théorèmes limites pour les applications Gibbs-Markov. Pour les tours localement Gibbs-Markov, on va utiliser la même idée, en perturbant les opérateurs de transfert de premier retour introduits dans le chapitre 2. Pour obtenir des informations sur l'application, on aura ensuite besoin d'un résultat spectral abstrait permettant d'estimer les inverses de certaines séries d'opérateurs (de la même façon que les résultats de la partie 2.2 ont permis d'étudier la vitesse de décorrélation). Cette partie est consacrée à la démonstration d'un théorème de ce type.

Rappelons que les suites d'opérateurs de renouvellement ont été définies dans la définition 2.2.1.

Théorème 4.3.1. *Soit R_n une suite d'opérateurs de renouvellement sur un espace de Banach $\mathcal{L}$. En particulier, soit $\mu > 0$ pour lequel le projecteur spectral P associé à la valeur propre 1 de $R(1) = \sum R_n$ satisfait $PR'(1)P = \mu P$.*

Soit $\alpha > 0$. Pour tous $n \geqslant 1$ et $t \in [-\alpha, \alpha]$, soit $R_n(t)$ un opérateur sur $\mathcal{L}$ tel que $t \mapsto R_n(t)$ soit continu en $t = 0$, avec $R_n(0) = R_n$ et $\|R_n(t)\| \leqslant Cr_n$ pour une certaine suite r_n telle que $\sum nr_n < \infty$. Pour $z \in \overline{\mathbb{D}}$ et $t \in [-\alpha, \alpha]$, notons

$$R(z,t) = \sum_{n=1}^{\infty} R_n(t) z^n.$$

C'est une perturbation continue de $R(z) = \sum R_n z^n$. Pour t petit et z proche de 1, $R(z,t)$ est proche de $R(1)$, si bien qu'il admet une unique valeur propre $\lambda(z,t)$ proche de 1. On suppose que $\lambda(1,t) = 1 - (c + o(1))\Lambda(|t|)$, pour $c \in \mathbb{C}$ avec $\mathrm{Re}(c) > 0$ et une certaine fonction continue $\Lambda : \mathbb{R}_+ \to \mathbb{R}_+$ qui ne s'annule qu'en 0.

Alors

1. *Il existe $\varepsilon_0 > 0$ tel que $\forall |t| < \varepsilon_0$, $I - R(z,t)$ est inversible pour tout $z \in \mathbb{D}$. On peut écrire $(I - R(z,t))^{-1} = \sum T_{n,t} z^n$.*
2. *De plus, il existe des fonctions $\varepsilon(t)$ et $\delta(n)$ qui tendent vers 0 quand $t \to 0$ et $n \to \infty$ telles que, $\forall |t| < \varepsilon_0$, $\forall n \in \mathbb{N}^*$,*

$$\left\| T_{n,t} - \frac{1}{\mu} \left(1 - \frac{c}{\mu} \Lambda(|t|)\right)^n P \right\| \leqslant \varepsilon(t) + \delta(n). \tag{4.23}$$

Par exemple, quand $\Lambda(t) = t^p$, ce résultat implique que $T_{n,t/n^{1/p}} \to \frac{e^{-c|t|^p/\mu}}{\mu} P$.

Remarque 4.3.2. On peut démontrer un théorème analogue lorsque la perturbation a lieu sur $[0,\alpha]$ au lieu de $[-\alpha,\alpha]$, avec exactement la même preuve. Ainsi, pour une perturbation sur $[-\alpha,\alpha]$ pour laquelle $\lambda(1,t) = 1 - a|t|^p + ib\,\mathrm{sgn}(t)|t|^p + o(|t|^p)$ avec $a > 0$, on peut encore obtenir la conclusion du théorème, en l'appliquant d'une part sur $[-\alpha,0]$ et d'autre part sur $[0,\alpha]$.

Dans les applications du théorème 4.3.1 à la preuve de théorèmes limites, les opérateurs R_n et $R_n(t)$ seront bien compris, tandis que $T_{n,t}$ sera relié à la fonction caractéristique $E(e^{itS_nf})$. Ainsi, la description de $T_{n,t}$ donnée par (4.23) permettra d'étudier la convergence des fonctions caractéristiques et donc d'obtenir des théorèmes limites.

Les résultats de la partie 2.2 décrivent précisément le comportement de $R(z) = R(z,0)$, il reste à comprendre l'influence du paramètre perturbatif t pour démontrer le théorème 4.3.1. Pour cela, on va utiliser les mêmes techniques d'analyse de Fourier que dans le paragraphe 2.2.2 . En particulier, on aura besoin du lemme de Wiener 2.2.3 et de la perturbation $\widetilde{R}(z)$ de $R(z)$ construite dans la preuve du théorème 2.2.5.

À partir de maintenant, et jusqu'à la fin de la preuve du théorème 4.3.1, les hypothèses et les notations de tous les résultats seront celles du théorème 4.3.1.

4.3.1 Inversibilité de $I - R(z,t)$

Dans ce paragraphe, on démontre que $I - R(z,t)$ est inversible lorsque $(z,t) \neq (1,0)$ et t est assez petit. Le principal problème vient de z proche de 1, où $R(z,t)$ a une valeur propre $\lambda(z,t)$ proche de 1. Étudions tout d'abord cette valeur propre.

Lemme 4.3.3. *Sur un voisinage de $(1,0)$, on peut écrire $\lambda(z,t) = 1 + (z-1)(\mu + a(z,t)) - (c + b(t))\Lambda(|t|)$ où les fonctions a et b tendent vers 0 respectivement en $(1,0)$ et 0.*

Démonstration. Comme

$$\frac{R(z,t) - R(1,t)}{z-1} = \sum_{n=0}^{\infty} z^n \left(\sum_{k=n+1}^{\infty} R_k(t) \right),$$

l'hypothèse de sommabilité assure que $(z,t) \mapsto \frac{R(z,t)-R(1,t)}{z-1}$ est continue en $t = 0$. En écrivant

$$\frac{P(z,t) - P(1,t)}{z-1} = \frac{1}{2i\pi} \int_{|u-1|=\delta} \frac{1}{uI - R(z,t)} \frac{R(z,t) - R(1,t)}{z-1} \frac{1}{uI - R(1,t)}\,\mathrm{d}u,$$

on obtient également la continuité de $(z,t) \mapsto \frac{P(z,t)-P(1,t)}{z-1}$ en $t = 0$. Ainsi, en appliquant une forme linéaire continue à l'équation

$$\frac{\lambda(z,t) - \lambda(1,t)}{z-1} P(z,t) = \frac{R(z,t) - R(1,t)}{z-1} P(z,t) + (R(1,t) - \lambda(1,t)) \frac{P(z,t) - P(1,t)}{z-1},$$

on obtient également la continuité de $G(z,t) = \frac{\lambda(z,t)-\lambda(1,t)}{z-1}$ en $t=0$. De plus $G(z,0) \to \mu$ quand $z \to 1$ d'après l'équation (2.12). On peut donc écrire $G(z,t) = \mu + a(z,t)$ où $a(z,t)$ s'annule en $(1,0)$. De plus, par hypothèse, $\lambda(1,t) = 1-(c+b(t))\Lambda(|t|)$ où b s'annule en 0. Ainsi,

$$\lambda(z,t) = \lambda(1,t) + (z-1)G(z,t) = 1-(c+b(t))\Lambda(|t|) + (z-1)(\mu + a(z,t)). \qquad \square$$

Proposition 4.3.4. *Il existe $\varepsilon_0 > 0$ tel que $\forall (z,t) \in (\overline{\mathbb{D}}\times[-\varepsilon_0,\varepsilon_0])\backslash\{(1,0)\}$, $I - R(z,t)$ est inversible.*

Démonstration. Supposons qu'il existe un voisinage V de $(1,0)$ tel que, sur $V\backslash\{(1,0)\}$, $I - R(z,t)$ est inversible. On peut supposer que V est de la forme $U\times[-\varepsilon_1,\varepsilon_1]$. Pour $w \in \overline{\mathbb{D}}\backslash U$, $I - R(w)$ est inversible, donc $I - R(z,t)$ est inversible dans un petit voisinage de $(w,0)$, qu'on peut prendre de la forme $U_w\times[-\varepsilon_w,\varepsilon_w]$. Par compacité, $\overline{\mathbb{D}}\backslash U$ est recouvert par un nombre fini des U_w, disons $U_{w_1},\dots,U_{w_n}$. Pour conclure, il suffit de prendre $\varepsilon_0 = \min(\varepsilon_1,\varepsilon_{w_1},\dots,\varepsilon_{w_n})$.

Ainsi, il suffit de construire un voisinage V comme ci-dessus. On doit montrer que la valeur propre $\lambda(z,t)$ de $R(z,t)$ est différente de 1 quand (z,t) est proche – et différent – de $(1,0)$. Cela vient essentiellement du développement asymptotique de $\lambda(z,t)$ donné par le lemme 4.3.3, mais il faut écrire les choses soigneusement.

Soit K tel que $\frac{\mu}{4} > \frac{3|c|}{2K}$. Le voisinage V sera l'ensemble des couples (z,t) tels que $|b(t)| \leqslant \mathrm{Re}(c)/2$ et $|a(z,t)| \leqslant \min\left(\frac{\mu}{4}, \frac{\mathrm{Re}(c)}{2K}\right)$.

Fixons $(z,t) \in V \cap (\overline{\mathbb{D}}\times\mathbb{R})$. Supposons que $\lambda(z,t) = 1$, on va montrer que $(z,t) = (1,0)$ en traitant deux cas : soit $|\,\mathrm{Im}(z-1)| \geqslant |z-1|/2$, soit $|\,\mathrm{Re}(z-1)| \geqslant |z-1|/2$.

1. Si $|\,\mathrm{Im}(z-1)| \geqslant |z-1|/2$. Supposons par exemple $\mathrm{Im}(z-1) \geqslant |z-1|/2$. Comme $\mu > 0$ et $z \in \overline{\mathbb{D}}$, on a $\mathrm{Re}(\mu(z-1)) \leqslant 0$ donc

$$\begin{aligned} 1 = \mathrm{Re}(\lambda(z,t)) &= 1 + \mathrm{Re}(\mu(z-1)) + \mathrm{Re}(a(z,t)(z-1)) - (\mathrm{Re}(c) + \mathrm{Re}(b(t))\Lambda(|t|) \\ &\leqslant 1 + |a(z,t)||z-1| - (\mathrm{Re}(c) - |b(t)|)\Lambda(|t|) \leqslant 1 + \frac{\mathrm{Re}(c)}{2K}|z-1| - \frac{\mathrm{Re}(c)}{2}\Lambda(|t|). \end{aligned}$$

   Ainsi, $|z-1| \geqslant K\Lambda(|t|)$. La partie imaginaire donne alors

$$0 = \mathrm{Im}(\lambda(z,t)) \geqslant (\mu - |a(z,t)|)\,\mathrm{Im}(z-1) - (|c| + |b(t)|)\Lambda(|t|) \geqslant \frac{\mu}{2}\frac{|z-1|}{2} - \frac{3|c|}{2}\frac{|z-1|}{K}.$$

   La définition de K implique que $|z-1| = 0$. Comme $|z-1| \geqslant K\Lambda(|t|)$ et Λ ne s'annule qu'en 0, on obtient aussi $t = 0$.

2. Si $|\,\mathrm{Re}(z-1)| \geqslant |z-1|/2$. Comme $\mathrm{Re}(z-1) \leqslant 0$, on a en fait $\mathrm{Re}(z-1) \leqslant -|z-1|/2$.

Alors

$$\begin{aligned} 1 = \mathrm{Re}(\lambda(z,t)) &\leqslant 1 + \mathrm{Re}(\mu(z-1)) + |a(z,t)||z-1| - (\mathrm{Re}(c) - |b(t)|)\Lambda(|t|) \\ &\leqslant 1 - \mu\frac{|z-1|}{2} + |a(z,t)||z-1| - \frac{\mathrm{Re}(c)}{2}\Lambda(|t|) \\ &\leqslant 1 - |z-1|\frac{\mu}{4} - \frac{\mathrm{Re}(c)}{2}\Lambda(|t|). \end{aligned}$$

Ainsi, $|z-1| = 0$, i.e. $z = 1$, et $\Lambda(|t|) = 0$ donc $t = 0$. □

En particulier, pour $|t| < \varepsilon_0$, la série entière $I - R(z,t)$ est partout inversible sur $\mathbb{D}$. Son inverse est également analytique par [DS58, lemme VII.6.13], et il peut donc s'écrire sous la forme $\sum T_{n,t} z^n$. De plus, pour $t \neq 0$, tout peut s'étendre continûment à la frontière $\partial\mathbb{D}$, ce qui implique que les $T_{n,t}$ sont les coefficients de Fourier de la fonction $(I - R(z,t))^{-1}$ vue comme une fonction de S^1 dans $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$.

4.3.2 Convergence de $\widetilde{R}(z,t)$

Dans ce paragraphe, on va travailler exclusivement sur S^1 (ainsi, z désignera un point de S^1 et *pas* de $\mathbb{D}$).

Pour utiliser des méthodes de séries de Fourier sur $R(z,t)$, on aura besoin qu'il ait une valeur propre proche de 1 bien définie pour tout z, ce qui n'est *a priori* pas le cas. On va donc remplacer $R(z,t)$ par une autre série $\widetilde{R}(z,t)$.

Construction de $\widetilde{R}(z,t)$.

On utilise la fonction $\widetilde{R}(z)$ qui a été construire dans la deuxième étape de la preuve du théorème 2.2.5. On pose $\widetilde{R}(z,t) = \widetilde{R}(z) - (R(z) - \sum_{n=1}^{\infty} z^n R_n(t))$. L'hypothèse de sommabilité uniforme du théorème 4.3.1, associée à la continuité de $t \mapsto R_n(t)$ en 0, assurent que l'application $t \mapsto \widetilde{R}(\cdot,t)$ est continue en $t = 0$ pour la norme de $A(\mathcal{A})$. Rappelons que cette algèbre de Banach introduite dans le paragraphe 2.2.1 est constituée des fonctions continues de S^1 dans $\mathcal{A}$ dont les coefficients sont sommables. En particulier, l'application $t \mapsto \widetilde{R}(\cdot,t)$ est également continue pour la norme du sup, si bien que $\widetilde{R}(z,t)$ a une valeur propre $\widetilde{\lambda}(z,t)$ proche de 1 bien définie pour t assez petit. De plus, pour z dans un voisinage de 1, on a $\widetilde{\lambda}(z,t) = \lambda(z,t)$, qui est différent de 1 quand $(z,t) \neq 0$. Si z n'est pas dans ce voisinage de 1, alors $\widetilde{\lambda}(z,t)$ est proche de $\widetilde{\lambda}(z)$, et est donc également différent de 1 si t est assez petit.

Le résultat principal de ce paragraphe est la proposition suivante. Rappelons que les hypothèses sont celles du théorème 4.3.1.

Proposition 4.3.5. *On a*

$$\left\| \left(\frac{I - \widetilde{R}(z,t)}{1 - (1 - \frac{c}{\mu}\Lambda(|t|))z} \right)^{-1} - \left(\frac{I - \widetilde{R}(z)}{1 - z} \right)^{-1} \right\|_{A(\mathcal{A})} \xrightarrow[t\to 0]{} 0.$$

Démonstration. Soient $\widetilde{P}(z,t)$ le projecteur spectral associé à la valeur propre $\widetilde{\lambda}(z,t)$ de $\widetilde{R}(z,t)$ et $\widetilde{Q}(z,t) = I - \widetilde{P}(z,t)$. On a alors

$$I - \widetilde{R}(z,t) = (1 - \widetilde{\lambda}(z,t))\widetilde{P}(z,t) + (I - \widetilde{R}(z,t)\widetilde{Q}(z,t))\widetilde{Q}(z,t).$$

Ainsi,

$$\left(\frac{I - \widetilde{R}(z,t)}{1 - (1 - \frac{c}{\mu}\Lambda(|t|))z}\right)^{-1} = \frac{1 - (1 - \frac{c}{\mu}\Lambda(|t|))z}{1 - \widetilde{\lambda}(z,t)}\widetilde{P}(z,t) + \left(1 - \left(1 - \frac{c}{\mu}\Lambda(|t|)\right)z\right)(I - \widetilde{R}(z,t)\widetilde{Q}(z,t))^{-1}\widetilde{Q}(z,t).$$

Comme $\widetilde{R}(z,t) \to \widetilde{R}(z)$ lorsque $t \to 0$, la continuité de l'inversion donne la convergence dans $A(\mathcal{A})$ de $\widetilde{P}(z,t)$ et $\widetilde{Q}(z,t)$ respectivement vers $\widetilde{P}(z)$ et $\widetilde{Q}(z)$. Comme $I - \widetilde{R}(z)\widetilde{Q}(z)$ est inversible pour tout $z \in S^1$, le lemme de Wiener montre que son inverse appartient également à $A(\mathcal{A})$. On obtient alors également la convergence de $(I - \widetilde{R}(z,t)\widetilde{Q}(z,t))^{-1}$ vers $(I - \widetilde{R}(z)\widetilde{Q}(z))^{-1}$ dans $A(\mathcal{A})$.

Il suffit donc de démontrer que $\frac{1-(1-\frac{c}{\mu}\Lambda(|t|))z}{1-\widetilde{\lambda}(z,t)} \to \frac{1-z}{1-\widetilde{\lambda}(z)}$ (dans $A(\mathbb{C})$) pour conclure. Mais cette convergence résulte du lemme suivant, du lemme de Wiener et de la continuité de l'inversion. □

Lemme 4.3.6. *On a*

$$\left\| \frac{1 - \widetilde{\lambda}(z,t)}{1 - (1 - \frac{c}{\mu}\Lambda(|t|))z} - \frac{1 - \widetilde{\lambda}(z)}{1 - z} \right\|_{A(\mathbb{C})} \xrightarrow[t\to 0]{} 0.$$

Démonstration. On va travailler la plupart du temps avec le cercle S^1 plongé dans $\mathbb{C}$, mais on aura parfois également besoin de le considérer comme $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$, auquel cas la variable sera notée θ.

Rappelons quelques résultats sur les séries de Fourier. Pour $|\theta| \leqslant \pi$ et $\varepsilon > 0$ petit, soit $\Delta_\varepsilon(\theta) = \sup\left(0, 1 - \frac{|\theta|}{\varepsilon}\right)$: cette fonction est supportée par $\{\theta \in [-\varepsilon, \varepsilon]\}$. Notons aussi $V_\varepsilon = 2\Delta_{2\varepsilon} - \Delta_\varepsilon$: cette fonction est égale à 1 sur $\{\theta \in [-\varepsilon, \varepsilon]\}$ et s'annule hors de $\{\theta \in [-2\varepsilon, 2\varepsilon]\}$. On utilisera V_ε et $1 - V_\varepsilon$ comme partition (continue) de l'unité. Par [Kah70, page 56], $\|V_\varepsilon\|_{A(\mathbb{C})} \leqslant 3$. De plus, V_ε est C^1 par morceaux, sa dérivée étant bornée par $1/\varepsilon$.

Si $g : \mathbb{T} \to \mathbb{C}$ est continue et C^1 par morceaux, alors ([Kah70, page 56, équation (1)])

$$\|g\|_{A(\mathbb{C})} \leqslant |c_0(g)| + \left(\frac{\pi}{12}\int_{-\pi}^{\pi} |g'(\theta)|^2 \,\mathrm{d}\theta\right)^{1/2}. \tag{4.24}$$

Finalement, si $f : S^1 \to \mathbb{C}$ appartient à $A(\mathbb{C})$ et satisfait $f(1) = 0$, alors ([Kah70, page 56])

$$\|fV_\varepsilon\|_{A(\mathbb{C})} \xrightarrow[\varepsilon\to 0]{} 0. \tag{4.25}$$

Fixons $|t| < \varepsilon_0$ et notons $\psi_t = V_{G(|t|)}$ pour $G(t) = \Lambda(t)^{1/3}$ et $\varphi_t = 1 - \psi_t$, sur S^1 ou $\mathbb{T}$. En notant $A_t(z) = \frac{1-\widetilde{\lambda}(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}$, on doit montrer que $A_t(z)$ tend dans $A(\mathbb{C})$ vers $B(z) = \frac{1-\widetilde{\lambda}(z)}{1-z}$. On va montrer que $A_t(z)\psi_t(z) - B(z)\psi_t(z) \to 0$ et $A_t(z)\varphi_t(z) - B(z)\varphi_t(z) \to 0$.

Dans cette preuve, on écrira $C(z)$ pour une fonction dont la norme dans $A(\mathbb{C})$ est bornée par une constante. Par exemple, comme $\|V_t\|_{A(\mathbb{C})} \leqslant 3$, on remplacera parfois $V_t(z)$ par $C(z)$. De la même façon, $\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z} = 1 - \frac{c}{\mu}\Lambda(|t|)\sum_{n=1}^{\infty}(1-\frac{c}{\mu}\Lambda(|t|))^{n-1}z^n$, si bien que $\left\|\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\right\|_{A(\mathbb{C})} \leqslant 1 + \frac{\frac{|c|}{\mu}\Lambda(|t|)}{1-\left|1-\frac{c}{\mu}\Lambda(|t|)\right|}$. Comme $\operatorname{Re} c > 0$, cette quantité reste bornée quand $t \to 0$. Ainsi, on pourra également remplacer $\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}$ par $C(z)$.

En utilisant exactement les mêmes équation que dans la preuve du lemme 4.3.3 (mais avec des tildes partout) et le lemme de Wiener, on vérifie que $K(z,t) = \frac{\widetilde{\lambda}(z,t)-\widetilde{\lambda}(1,t)}{z-1} - \frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(1)}{z-1}$ tend vers 0 dans $A(\mathbb{C})$ quand $t \to 0$.

Étape 1 (près de $z = 1$) : $A_t(z)\psi_t(z) - B(z)\psi_t(z)$ tend vers 0 dans $A(\mathbb{C})$.

On sait que $\frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(1)}{z-1} \in A(\mathbb{C})$ d'après la troisième étape de la preuve du théorème 2.2.5. Comme ce quotient tend vers μ quand $z \to 1$, on peut écrire $\frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(1)}{z-1} = \mu + F(z)$ où $F \in A(\mathbb{C})$ et $F(1) = 0$. Comme $\widetilde{R}(z) = R(z)$ dans un voisinage de 1, on a $\widetilde{\lambda}(1,t) = \lambda(1,t) = 1 - (c + b(t))\Lambda(|t|)$. Cela donne

$$\widetilde{\lambda}(z,t) = 1 - (c + b(t))\Lambda(|t|) + (z-1)\left(K(z,t) + F(z) + \mu\right).$$

Ainsi,

$$\begin{aligned} A_t(z) &= \frac{1-\widetilde{\lambda}(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z} \\ &= \mu + \frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\left(F(z) + K(z,t) + c\Lambda(|t|)\right) + \frac{\Lambda(|t|)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}b(t). \end{aligned}$$

Par conséquent,

$$\begin{aligned} &A_t(z)\psi_t(z) \\ &\quad = \mu\psi_t(z) + C(z)F(z)\psi_t(z) + C(z)(K(z,t) + \Lambda(|t|)) + \frac{\Lambda(|t|)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}b(t)\psi_t(z). \end{aligned}$$

On vérifie aisément que $\frac{\Lambda(|t|)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}$ est borné dans $A(\mathbb{C})$. De plus, $\mu = B(z) - F(z)$, et le terme $F(z)\psi_t(z)$ peut être inclus dans $C(z)F(z)\psi_t(z)$. Finalement,

$$A_t(z)\psi_t(z) - B(z)\psi_t(z) = C(z)F(z)\psi_t(z) + C(z)(\Lambda(|t|) + K(z,t) + b(t)).$$

Dans le terme le plus à droite, tout tend vers 0 quand $t \to 0$. De plus, (4.25) montre que $F(z)\psi_t(z) \to 0$. Cela conclut la première étape.

Étape 2 (loin de $z = 1$) : $A_t(z)\varphi_t(z) - B(z)\varphi_t(z)$ tend vers 0 dans $A(\mathbb{C})$.

On part de

$$\frac{1-\widetilde{\lambda}(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z} - \frac{1-\widetilde{\lambda}(z)}{1-z}$$
$$= \frac{1-\widetilde{\lambda}(z)}{1-z}\left(\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z} - 1\right) + \frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(z,t)}{1-z}.$$

On va voir que chaque terme à droite, multiplié par φ_t, tend vers 0 dans $A(\mathbb{C})$ quand $t \to 0$.

On va utiliser le fait que $|z-1| \geqslant CG(|t|)$ sur le support de φ_t. Cela implique que $|1-(1-\frac{c}{\mu}\Lambda(|t|))z| \geqslant CG(|t|) - D\Lambda(|t|) \geqslant C'G(|t|)$ puisque $\Lambda(t) = o(G(t))$. Dans la suite, C désignera une constante générique.

Le premier terme.

Comme $\frac{1-\widetilde{\lambda}(z)}{1-z} \in A(\mathbb{C})$, il suffit de montrer que $f = \left(\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z} - 1\right)\varphi_t(z)$ tend vers 0. On va utiliser (4.24).

On a

$$|f(z)| \leqslant \frac{\frac{|c|}{\mu}\Lambda(|t|)}{CG(|t|)} \leqslant C\Lambda(|t|)^{2/3}.$$

En particulier, $|c_0(f)| \leqslant C\Lambda(|t|)^{1/3}$.

Pour la dérivée, si on note $f(\theta) = h(\theta)\varphi_t(\theta)$, on a $f' = h'\varphi_t + h\varphi_t'$. Comme $|h| \leqslant C\Lambda(|t|)^{2/3}$ et $|\varphi_t'| \leqslant 1/G(|t|)$, on obtient $|h\varphi_t'| \leqslant C\Lambda(|t|)^{1/3}$.

Pour estimer h', on écrit

$$h(\theta) = -\frac{\frac{c}{\mu}\Lambda(|t|)}{1-\frac{c}{\mu}\Lambda(|t|)} + \frac{\frac{c}{\mu}\Lambda(|t|)}{1-\frac{c}{\mu}\Lambda(|t|)}\frac{1}{1-(1-\frac{c}{\mu}\Lambda(|t|))e^{i\theta}}.$$

Ainsi,

$$h'(\theta) = i\frac{c\Lambda(|t|)}{\mu}\frac{1}{(1-(1-\frac{c}{\mu}\Lambda(|t|))e^{i\theta})^2}.$$

Sur le support de φ_t, le module du dénominateur est $\geqslant (CG(|t|))^2$. On obtient donc $|h'\varphi_t| \leqslant C\Lambda(|t|)^{1/3}$.

On a montré que $|f'| \leqslant C\Lambda(|t|)^{1/3}$. Comme $|c_0(f)| \leqslant C\Lambda(|t|)^{1/3}$, (4.24) implique donc que $\|f\|_{A(\mathbb{C})} \leqslant C\Lambda(|t|)^{1/3}$.

Le deuxième terme.

Comme $\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}$ est borné dans $A(\mathbb{C})$, il suffit de montrer que $g(z) = \frac{\widetilde{\lambda}(z,t)-\widetilde{\lambda}(z)}{1-z}\varphi_t(z)$ tend vers 0 dans $A(\mathbb{C})$.

Écrivons encore $K(z,t) = \frac{\widetilde{\lambda}(z,t)-\widetilde{\lambda}(1,t)}{z-1} - \frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(1)}{z-1}$. Cette différence tend vers 0 quand $t \to 0$. On a

$$g(z) = K(z,t)\varphi_t(z) + (\widetilde{\lambda}(1) - \widetilde{\lambda}(1,t))\frac{\varphi_t(z)}{1-z}.$$

Le premier terme $K(z,t)\varphi_t(z)$ tend vers 0. Pour le second, posons $h(z) = \frac{\varphi_t(z)}{1-z}$. On a $|c_0(h)| \leqslant \frac{1}{CG(|t|)}$ et

$$|h'(\theta)| = \left| i\frac{\varphi_t(\theta)}{(1-e^{i\theta})^2} + \frac{1}{1-e^{i\theta}}\varphi_t'(\theta)\right| \leqslant \frac{1}{C^2G(|t|)^2} + \frac{1}{CG(|t|)}\frac{1}{G(|t|)}.$$

L'équation (4.24) montre que $\|h\|_{A(\mathbb{C})} \leqslant \frac{C}{G(|t|)^2} = O(\Lambda(|t|)^{-2/3})$. Comme $\widetilde{\lambda}(1) - \widetilde{\lambda}(1,t) = O(\Lambda(|t|))$, on obtient finalement $\left\|(\widetilde{\lambda}(1) - \widetilde{\lambda}(1,t))\frac{\varphi_t(z)}{1-z}\right\|_{A(\mathbb{C})} = O(\Lambda(|t|)^{1/3})$, ce qui conclut la deuxième étape de la preuve du lemme 4.3.6. □

4.3.3 Conclusion de la preuve du théorème 4.3.1

Proposition 4.3.7. *On a*

$$\left\| \left(\frac{I-R(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\right)^{-1} - \left(\frac{I-R(z)}{1-z}\right)^{-1}\right\|_{A(\mathcal{A})} \xrightarrow[t\to 0]{} 0.$$

Démonstration. Soit χ_1, χ_2 une partition C^∞ de l'unité de S^1 telle que $R = \widetilde{R}$ sur le support de χ_1, et posons $A(z,t) = \left(\frac{I-R(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\right)^{-1}$.

La proposition 4.3.5 montre que $\chi_1(z)A(z,t) \to \chi_1(z)A(z,0)$. Pour le terme $\chi_2(z)A(z,t)$, on peut modifier R sur un petit voisinage de 1 pour que $I - R(z)$ soit partout inversible, ce qui ne change pas $\chi_2(z)A(z,t)$. Le lemme de Wiener et la continuité de l'inversion donnent alors que $\chi_2(z)A(z,t) \to \chi_2(z)A(z,0)$. □

Preuve du théorème 4.3.1. Pour conclure la preuve, on va construire des fonctions $\varepsilon(t)$ et $\delta(n)$ qui satisfont (4.23).

La proposition 4.3.7 montre qu'on peut écrire

$$\left(\frac{I-R(z,t)}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\right)^{-1}=\left(\frac{I-R(z)}{1-z}\right)^{-1}+F_t(z)$$

où $F_t(z)\in A(\mathcal{A})$ tend vers 0 quand $t\to 0$. De plus, la proposition 2.2.9 montre que

$$\left(\frac{I-R(z)}{1-z}\right)^{-1}=\frac{1}{\mu}P+(1-z)A(z)$$

où $A(z)=\sum A_n z^n$ avec $A_n\to 0$. On obtient

$$\begin{aligned}\sum T_{n,t}z^n&=(I-R(z,t))^{-1}\\&=\frac{1}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}\frac{1}{\mu}P+\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}A(z)+\frac{1}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}F_t(z).\end{aligned}$$

Dans le premier terme, le coefficient de z^n est $\frac{1}{\mu}\left(1-\frac{c}{\mu}\Lambda(|t|)\right)^n P$ et correspond à (4.23). On doit vérifier que les deux autres termes peuvent être majorés par $\varepsilon(t)+\delta(n)$.
Soit $\delta(n)$ le maximum pour $t\in[-\varepsilon_0,\varepsilon_0]$ du coefficient $c_n(t)$ de z^n dans $\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}A(z)$. On va montrer que $\delta(n)\to 0$ quand $n\to\infty$. On a

$$\frac{1-z}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}=1-\frac{c}{\mu}\Lambda(|t|)\sum_{n=1}^{\infty}\left(1-\frac{c}{\mu}\Lambda(|t|)\right)^{n-1}z^n.$$

Ainsi,

$$c_n(t)=A_n-\frac{c}{\mu}\Lambda(|t|)\sum_{k=0}^{n-1}\left(1-\frac{c}{\mu}\Lambda(|t|)\right)^{n-k-1}A_k.$$

Comme $\mathrm{Re}(c)>0$ et $\Lambda(t)\to 0$ quand $t\to 0$, on a $\left|1-\frac{c}{\mu}\Lambda(|t|)\right|\leqslant 1-\frac{\mathrm{Re}\,c}{2\mu}\Lambda(|t|)$ si t est assez petit. Soit C qui majore $\|A_n\|$ pour $n\in\mathbb{N}$. Fixons $\varepsilon>0$, et N tel que $\forall k\geqslant N$, $\|A_k\|\leqslant\varepsilon$. Alors, si $n\geqslant N$,

$$\begin{aligned}\|c_n(t)\|&\leqslant\|A_n\|+\frac{|c|}{\mu}\Lambda(|t|)\sum_{k=0}^{N-1}\left(1-\frac{\mathrm{Re}\,c}{2\mu}\Lambda(|t|)\right)^{n-k-1}\|A_k\|\\&\quad+\frac{|c|}{\mu}\Lambda(|t|)\sum_{k=N}^{n-1}\left(1-\frac{\mathrm{Re}\,c}{2\mu}\Lambda(|t|)\right)^{n-k-1}\varepsilon\\&\leqslant\varepsilon+N\frac{|c|}{\mu}C\Lambda(|t|)\left(1-\frac{\mathrm{Re}\,c}{2\mu}\Lambda(|t|)\right)^{n-N}+E\varepsilon.\end{aligned}$$

Si ε_0 est assez petit, $\Lambda(|t|)\left(1-\frac{\mathrm{Re}\, c}{2\mu}\Lambda(|t|)\right)^{n-N}$ tend uniformément vers 0 sur $[-\varepsilon_0,\varepsilon_0]$ quand $n\to\infty$. Pour n assez grand, on obtient donc $\|c_n(t)\|\leqslant(2+E)\varepsilon$, et $\delta(n)\leqslant(2+E)\varepsilon$.

Majorons finalement les coefficients de $\frac{1}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}F_t(z)$, ce qui donnera le terme $\varepsilon(t)$. Les coefficients de Fourier de $\frac{1}{1-(1-\frac{c}{\mu}\Lambda(|t|))z}$ sont $\leqslant 1$, ce qui implique que le coefficient de z^n dans le produit est au plus $\sum_{k=0}^n\|F_t(z)_k\|\leqslant\|F_t\|_{A(\mathcal{A})}$. Comme $\|F_t\|_{A(\mathcal{A})}\to 0$ quand $t\to 0$, on peut prendre $\varepsilon(t)=\|F_t\|_{A(\mathcal{A})}$ pour conclure. □

4.4 Tours localement Gibbs-Markov : méthode spectrale

Dans cette partie, on s'intéresse à une tour localement Gibbs-Markov transitive $T:X\to X$ de base Y et préservant une mesure de probabilité m. On n'utilisera que des distances sur X de la forme $d_\tau(x,y)=\tau^{s(x,y)}$ où $\tau<1$. Comme toute distance pour laquelle T est Gibbs-Markov est majorée par une telle distance d'après la remarque 1.1.5, on ne perd pas en généralité. On notera $D_\tau f(a)$ la meilleure constante de Lipschitz de f sur a pour la distance d_τ. On notera $\varphi_Y:Y\to\mathbb{N}^*$ le temps de retour à Y et α_0 la partition induite sur Y par la partition de Markov de T.

On dira qu'une fonction $f:X\to\mathbb{R}$ appartient à $I(X,Y)$ si :

1. $\forall N\in\mathbb{N}$, il existe des constantes C_N et $0<\tau_N<1$ telles que, sur $Y_N=\bigcup_0^{N-1}T^i(Y)$, on ait $|f|\leqslant C_N$ et, si $x,y\in Y_N$ satisfont $s(x,y)\geqslant 1$, alors $|f(x)-f(y)|\leqslant C_N\tau_N^{s(x,y)}$.
2. Notons $f_Y(x)=\sum_{k=0}^{\varphi_Y(x)-1}f(T^kx)$ l'application induite par f sur Y. Alors il existe $\tau<1$ tel que $\sum_{a\in\alpha_0}m[a]D_\tau f_Y(a)<\infty$.
3. Il existe $\tau<1$ et $0<\eta\leqslant 1$ tels que $\sum_{a\in\alpha_0}\varphi_Y(a)m[a]D_\tau f_Y(a)^\eta<\infty$.

Les fonctions de $I(X,Y)$ induisent des "bonnes" fonctions sur la base Y, ce qui va permettre de démontrer des théorèmes limites en utilisant les résultats spectraux de la partie 4.3.

Par exemple, si f est bornée et hölderienne et l'application induite f_Y est également hölderienne sur Y, alors ces conditions sont satisfaites pour $\eta=1$ car $\sum_{a\in\alpha_0}\varphi_Y(a)m(a)=m(X)<\infty$, par la formule de Kac. Les conditions plus faibles dans la définition de $I(X,Y)$ sont utiles dans les applications, par exemple quand f est non bornée.

Si f est bornée et hölderienne et $m[\varphi_Y>n]=O(1/n^\beta)$ pour un certain $\beta>1$, alors $f\in I(X,Y)$: la première condition est claire, la seconde résulte de la formule de Kac car $D_\tau f_Y(a)\leqslant C\varphi_Y(a)$, et la troisième est satisfaite pour tout $\eta<\min(1,\beta-1)$.

Théorème 4.4.1. *Soit $T:X\to X$ une tour localement Gibbs-Markov transitive de base Y, préservant une mesure de probabilité m. Soit $f\in I(X,Y)$ une fonction intégrable de moyenne nulle. Alors*

– *Si $f_Y\in L^2(Y)$, il existe une constante $\sigma^2\geqslant 0$ telle que $\frac{1}{\sqrt{n}}S_nf\to\mathcal{N}(0,\sigma^2)$. De plus, $\sigma^2=0$ si et seulement si il existe une fonction mesurable $\chi:X\to\mathbb{R}$ telle que $f=\chi-\chi\circ T$*

presque partout.
- *Si* $m(|f_Y| > x) = x^{-2}l(x)$ *avec* $L(x) := 2\int_1^x \frac{l(u)}{u}\mathrm{d}u$ *à variation lente et non bornée, soit* $B_n \to +\infty$ *telle que* $nL(B_n) = B_n^2$. *Alors* $\frac{1}{B_n}S_n f \to \mathcal{N}(0,1)$.
- *Si* $m(f_Y > x) = (c_1 + o(1))x^{-p}L(x)$ *et* $m(f_Y < -x) = (c_2 + o(1))x^{-p}L(x)$, *où* L *est une fonction à variation lente,* $p \in]1,2[$, *et* $c_1, c_2 \geqslant 0$ *satisfont* $c_1 + c_2 > 0$, *soit* $B_n \to +\infty$ *telle que* $nL(B_n) = B_n^p$. *Alors* $\frac{1}{B_n}S_n f \to Z_{p,c,\beta}$ *où* $c = (c_1 + c_2)\Gamma(1-p)\cos\left(\frac{p\pi}{2}\right)$ *et* $\beta = \frac{c_1 - c_2}{c_1 + c_2}$.

Ce théorème est analogue au théorème 4.2.4, au moins en ce qui concerne les conclusions : il affirme que, si la distribution de f_Y est dans le domaine d'attraction d'une certaine loi normale ou stable, alors les sommes de Birkhoff de f convenablement renormalisées convergent vers cette loi, la normalisation étant la même que dans le cas indépendant.

Il y a cependant des différences entre ces deux théorèmes : le théorème 4.2.4 fait intervenir des hypothèses sur le temps de retour, tandis qu'il n'y a aucune hypothèse de ce type dans le théorème 4.4.1. D'un autre côté, les hypothèses de régularité sont un peu plus fortes dans le théorème 4.4.1 : d'une part, il y a la condition supplémentaire $\sum_{a\in\alpha_0}\varphi_Y(a)m[a]D_\tau f_Y(a)^\eta < \infty$, et d'autre part la première condition dans la définition de $I(X,Y)$ requiert que f soit bornée et hölderienne sur toute la tour (même si les constantes peuvent dépendre de la hauteur dans la tour). Cependant, dans la pratique, toutes ces conditions sont vérifiées, si bien qu'on peut souvent appliquer indifféremment les théorèmes 4.2.4 ou 4.4.1. Par exemple, tous les exemples du paragraphe 4.2.3 peuvent être étudiés à l'aide du théorème 4.4.1.

En revanche, les techniques de preuve sont très différentes : dans les deux cas, on utilise des informations sur l'application induite T_Y, mais c'est le seul point commun aux deux arguments. Il ne faut donc pas s'étonner que les possibilités de généralisation soient très différentes : on verra dans le chapitre 5 que l'approche du théorème 4.4.1 est très féconde pour démontrer des résultats limites plus subtils, qui ne semblent pas accessibles par des méthodes élémentaires comme celle du théorème 4.2.4.

Remarque 4.4.2. Le théorème reste en fait valable pour des classes de fonctions un peu plus grandes que $I(X,Y)$: au cours de la preuve, on verra que ce qui est important est qu'il existe des parties de l'espace de mesure arbitrairement grandes sur lesquelles f induit des fonctions assez régulières. Lorsque $f \in I(X,Y)$, on prend pour ces parties $\bigcup_0^{N-1} T^i(Y)$ pour $N \to \infty$, mais d'autres choix sont possibles.

4.4.1 Réduction au cas mélangeant

Pour démontrer le théorème 4.4.1, on va appliquer le théorème 4.3.1 aux opérateurs de transfert de premier retour. Pour cela, il faut qu'ils forment une suite d'opérateurs de renouvellement, ce qui n'est vrai que lorsque $T : X \to X$ est Markov-mélangeante. Dans ce paragraphe, on va voir qu'il suffit de traiter le cas Markov-mélangeant puisqu'il implique le cas Markov-transitif :

Proposition 4.4.3. *Si le théorème 4.4.1 est vrai pour les tours Markov-mélangeantes, alors il est également vrai pour les tours Markov-transitives.*

Démonstration. Soient $T : X \to X$ et $f : X \to \mathbb{R}$ vérifiant les hypothèses du théorème 4.4.1. D'après la proposition 1.3.14, il existe un entier $d \geqslant 1$ et une partition $X_0, \dots, X_{d-1}$ de X telle que T envoie X_i dans X_{i+1} pour $0 \leqslant i \leqslant d-1$, et telle que la restriction de T^d à chaque X_i est une tour localement Gibbs-Markov mélangeante. La base de la tour $T^d : X_i \to X_i$ est obtenue en prenant les éléments de α contenus dans X_i et de hauteur au plus $d-1$ dans la tour X.

On traite tout d'abord le troisième cas du théorème 4.4.1, i.e. on suppose que $m(f_Y > x) = (c_1 + o(1))x^{-p}L(x)$ et $m(f_Y < -x) = (c_2 + o(1))x^{-p}L(x)$. Soit $0 \leqslant i \leqslant d-1$, et notons f_{Y_i} la fonction induite par $S_d f = \sum_0^{d-1} f \circ T^k$ sur la base Y_i de la tour X_i. Soit φ_i le temps de retour de Y_i à lui-même pour l'application T^d, alors $f_{Y_i}(y) = \sum_{k=0}^{d\varphi_i(y)-1} f(T^k y)$. Soit $a \in \alpha_0$. Si $\varphi_Y(a) < d$, alors f_Y est bornée sur $a \times \{0\}$ par $d\,\|f\|_\infty$, et f_{Y_i} est également bornée par la même constante sur l'éventuel élément de Y_i de la forme $a \times \{k\}$ avec $k < d-1$. Sinon, soit $k < d-1$ tel que $a \times \{k\} \subset Y_i$. Pour $y \in a$, on a $f_Y(y,0) = f_{Y_i}(y,k) \pm d\,\|f\|_\infty$. Cela implique que f_{Y_i} vérifie aussi $m(f_{Y_i} > x) = (c_1 + o(1))x^{-p}L(x)$ et $m(f_{Y_i} < -x) = (c_2 + o(1))x^{-p}L(x)$. De la même façon, on vérifie que $f_{Y_i} \in I(X_i, Y_i)$ en utilisant que $f_Y \in I(X, Y)$. Ainsi, on peut appliquer le théorème 4.4.1 à la fonction $(S_d f)_{|X_i}$ dans la tour localement Gibbs-Markov mélangeante $T^d : X_i \to X_i$ munie de la mesure de probabilité $m_i = \frac{m_{|X_i}}{m(X_i)}$, et on obtient que, pour tout intervalle I de $\mathbb{R}$,

$$m\{y \in X_i \mid S_{dn}f(y)/B_{dn} \in I\} \to m(X_i)P(Z \in I)$$

où $Z = Z_{p,c,\beta}$ est comme dans la conclusion du théorème 4.4.1. En sommant sur les différents X_i, on trouve que $S_{dn}f/B_{dn} \to Z$ sur X. On étend ensuite ce résultat à des temps arbitraires, de la forme $dn + r$ avec $0 \leqslant r \leqslant d-1$, en remarquant que $S_{dn}f$ et $S_{dn+r}f$ sont à distance bornée l'un de l'autre sur des ensembles de mesure arbitrairement proche de 1, tandis que $B_n \to +\infty$.

Le deuxième cas du théorème est analogue.

Dans le cas du théorème central limite, on montre par le même type d'argument que $f_{Y_i} \in L^2(Y_i)$ pour tout i, si bien que $S_{dn}f/\sqrt{dn}$ converge en loi sur X_i vers une gaussienne $\mathcal{N}(0, \sigma_i^2)$, mais la variance dépend *a priori* de i. Pour $x \in X_i$, on a $\frac{1}{\sqrt{dn}} S_{dn}f(Tx) = \frac{1}{\sqrt{dn}} S_{dn}f(x) + \frac{f(T^{dn}x) - f(x)}{\sqrt{dn}}$. Comme $\frac{f(T^{dn}x)-f(x)}{\sqrt{dn}}$ tend vers 0 en probabilité, on en déduit que les lois limites de $S_{dn}f/\sqrt{dn}$ sur X_i et sur X_{i+1} sont les mêmes, i.e. que $\sigma_i = \sigma_{i+1}$. En notant σ la valeur commune des σ_i, on conclut comme plus haut que $S_n f/\sqrt{n} \to \mathcal{N}(0, \sigma^2)$.

Finalement, si $\sigma^2 = 0$, alors $S_d f$ est un cobord sur X_0, i.e. il s'écrit $\chi - \chi \circ T^d$ pour une certaine fonction mesurable $\chi : X_0 \to \mathbb{R}$. On étend χ à tout X en posant, pour $x \in X_0$ et $0 \leqslant k \leqslant d-1$, $\chi(T^k x) = \chi(x) - \sum_0^{k-1} f(T^k x)$. On vérifie alors que $f = \chi - \chi \circ T$ sur X, et f est donc un cobord pour T. □

La suite de cette partie est consacrée à la démonstration du théorème 4.4.1 dans le cas Markov-mélangeant. En fait, on va essentiellement s'intéresser au cas du théorème central limite, i.e. le cas $f_Y \in L^2$, et on verra dans un dernier paragraphe comment traiter les deux autres cas. On va commencer par travailler sur Y, avant de voir qu'on peut en fait faire le même raisonnement sur $\bigcup_0^{N-1} T^i(Y)$.

4.4.2 Le résultat local sur Y

Dans ce paragraphe, on suppose que f satisfait les hypothèses du premier cas du théorème 4.4.1, i.e. $f \in I(X,Y)$ et $f_Y \in L^2(Y)$.

Rappelons que les opérateurs de transfert de premier retour R_n ont été définis dans la partie 2.1, où on a également vu qu'ils formaient une suite d'opérateurs de renouvellement puisque T est Markov-mélangeante. On perturbe ces opérateurs en posant $R_n(t) = R_n(e^{itf_Y}\cdot) = R_n(e^{itS_nf}\cdot)$. On perturbe également les opérateurs T_n de la partie 2.1 par $T_{n,t} = T_n(e^{itS_nf}\cdot)$. Comme les sommes de Birkhoff de f pour T et de f_Y pour T_Y coïncident, l'équation du renouvellement (2.3) est encore valide pour ces perturbations, i.e.

$$T_{n,t} = \sum_{k_1+\ldots+k_l=n} R_{k_1}(t)\ldots R_{k_l}(t).$$

Cette équation peut aussi s'écrire $\sum T_{n,t}z^n = \left(I - \sum R_n(t)z^n\right)^{-1}$.

Pour des raisons techniques, on va faire agir les opérateurs non pas sur $\mathcal{L}_\tau(Y)$ l'espace des fonctions lipschitziennes pour la distance d_τ, mais plutôt sur $\mathcal{L}_{\tau^\eta}(Y)$ où η est comme dans la définition de $I(X,Y)$. Comme $\tau^\eta \geqslant \tau$, la distorsion de T_Y est également τ^η-hölderienne, si bien que l'application T est encore une tour localement Gibbs-Markov pour la distance d_{τ^η}. En particulier, toutes les estimées sur les tours localement Gibbs-Markov continuent à s'appliquer. Par exemple, le lemme 2.1.2 implique que $\|R_n\|_{\mathcal{L}_{\tau^\eta}} = O(m[\varphi_Y = n])$.

Lemme 4.4.4. *Il existe des constantes r_n et $\alpha > 0$ telles que, $\forall n \geqslant 1, \forall t \in [-\alpha,\alpha]$, $\|R_n(t)\|_{\mathcal{L}_{\tau^\eta}} \leqslant r_n$, avec $\sum nr_n < \infty$.*

De plus, pour tout $n \geqslant 1$, l'application $t \mapsto R_n(t)$ est continue en $t = 0$.

Démonstration. On va utiliser les opérateurs $M_{\underline{a}_Y,t}$ introduits dans le paragraphe 4.1.1, pour l'application Gibbs-Markov T_Y. On a $R_n(t) = \sum M_{\underline{a}_Y,t}$ où la somme porte sur tous les cylindres $\underline{a}_Y = [a]_Y$ avec $\varphi_Y(a) = n$. En additionnant les estimées données par le lemme 4.1.2, on obtient

$$\begin{aligned}\|R_n(t)\|_{\mathcal{L}_{\tau^\eta}} &\leqslant \|R_n\|_{\mathcal{L}_{\tau^\eta}} + B\left(\sum_{a\in\alpha_0,\varphi_Y(a)=n} m[a]D_\tau f_Y(a)^\eta + \int_{\{\varphi_Y=n\}} |e^{itf_Y}-1|\,\mathrm{d}m\right)\\ &\leqslant Cm[\varphi_Y = n] + B\sum_{a\in\alpha_0,\varphi_Y(a)=n} m[a]D_\tau f_Y(a)^\eta + 2Bm[\varphi_Y = n] =: r_n.\end{aligned}$$

On sait que $\sum_{a\in\alpha_0}\varphi_Y(a)m(a)D_\tau f_Y(a)^\eta<\infty$ par définition de $I(X,Y)$, et que $\sum nm[\varphi_Y=n]=m(X)<\infty$ par la formule de Kac. Cela montre que $\sum nr_n<\infty$.

Finalement, pour obtenir la continuité de $t\mapsto R_n(t)$ en $t=0$, on somme encore les estimées du lemme 4.1.2, mais cette fois dans l'espace $\mathcal{L}_{\tau^\eta}$ (i.e. les quantités τ et η du lemme sont prises égales respectivement à τ^η et 1). La somme $\sum m[a]D_{\tau^\eta}f_Y(a)$ est finie puisqu'elle est majorée par $\sum m[a]D_\tau f_Y(a)$ qui est elle-même finie par définition de $I(X,Y)$. On obtient $\|R_n(t)-R_n\|_{\mathcal{L}_{\tau^\eta}}\leqslant B|t|$. □

On va appliquer le théorème 4.3.1 aux opérateurs $R_n(t)$ agissant sur $\mathcal{L}_{\tau^\eta}$. Il faut pour cela un contrôle de la norme de $R_n(t)$, donné par le lemme 4.4.4, et d'autre part un contrôle de la valeur propre $\lambda(1,t)$ de $R(1,t)=\sum R_n(t)=\widehat{T}_Y(e^{itf_Y}\cdot)$. Comme T_Y est Gibbs-Markov et $f_Y\in L^2$, le théorème 4.1.4 assure que $\lambda(t)=1-\frac{\sigma^2}{2m(Y)}t^2+o(t^2)$ pour un certain $\sigma^2\geqslant 0$. Notons que cette normalisation en $\frac{\sigma^2}{2m(Y)}$ est la normalisation naturelle étant donné que $m(Y)<1$.

Si $\sigma^2\neq 0$, toutes les hypothèses du théorème 4.3.1 sont satisfaites pour $\Lambda(t)=t^2$ et $c=\frac{\sigma^2}{2m(Y)}$, et on obtient :

Proposition 4.4.5. *Supposons que $\sigma^2>0$. Alors, $\forall t\in\mathbb{R}$,*

$$\int_{T^{-n}Y\cap Y}e^{itS_nf/\sqrt{n}}\,\mathrm{d}m\to m(Y)^2e^{-\sigma^2t^2/2}.$$

Démonstration. Si t est fixé, $t/\sqrt{n}\in[-\varepsilon_0,\varepsilon_0]$ pour n assez grand, si bien qu'on peut appliquer le théorème 4.3.1. La constante μ est égale à $\frac{1}{m(Y)}$ d'après le lemme 2.1.5.

D'une part,

$$\begin{aligned}\int T_{n,t/\sqrt{n}}1\,\mathrm{d}m&=\int 1_Y\widehat{T}^n(e^{itS_nf/\sqrt{n}}1_Y)\,\mathrm{d}m=\int 1_Y\circ T^n\cdot e^{itS_nf/\sqrt{n}}1_Y\,\mathrm{d}m\\&=\int_{T^{-n}Y\cap Y}e^{itS_nf/\sqrt{n}}\,\mathrm{d}m,\end{aligned}$$

et d'autre part

$$\int T_{n,t/\sqrt{n}}1\,\mathrm{d}m=\int\left(1-\frac{\sigma^2}{2\mu m(Y)}\frac{t^2}{n}\right)^n\frac{1}{\mu}P(1)\,\mathrm{d}m+o(1)=\frac{m[Y]}{\mu}e^{-\sigma^2t^2/2}+o(1)$$

en utilisant $P(1)=1_Y$, et $1/\mu=m(Y)$ donc $\mu m(Y)=1$. □

Si $\sigma^2=0$, on ne peut pas utiliser le théorème 4.3.1, mais on a néanmoins le résultat suivant :

Proposition 4.4.6. *Supposons que $\sigma^2=0$. Alors f est un cobord sur X, i.e. il existe une fonction mesurable $\chi:X\to\mathbb{R}$ telle que $f=\chi-\chi\circ T$.*

Démonstration. Le théorème 4.1.4 montre que, si $\sigma^2 = 0$, alors f_Y est un cobord, i.e. on peut l'écrire sous la forme $f_Y = \psi - \psi \circ T_Y$ pour une certaine fonction $\psi : Y \to \mathbb{R}$. Il faut revenir de f_Y à f.

On étend ψ à tout l'espace X par $\psi(x) = 0$ si $x \notin Y$. Alors, en posant $\widetilde{f} = f - \psi + \psi \circ T$, on a pour $x \in Y$

$$\sum_0^{\varphi_Y(x)-1} \widetilde{f}(T^i x) = f_Y(x) - \psi(x) + \psi(T_Y x) = 0.$$

Posons $\delta(x) = \sum_0^{\varphi_Y(x)-1} \widetilde{f}(T^i x)$, où φ_Y désigne le premier retour (respectivement la première entrée) dans Y si $x \in Y$ (respectivement $x \notin Y$). Alors, si $Tx \notin Y$, on a $\widetilde{f}(x) = \delta(x) - \delta(Tx)$, tandis que si $Tx \in Y$ alors $\delta(Tx) = 0$ et $\delta(x) = \widetilde{f}(x)$ donc $\widetilde{f}(x) = \delta(x) - \delta(Tx)$. Ainsi, l'égalité $\widetilde{f} = \delta - \delta \circ T$ a lieu partout, puis $f = (\psi + \delta) - (\psi + \delta) \circ T$. □

4.4.3 Le résultat local sur $\bigcup_{i=0}^{N-1} T^i(Y)$

On suppose encore que $f \in I(X, Y)$ et que $f_Y \in L^2$. La proposition 4.4.5 décrit alors $\int_{T^{-n}Y \cap Y} e^{itS_n f/\sqrt{n}}$, alors qu'on aurait besoin de connaître $\int e^{itS_n f/\sqrt{n}}$ pour démontrer le premier cas du théorème 4.4.1. Pour estimer cette intégrale, on va utiliser des parties de plus en plus grandes de X auxquelles on pourra encore appliquer (une version légèrement modifiée de) la proposition 4.4.5.

Soit $Z = Z_N = \bigcup_{i=0}^{N-1} T^i(Y)$ la réunion des N premiers étages de la tour. Alors l'application T_Z induite par T sur Z est markovienne pour la partition α_N obtenue en restreignant α à Z, mais elle n'est pas Gibbs-Markov, puisque les images des éléments de la base n'ont pas une mesure uniformément minorée. Cependant, sur les éléments qui n'ont pas une grande image (i.e. ceux qui ne reviennent pas à la base de la tour), il n'y a pas de distorsion : T_Z est un isomorphisme qui préserve la mesure entre chacun de ces éléments et leurs images. Cela permet de faire marcher toutes les preuves sur les applications Gibbs-Markov, en distinguant à chaque fois les éléments qui reviennent à la base et les autres. Plus précisément, en répétant avec de légères variations les preuves dans le cas Gibbs-Markov, on peut démontrer les résultats suivants, qui sont les analogues dans ce cadre des résultats valables pour les applications Gibbs-Markov. Les preuves détaillées sont données dans [Gou04a].

1. [Corollaire 1.1.14] Notons $\mathcal{L}_\tau(Z)$ les fonctions sur Z qui sont lipschitziennes pour la distance $d_\tau(x, y) = \tau^{s(x,y)}$ sur chaque élément de α_N, et bornées. Alors $\widehat{T}_Z$ agit continûment sur $\mathcal{L}_\tau(Z)$. De plus, pour tout $n \geqslant N$, $\widehat{T}_Z^n$ vérifie une inégalité de Doeblin-Fortet.
   Pour la preuve de la continuité, on traite séparément les points qui reviennent à la base et les autres. Pour l'inégalité de Doeblin-Fortet, on utilise le fait que tous les éléments de la partition retournent à la base en temps $< N$, ce qui donne une grande image.
2. [Proposition 1.1.15] Soit $u : Z \to \mathbb{R}$ intégrable avec $\sum_{a \in \alpha_N} m[a] Du(a) < \infty$ et qui, sur l'ensemble des points qui ne retournent pas à la base, est bornée et lipschitzienne pour

d_τ. Alors $\widehat{T}_Z u \in \mathcal{L}_\tau(Z)$.

3. [Proposition 1.1.17] L'opérateur $\widehat{T}_Z$ agissant sur $\mathcal{L}_\tau(Z)$ a une valeur propre simple isolée en 1, le projecteur spectral correspondant étant donné par $Pu = \frac{\int_Z u}{m[Z]}$.
4. Les opérateurs de premier retour $R_n(Z)$ forment encore une suite d'opérateurs de renouvellement.
   La seule petite différence dans les arguments est que, dans la preuve du lemme 2.1.4, $R(z)^n$ ne vérifie une inégalité de Doeblin-Fortet que pour $n \geqslant N$.
5. Le théorème 4.1.4 est valable pour toute fonction $h : Z \to \mathbb{R}$ avec $\sum_{a\in\alpha_N} m[a]Dh(a) < \infty$, $h \in L^2(Z)$ et $\int h = 0$, et qui, sur l'ensemble des points qui ne retournent pas à la base, est bornée et lipschitzienne pour d_τ.
   En effet, la même preuve s'applique encore, en notant que les hypothèses supplémentaires sur h impliquent que $\widehat{T}_Z h \in \mathcal{L}_\tau(Z)$ d'après le point 2.

Pour pouvoir appliquer les arguments du paragraphe 4.4.2, on aura en plus besoin du lemme suivant :

Lemme 4.4.7. *Il existe $\tau < 1$ tel que l'application f_Z induite par f sur Z vérifie les propriétés suivantes :*

$$\sum_{a\in\alpha_N} \varphi_Z(a) m[a] D_\tau f_Z(a)^\eta < \infty$$

et

$$\sum_{a\in\alpha_N} m[a] D_{\tau^\eta} f_Y(a) < \infty.$$

Démonstration. On sait que ces propriétés sont vérifiées par f_Y sur Y. On va démontrer la première, la seconde étant analogue.

Soient A l'ensemble des éléments de la partition de Z qui ne retournent pas à la base Y de la tour à la prochaine itération de T_Z, et B l'ensemble de ceux qui reviennent à la base. Si $a \in A$, alors $\varphi_Z = 1$ sur a et $f_Z = f$ est uniformément τ_N-hölderienne sur ces ensembles, d'après la définition de $I(X, Y)$. Par conséquent, si $\tau \geqslant \tau_N$,

$$\sum_{a\in A} \varphi_Z[a] m[a] D_\tau f_Z(a)^\eta \leqslant C \sum_{a\in A} m[a] \leqslant C m[Z] < \infty.$$

Soient $a \in B$ et $x, y \in a$. Soient x' et y' les points dans la base de la tour correspondant à x et y, et a' l'élément de partition correspondant : $x = T^i x'$ et $y = T^i y'$ pour un certain $i < N$. Alors $f_Y(x') = \sum_0^{i-1} f(T^j x') + f_Z(x)$, et on a la même expression pour y, donc si $\tau \geqslant \tau_N$

$$\begin{aligned} |f_Z(x) - f_Z(y)| &\leqslant |f_Y(x') - f_Y(y')| + \sum_{j=0}^{i-1} |f(T^j x') - f(T^j y')| \\ &\leqslant C D_\tau f_Y(a') \tau^{s(x',y')} + CN\tau^{s(x',y')} \leqslant (C D_\tau f_Y(a') + CN)\tau^{s(x,y)} \end{aligned}$$

en utilisant $s(x',y') = s(x,y)$. Comme $\varphi_Z(a) \leqslant \varphi_Y(a')$, on obtient

$$\sum_{a\in B} \varphi_Z[a]m[a]D_\tau f_Z(a)^\eta \leqslant \sum_{a'\in\alpha_0} \varphi_Y[a']m[a'](CD_\tau f_Y(a') + CN)^\eta.$$

Cette dernière somme est finie par la formule de Kac et l'hypothèse sur f_Y. □

Les estimées du lemme 4.4.7 permettent alors de démontrer l'analogue sur Z du lemme 4.4.4.

De plus, on démontre de la même manière que $f_Z \in L^2(Z)$. La fonction f, appartenant à $I(X,Y)$, est bornée et lipschitzienne sur Z. Comme $f_Z(x) = f(x)$ si x ne retourne pas à la base de la tour, f_Z est bornée et lipschitzienne sur l'ensemble des points de Z dont l'image par T_Z n'est pas dans la base de la tour. Le théorème 4.1.4 s'applique donc à f_Z, et il montre que la valeur propre proche de 1 de $\widehat{T}_{Z,t}$ peut se mettre sous la forme $\lambda_Z(t) = 1 - \frac{\sigma_N^2}{2m(Z)}t^2 + o(1)$ pour un certain $\sigma_N \geqslant 0$.

On a vérifié toutes les hypothèses nécessaires pour appliquer le théorème 4.4.1 aux opérateurs de transfert de premier retour $R_n(Z)$, dans le cas $\sigma_N > 0$. On démontre donc de la même façon que dans le paragraphe 4.4.2 l'analogue suivant des propositions 4.4.5 et 4.4.6 :

Proposition 4.4.8. *Si $\sigma_N = 0$, alors f est un cobord sur X.*
Sinon, $\int_{Z\cap T^{-n}Z} e^{itS_nf/\sqrt{n}}\,\mathrm{d}m \to m(Z)^2 e^{-\sigma_N^2 t^2/2}$.

4.4.4 Démonstration du théorème central limite

Dans ce paragraphe, on démontre le premier cas du théorème 4.4.1. Pour cela, on aura besoin de voir que la variance σ_N sur Z_N ne dépend pas de N :

Proposition 4.4.9. *σ_N est indépendant de N.*

Démonstration. On fixe $N > 1$, et on va montrer que $\sigma_N = \sigma_1$. On écrira $Z = Z_N$ et $f_Z(x) = \sum_0^{\varphi_Z(x)-1} f(x)$. Rappelons que f_Y désigne la même fonction, mais sur la base Y et avec le temps de retour φ_Y. On notera $\widetilde{\varphi}$ le temps de retour de Y à lui-même pour l'application T_Z induite par T sur Z. Avec ces notations, pour $x \in Y$, $f_Y(x) = \sum_{j<\widetilde{\varphi}(x)} f_Z(T_Z^j x)$.

Le théorème 4.1.4 donne que $\sigma_1^2 = \int f_Y^2 + 2\int af_Y$ où $a = (I - \widehat{T}_Y)^{-1}(\widehat{T}_Y f_Y)$ et $\widehat{T}_Y$ est l'opérateur de transfert associé à T_Y. De la même façon, $\sigma_N^2 = \int f_Z^2 + 2\int bf_Z$ où $b = (I - \widehat{T}_Z)^{-1}(\widehat{T}_Z f_Z)$.

On définit une fonction c sur Z par $c(x,i) = a(x) + \sum_{j=0}^{i-1} f_Z(T_Z^j x)$ pour x dans la base de la tour et $i < \widetilde{\varphi}(x)$. Alors, si ψ est une fonction bornée sur Z, on va montrer que

$$\int_Z c\cdot(\psi - \psi\circ T_Z) = \int_Z f_Z\cdot\psi\circ T_Z. \tag{4.26}$$

Notons que la définition de a implique que, pour toute fonction χ sur Y, $\int_Y a\cdot(\chi-\chi\circ T_Y) = \int_Y f_Y\cdot\chi\circ T_Y$. Ainsi,

$$\begin{aligned}
\int_Z c\cdot(\psi-\psi\circ T_Z) &= \int_Y \sum_{i<\widetilde{\varphi}(x)} \left(a(x)+\sum_{j<i} f_Z(T_Z^j x)\right)(\psi(T_Z^i x)-\psi(T_Z^{i+1}x)) \\
&= \int_Y a(x)(\psi(x)-\psi(T_Y x)) + \sum_{j<i<\widetilde{\varphi}(x)} f_Z(T_Z^j x)(\psi(T_Z^i x)-\psi(T_Z^{i+1}x)) \\
&= \int_Y f_Y(x)\psi(T_Y x) + \sum_{j<\widetilde{\varphi}(x)-1} f_Z(T^j x)(\psi(T_Z^{j+1}x)-\psi(T_Y x)) \\
&= \int_Y \sum_{j<\widetilde{\varphi}(x)} f_Z(T_Z^j x)\psi(T_Y x) + \sum_{j<\widetilde{\varphi}(x)-1} f_Z(T^j x)(\psi(T_Z^{j+1}x)-\psi(T_Y x)) \\
&= \int_Y \sum_{j<\widetilde{\varphi}(x)-1} f_Z(T_Z^j x)\psi(T_Z^{j+1}x) + f_Z(T_Z^{\widetilde{\varphi}(x)-1}x)\psi(T_Y x) \\
&= \int_Y \sum_{j<\widetilde{\varphi}(x)} f_Z(T_Z^j x)\psi(T_Z^{j+1}x) = \int_Z f_Z(x)\psi(T_Z x).
\end{aligned}$$

L'équation (4.26) implique que $(I-\widehat{T}_Z)c = \widehat{T}_Z f_Z$. Par définition de b, on a aussi $(I-\widehat{T}_Z)b = \widehat{T}_Z f_Z$. Ainsi, $c-b$ est dans le noyau de $I-\widehat{T}_Z$, c'est donc une constante K, et $c = b+K$. On calcule finalement σ_1^2, en utilisant $\int f_Z = 0$:

$$\begin{aligned}
\sigma_1^2 &= \int_Y f_Y^2 + 2f_Y a = \int_Y \left(\sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)\right)^2 + 2\sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)a(x) \\
&= \int_Y \sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)^2 + 2\sum_{j<i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)f_Z(T_Z^j x) \\
&\qquad + 2\sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)\left(c(T_Z^i x) - \sum_{j<i} f_Z(T_Z^j x)\right) \\
&= \int_Y \sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)^2 + 2\sum_{i<\widetilde{\varphi}(x)} f_Z(T_Z^i x)c(T_Z^i x) = \int_Z f_Z(x)^2 + 2f_Z(x)c(x) \\
&= \int_Z f_Z(x)^2 + 2f_Z(x)(b(x)+K) = \int_Z f_Z(x)^2 + 2f_Z(x)b(x) = \sigma_N^2. \qquad \square
\end{aligned}$$

Conclusion de la preuve du premier cas du théorème 4.4.1. Soit σ^2 le nombre égal à tous les σ_N, comme le garantit la proposition 4.4.9.

Si $\sigma^2 = 0$, alors la proposition 4.4.6 montre que f est un cobord. Ainsi, $S_n f/\sqrt{n}$ tend vers 0 en probabilité, et le théorème est prouvé.

Supposons maintenant que $\sigma^2 > 0$. Soit $t \in \mathbb{R}$, on veut montrer que $\int_X e^{itS_n f/\sqrt{n}} \to e^{-\sigma^2 t^2/2}$. Soit $\varepsilon > 0$. Soit $Z = Z_N$ avec N assez grand pour que $m(Z) \geqslant 1 - \varepsilon$. La proposition 4.4.8 appliquée dans Z_N montre qu'il existe n_0 tel que, pour $n \geqslant n_0$,

$$\left| \int_{T^{-n}Z \cap Z} e^{itS_n f/\sqrt{n}} - m(Z)^2 e^{-\sigma^2 t^2/2} \right| \leqslant \varepsilon.$$

De plus,

$$\begin{aligned} \left| \int_X e^{itS_n f/\sqrt{n}} - \int_{T^{-n}Z \cap Z} e^{itS_n f/\sqrt{n}} \right| &\leqslant m(X \backslash (Z \cap T^{-n}Z)) \\ &\leqslant m(X \backslash Z) + m(X \backslash T^{-n}Z) = 2m(X \backslash Z) \leqslant 2\varepsilon. \end{aligned}$$

Finalement,

$$\left| e^{-\sigma^2 t^2/2} - m(Z)^2 e^{-\sigma^2 t^2/2} \right| \leqslant 1 - m(Z)^2 \leqslant 2\varepsilon.$$

Ces trois inégalités montrent que $\left| \int_X e^{itS_n f/\sqrt{n}} - e^{-\sigma^2 t^2/2} \right| \leqslant 5\varepsilon$, ce qui conclut la preuve. □

4.4.5 Preuve des autres cas du théorème 4.4.1

La preuve est quasiment la même que celle du théorème central limite. On ne considère que le cas $p \in]1, 2[$, le dernier cas étant analogue.

On travaille sur $Z = Z_N$ la réunion des N premiers étages de la tour. En posant $f_Z(x) = \sum_0^{\varphi_Z(x)-1} f(x)$, le lemme 4.4.7 montre que $\sum \varphi_Z[a] m[a] D_\tau f_Z(a)^\eta < \infty$ et $\sum m[a] D_\tau f_Z(a) < \infty$. Pour appliquer le théorème 4.1.7 pour obtenir le développement de la valeur propre $\lambda(t)$ de $\widehat{T}_{Z,t}$, on doit estimer $m[f_Z > x]$ et $m[f_Z < -x]$ pour x grand.

Sur la base Y de la tour, c_1 est caractérisé par $m[f_Y > x] = (c_1 + o(1))x^{-p}L(x)$. On va montrer qu'on a aussi $m[f_Z > x] = (c_1 + o(1))x^{-p}L(x)$. Soient Δ le dernier étage de la tour dans Z, et Λ sa projection sur la base Y. Alors f_Z est bornée sur $Z \backslash \Delta$ et f_Y est bornée sur $Y \backslash \Lambda$, donc il suffit de considérer Δ et Λ. Soient $x \in \Lambda$ et $y = T^{N-1}x$ son image dans Δ. Alors $f_Z(y) = f_Y(x) + \sum_{i=0}^{N-2} f(T^i x)$. Comme la somme est bornée par $C = N \left\| f_{|Z} \right\|_\infty$, on obtient $f_Z(y) = f_Y(x) \pm C$. Ainsi, si x est assez grand, $m[f_Y > x - C] \geqslant m[f_Z > x] \geqslant m[f_Y > x + C]$. Comme $m[f_Y > x] = (c_1 + o(1))x^{-p}L(x)$, on obtient la même estimée pour $m[f_Z > x]$ en utilisant la variation lente de L. De la même façon, $m[f_Z < -x] = (c_2 + o(1))x^{-p}L(x)$.

En utilisant le théorème 4.1.7 on obtient que $\lambda(t) = 1 - a|t|^p L(1/|t|) + ib\,\mathrm{sgn}(t)|t|^p L(1/|t|) + o(|t|^p L(1/|t|))$, pour $a = \frac{c}{m(Z)}$ et $b = \frac{c}{m(Z)}\beta \tan\left(\frac{p\pi}{2}\right)$.

On utilise ensuite la remarque 4.3.2, avec $\Lambda(t) = t^p L(1/t)$, pour obtenir

$$\left\| T_{n,t} - \frac{1}{\mu}\left(1 - c|t|^p L(1/|t|) + ic\beta \tan\left(\frac{p\pi}{2}\right) \mathrm{sgn}(t)|t|^p L(1/|t|)\right)^n P \right\| \leqslant \varepsilon(t) + \delta(n).$$

Si B_n est défini par $B_n^p = nL(B_n)$, on a

$$n\frac{|t|^p}{B_n^p}L\left(\frac{B_n}{|t|}\right) = |t|^p\frac{L(B_n/|t|)}{L(B_n)} \to |t|^p$$

puisque L est à variation lente. Ainsi,

$$T_{n,t/B_n} \to \frac{1}{\mu}e^{-c|t|^p+ic\beta\tan\left(\frac{p\pi}{2}\right)\operatorname{sgn}(t)|t|^p}P.$$

En appliquant ce résultat à la fonction 1 et en intégrant, on obtient

$$\int_{T^{-n}Z\cap Z} e^{itS_nf/B_n}\,\mathrm{d}m \to m(Z)^2e^{-c|t|^p(1-i\beta\operatorname{sgn}(t)\tan(\frac{p\pi}{2}))},$$

et l'exponentielle qui apparaît ici est la fonction caractéristique de la loi stable $Z_{p,c,\beta}$.

On fait ensuite tendre N vers $+\infty$ et on conclut la preuve comme celle du théorème central limite. □

Chapitre 5

Théorèmes de la limite locale et de Berry-Esseen

Sommaire

5.1 Un résultat spectral perturbatif quantitatif 190
5.1.1 Estimées techniques préliminaires 191
5.1.2 Démonstration du théorème 5.1.1 194
5.2 L'estimée clé 197
5.2.1 Perturbation des opérateurs de transfert 198
5.2.2 Démonstration du théorème 5.2.1 201
5.3 Le théorème de la limite locale 203
5.3.1 Problèmes de périodicité 203
5.3.2 Le cas apériodique 205
5.3.3 Le cas périodique 208
5.4 Le théorème de Berry-Esseen 208
5.5 Applications 211
5.5.1 Le cas Markov-transitif 211
5.5.2 Apériodicité et extensions 212
5.5.3 Applications ayant un modèle localement Gibbs-Markov 215
5.5.4 Exemples 217

Dans le chapitre précédent, on s'est intéressé aux théorèmes limites les plus simples pour les sommes de Birkhoff d'applications markoviennes, c'est-à-dire la convergence en loi de la suite $\frac{1}{B_n} S_n f$ pour une certaine normalisation $B_n \to +\infty$. En théorie des probabilités, il existe de nombreux raffinements de ces théorèmes limites, par exemple donnant la vitesse de convergence (théorème de Berry-Esseen), ou encore des renseignements plus précis sur certaines probabilités (théorème de la limite locale). Dans le cas des applications Gibbs-Markov, ces

théorèmes sont assez faciles à démontrer par des méthodes spectrales perturbatives en utilisant le développement de la valeur propre donné par le théorème 4.1.4 et le fait que toutes les erreurs sont exponentiellement petites (voir par exemple [Bro96]).

Dans ce chapitre, qui correspond essentiellement à l'article [Gou05], on va démontrer ces théorèmes dans le cas des tours localement Gibbs-Markov. Les preuves reposeront sur le développement perturbatif des opérateurs de transfert de premier retour, comme dans la partie 4.4, mais en contrôlant précisément les termes d'erreur. Ce contrôle proviendra d'une version quantitative du théorème 4.3.1, qu'on va démontrer dans la partie 5.1. On l'appliquera ensuite dans les parties 5.3 et 5.4 pour démontrer le théorème de la limite locale et le théorème de Berry-Esseen dans certaines tours localement Gibbs-Markov. Enfin, dans une dernière partie, on montrera comment en déduire le même type de résultats pour des applications non-uniformément dilatantes générales : contrairement aux théorèmes limites démontrés dans le chapitre précédent, la réduction au cas des tours localement Gibbs-Markov demande ici des arguments non triviaux, particulièrement pour vérifier certaines hypothèses d'apériodicité.

5.1 Un résultat spectral perturbatif quantitatif

Si $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ sont deux suites réelles, on note $(a_n)\star(b_n)$ la suite $c_n = \sum_{k=0}^n a_k b_{n-k}$. Si $(a_n)_{n\in\mathbb{Z}}$ et $(b_n)_{n\in\mathbb{Z}}$ sont deux suites sommables indexées par $\mathbb{Z}$, on pose aussi $(a_n)\star(b_n) = \sum_{k=-\infty}^{\infty} a_k b_{n-k}$.

Le but de cette partie est de démontrer le résultat suivant, qui nous permettra plus tard d'estimer des fonctions caractéristiques de la forme $\int e^{itS_n f}$ avec un terme d'erreur explicite.

Rappelons que les suites d'opérateurs de renouvellement ont été définies dans la définition 2.2.1.

Théorème 5.1.1. *Soit $\beta > 2$. Soit $(R_n)_{n\geqslant 1}$ une suite d'opérateurs de renouvellement sur un espace de Banach $\mathcal{L}$, avec $\sum_{k=n+1}^{\infty} \|R_k\| = O(1/n^\beta)$. En particulier, soit $\mu > 0$ tel que le projecteur spectral P associé à la valeur propre 1 de $R(1) = \sum R_n$ satisfait $PR'(1)P = \mu P$.*

Soit $\alpha > 0$. Pour tous $n \geqslant 1$ et $t \in [-\alpha, \alpha]$, soit $R_n(t)$ un opérateur sur $\mathcal{L}$ tel que $\sum_{k=n+1}^{\infty} \|R_k(t) - R_k\| \leqslant C|t|/n^{\beta-1}$. Pour $z \in \overline{\mathbb{D}}$ et $t \in [-\alpha, \alpha]$, notons $R(z,t) = \sum z^n R_n(t)$. Soit $\lambda(z,t)$ la valeur propre proche de 1 de $R(z,t)$, pour (z,t) proche de $(1,0)$. On suppose que $\lambda(1,t) = 1 - \Xi(t)$ avec $\Xi(t) \sim ct^2$ pour une certaine constante c telle que $\mathrm{Re}(c) > 0$.

Alors, pour t assez petit, $I - R(z,t)$ est inversible pour tout $z \in \mathbb{D}$. Notons son inverse $\sum T_{n,t} z^n$. Alors il existe $\varepsilon_0 > 0$, $d > 0$ et $C > 0$ tels que, pour tout $t \in [-\varepsilon_0, \varepsilon_0]$, pour tout $n \in \mathbb{N}^*$,

$$\left\| T_{n,t} - \frac{1}{\mu}\left(1 - \frac{1}{\mu}\Xi(t)\right)^n P \right\| \leqslant \frac{C}{n^{\beta-1}} + C|t| \left(\frac{1}{n^{\beta-1}}\right) \star \left(1 - dt^2\right)^n . \tag{5.1}$$

5.1.1 Estimées techniques préliminaires

La preuve du théorème 5.1.1 se fera essentiellement en travaillant dans certaines algèbres de Banach judicieusement choisies, qui satisferont le lemme de Wiener. Elles sont analogues aux algèbres introduites dans la partie 2.2.4.

Soit $\mathcal{A}$ une algèbre de Banach (la plupart du temps, on prendra $\mathcal{A} = \mathrm{Hom}(\mathcal{L}, \mathcal{L})$). Pour $\gamma > 1$, on note $\mathcal{O}_\gamma(\mathcal{A})$ l'ensemble des séries $\sum_{n\in\mathbb{Z}} A_n z^n$ où $A_n \in \mathcal{A}$, $\sum \|A_n\| < \infty$ et $\sup_{n\in\mathbb{Z}} |n|^\gamma \|A_n\| < \infty$. Cette définition est analogue à la définition 2.2.14 pour la suite $u_n = 1/n^\gamma$, la différence étant qu'on exige également un contrôle en $-\infty$. On montre comme dans la proposition 2.2.15 qu'il existe sur $\mathcal{O}_\gamma(\mathcal{A})$ une norme équivalente à $\sum \|A_n\| + \sup_{n\in\mathbb{Z}} |n|^\gamma \|A_n\|$ qui en fait une algèbre de Banach. De plus, cette algèbre de Banach satisfait le lemme de Wiener, la preuve du théorème 2.2.16 s'adaptant immédiatement.

On aura également besoin de l'algèbre de Banach $\mathcal{O}^+_\gamma(\mathcal{A})$ donnée par l'ensemble des séries $\sum A_n z^n \in \mathcal{O}_\gamma(\mathcal{A})$ avec $A_n = 0$ pour $n < 0$. C'est une sous-algèbre fermée de $\mathcal{O}_\gamma(\mathcal{A})$, et elle satisfait également un lemme de Wiener : si $\sum A_n z^n \in \mathcal{O}^+_\gamma(\mathcal{A})$ et si cette somme est inversible pour tout $z \in \overline{\mathbb{D}}$, alors cette série admet un inverse dans $\mathcal{O}^+_\gamma(\mathcal{A})$. La preuve du théorème 2.2.16 s'adapte en effet directement à ce cas.

Notation. *Soient $f : [-\alpha, \alpha] \times \mathbb{Z} \to \mathbb{R}_+$ pour un certain $\alpha > 0$ et $\mathcal{A}$ une algèbre de Banach. On note $O_\mathcal{A}(f(t,n))$ l'ensemble des séries $\sum_{n=-\infty}^{\infty} c_n(t) z^n$ où $c_n : [-\alpha, \alpha] \to \mathcal{A}$ est telle qu'il existe $\alpha' > 0$ et $C > 0$ tels que*

$$\forall t \in [-\alpha', \alpha'], \forall n \in \mathbb{Z},\ \|c_n(t)\| \leqslant C f(t,n).$$

On omettra souvent l'indice dans $O_\mathcal{A}$. Comme d'habitude, on écrira souvent $\sum c_n(t) z^n = O(f(t,n))$ au lieu de la notation formellement plus correcte $\sum c_n(t) z^n \in O(f(t,n))$. On écrira également $O(g(n))$ pour l'ensemble des séries de la forme $\sum A_n z^n$ avec $\|A_n\| \leqslant C g(n)$ pour une certaine constante C. C'est un cas particulier de la notation précédente, dans lequel les fonctions $f(t,n)$ sont indépendantes de t. Jusqu'à la fin de la partie 5.1, le symbole O se référera exclusivement à cette notation.

Remarque 5.1.2. Il ne faut pas confondre les notations O et $\mathcal{O}$: il y a des similarités, ce qui explique qu'on ait utilisé la même lettre, mais la notation calligraphique $\mathcal{O}$ indique en plus qu'on a affaire à une algèbre de Banach. En particulier, on peut y utiliser la continuité de l'inversion ou le lemme de Wiener.

Avec ces notations, le lemme de Wiener dans l'algèbre $\mathcal{O}_\gamma(\mathcal{A})$ se reformule comme suit : si $\sum A_n z^n = O(1/(|n|+1)^\gamma)$ pour $\gamma > 1$ et $\sum A_n z^n$ est inversible pour tout $z \in S^1$, alors $(\sum A_n z^n)^{-1} = O(1/(|n|+1)^\gamma)$. Le fait que $\mathcal{O}_\gamma(\mathcal{A})$ soit une algèbre de Banach implique aussi que, pour $\gamma > 1$,

$$O\left(\frac{1}{(|n|+1)^\gamma}\right) \star O\left(\frac{1}{(|n|+1)^\gamma}\right) \subset O\left(\frac{1}{(|n|+1)^\gamma}\right), \tag{5.2}$$

i.e., si deux séries $\sum_{n\in\mathbb{Z}} A_n z^n$ et $\sum_{n\in\mathbb{Z}} B_n z^n$ (avec $A_n, B_n \in \mathcal{A}$) vérifient $\sup_{n\in\mathbb{Z}}(|n|+1)^\gamma \|A_n\| < \infty$ et $\sup_{n\in\mathbb{Z}}(|n|+1)^\gamma \|B_n\| < \infty$, alors la série $\sum C_n z^n := (\sum A_n z^n)(\sum B_n z^n)$ satisfait aussi $\sup_{n\in\mathbb{Z}}(|n|+1)^\gamma \|C_n\| < \infty$.

Pour alléger les notations, on notera souvent t au lieu de $|t|$ dans ce qui suit. De manière équivalente, le lecteur peut considérer que les preuves sont écrites pour $t \geqslant 0$. On notera aussi $\frac{1}{|n|^\gamma}$ au lieu de $\frac{1}{(|n|+1)^\gamma}$, en négligeant le problème en $n = 0$.

Lemme 5.1.3. *Lorsque $\gamma > 1$ et $d > 0$,*

$$O\left(\frac{1}{|n|^\gamma}\right) \star O\left(1_{n\geqslant 0} t^2 (1-dt^2)^n\right) \subset O\left(\frac{1}{|n|^\gamma} + 1_{n\geqslant 0} t^2 \left(1 - \frac{d}{2}t^2\right)^n\right).$$

Démonstration. Pour $n < 0$, le n-ième terme de la convolution entre $\frac{1}{|n|^\gamma}$ et $1_{n\geqslant 0} t^2(1-dt^2)^n$ est majoré par $\sum_{k=0}^{\infty} t^2(1-dt^2)^k \frac{1}{|n|^\gamma} \leqslant \frac{C}{|n|^\gamma}$. Pour $n \geqslant 0$, il est majoré par

$$\sum_{k=-\infty}^{n/2} \frac{1}{|k|^\gamma} t^2 (1-dt^2)^{n/2} + \sum_{k=n/2}^{n} \frac{1}{(n/2)^\gamma} t^2 (1-dt^2)^{n-k} \leqslant C t^2 (1-dt^2)^{n/2} + \frac{C}{n^\gamma}.$$

Comme $\sqrt{1-dt^2} \leqslant 1 - \frac{d}{2}t^2$, on obtient la conclusion. □

Lemme 5.1.4. *Soient $\gamma > 1$ et $d > 0$. Soit $G_t(z) = O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 (1-dt^2)^n\right)$, et supposons que $F(z) = O(1/|n|^\gamma)$ est inversible pour tout $z \in S^1$. Alors $\left[F(z) + G_t(z)\right]^{-1} = F(z)^{-1} + O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 \left(1 - \frac{d}{64}t^2\right)^n\right)$.*

Démonstration. On suppose tout d'abord que $F(z) = 1$. Si on pose $H_t(z) = \sum_{n\in\mathbb{Z}} \frac{t}{|n|^\gamma} z^n + \sum_{n\in\mathbb{N}} t^3(1-dt^2)^n z^n$, la norme du coefficient de z^n dans $[1+G_t(z)]^{-1}$ est majorée par le coefficient de z^n dans $[1-H_t(z)]^{-1}$. Ainsi, il suffit de considérer $\frac{1}{1-tK(z)-\frac{t^3}{1-(1-dt^2)z}}$ où $K(z) = \sum \frac{z^n}{|n|^\gamma}$. Notons que

$$\frac{1}{1-tK(z)-\frac{t^3}{1-(1-dt^2)z}} = \frac{1}{1-tK(z)-t^3} \frac{1-(1-dt^2)z}{1-(1-dt^2)\left[1+\frac{t^3}{1-tK(z)-t^3}\right]z}. \tag{5.3}$$

Pour t assez petit, $\left|(1-dt^2)\left[1+\frac{t^3}{1-tK(z)-t^3}\right]\right| < 1$, donc

$$\begin{aligned}
&\frac{1-(1-dt^2)z}{1-(1-dt^2)\left[1+\frac{t^3}{1-tK(z)-t^3}\right]z} \\
&\qquad = \left(1-(1-dt^2)z\right) \sum_{n=0}^{\infty} (1-dt^2)^n \left[1+\frac{t^3}{1-tK(z)-t^3}\right]^n z^n \qquad (5.4)\\
&\qquad = 1 + \frac{tz(1-dt^2)}{1-tK(z)-t^3} t^2 \sum_{n=0}^{\infty} (1-dt^2)^n \left[1+\frac{t^3}{1-tK(z)-t^3}\right]^n z^n.
\end{aligned}$$

Étudions tout d'abord la somme. Soit $\mathcal{A} = \mathcal{O}_\gamma(\mathbb{C})$ l'algèbre de Banach des séries dont les coefficients sont dans $O(1/|n|^\gamma)$. Comme $K(z) \in \mathcal{A}$, on a $1 - tK(z) - t^3 = 1 + O_{\mathcal{A}}(t)$. Puisque l'inversion est lipschitzienne sur une algèbre de Banach, on obtient $\frac{t^3}{1-tK(z)-t^3} = t^3 + O_{\mathcal{A}}(t^4)$. Par conséquent, pour t assez petit, $\left\| \left[1 + \frac{t^3}{1-tK(z)-t^3}\right]^n \right\|_{\mathcal{A}} \leqslant C(1+2t^3)^n$. Lorsque t est assez petit pour que $(1-dt^2)(1+2t^3) \leqslant \left(1 - \frac{d}{2}t^2\right)$, le coefficient c_p de z^p dans $t^2 \sum_{n=0}^\infty (1-dt^2)^n \left[1 + \frac{t^3}{1-tK(z)-t^3}\right]^n z^n$ vérifie donc

$$\begin{aligned} c_p = t^2 \sum_{n=0}^\infty (1-dt^2)^n \left(\left[1 + \frac{t^3}{1-tK(z)-t^3}\right]^n\right)_{-n+p} &\leqslant C \sum_{n=0}^\infty t^2 (1-dt^2)^n \frac{(1+2t^3)^n}{|-n+p|^\gamma} \\ &\leqslant C \sum_{n=0}^\infty t^2 \left(1 - \frac{d}{2}t^2\right)^n \frac{1}{|-n+p|^\gamma}. \end{aligned}$$

On retrouve la même expression que dans la convolution entre $O\left(1_{n\geqslant 0} t^2 \left(1 - \frac{d}{2}t^2\right)^n\right)$ et $O(1/|n|^\gamma)$, qu'on a déjà estimée dans le lemme 5.1.3. Ainsi, on obtient

$$t^2 \sum_{n=0}^\infty (1-dt^2)^n \left[1 + \frac{t^3}{1-tK(z)-t^3}\right]^n z^n = O\left(\frac{1}{|p|^\gamma} + 1_{p\geqslant 0} t^2 \left(1 - \frac{d}{4}t^2\right)^p\right).$$

Comme $\frac{tz(1-dt^2)}{1-tK(z)-t^3} = O(t/|n|^\gamma)$, une autre convolution donne

$$(5.4) = 1 + O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 \left(1 - \frac{d}{8}t^2\right)^n\right).$$

En multipliant par $\frac{1}{1-tK(z)-t^3} = 1 + O\left(\frac{t}{|n|^\gamma}\right)$, on obtient

$$(5.3) = 1 + O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 \left(1 - \frac{d}{16}t^2\right)^n\right).$$

Cela conclut la preuve du cas $F(z) = 1$.

On traite maintenant le cas de $F(z)$ arbitraire. Notons que

$$\left[F(z) + G_t(z)\right]^{-1} = \left[1 + F(z)^{-1} G_t(z)\right]^{-1} F(z)^{-1}.$$

Par le lemme de Wiener, $F(z)^{-1} = O\left(\frac{1}{|n|^\gamma}\right)$, et le lemme 5.1.3 donne donc $F(z)^{-1} G_t(z) = O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 \left(1 - \frac{d}{2}t^2\right)^n\right)$. Ainsi, le cas $F(z) = 1$ implique que

$$\left[1 + F(z)^{-1} G_t(z)\right]^{-1} = 1 + O\left(\frac{t}{|n|^\gamma} + 1_{n\geqslant 0} t^3 \left(1 - \frac{d}{32}t^2\right)^n\right).$$

On obtient le résultat en convolant une dernière fois avec $F(z)^{-1}$. □

5.1.2 Démonstration du théorème 5.1.1

Soit $\Xi(t)$ comme dans le théorème 5.1.1. On fixe une fois pour toutes $d > 0$ tel que $|1-\frac{1}{\mu}\Xi(t)| \leqslant 1 - dt^2$ pour t assez petit, et on restreint t à un intervalle dans lequel cette inégalité est vraie. L'inversibilité de $R(z,t)$ pour $z \in \mathbb{D}$ et t assez petit est démontrée dans la proposition 4.3.4. On va utiliser la perturbation $\widetilde{R}(z,t)$ de $R(z,t)$ définie dans le paragraphe 4.3.2. Étudions tout d'abord sa valeur propre $\widetilde{\lambda}(z,t)$ proche de 1.

Lemme 5.1.5. *On a*

$$\frac{1-\widetilde{\lambda}(z,t)}{1-(1-\frac{1}{\mu}\Xi(t))z} = \frac{1-\widetilde{\lambda}(z)}{1-z} + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{2}t^2\right)^n\right).$$

Démonstration. Soient $K(z,t) = \frac{\widetilde{\lambda}(1,t)-\widetilde{\lambda}(z,t)}{1-z} - \frac{1-\widetilde{\lambda}(z)}{1-z}$ et $B(z) = \frac{1-\widetilde{\lambda}(z)}{1-z}$. Rappelons que $\widetilde{\lambda}(1,t) = \lambda(1,t) = 1-\Xi(t)$. On a alors

$$\widetilde{\lambda}(z,t) = 1 - \Xi(t) + (z-1)(K(z,t)+B(z)).$$

Ainsi,

$$\begin{aligned}\frac{1-\widetilde{\lambda}(z,t)}{1-(1-\frac{1}{\mu}\Xi(t))z} - B(z) &= \frac{\Xi(t) + (1-z)(K(z,t)+B(z)) - \left[1-(1-\frac{1}{\mu}\Xi(t))z\right]B(z)}{1-(1-\frac{1}{\mu}\Xi(t))z} \\ &= K(z,t)\frac{1-z}{1-(1-\frac{1}{\mu}\Xi(t))z} + \Xi(t)\frac{1-zB(z)/\mu}{1-(1-\frac{1}{\mu}\Xi(t))z} \\ &= I + II.\end{aligned}$$

Étudions tout d'abord I. On a

$$\frac{1-z}{1-(1-\frac{1}{\mu}\Xi(t))z} = 1 - \frac{1}{\mu}\Xi(t)\sum_{n=1}^{\infty}\left(1-\frac{1}{\mu}\Xi(t)\right)^{n-1}z^n = 1 + O(1_{n\geqslant 0}t^2(1-dt^2)^n). \tag{5.5}$$

Soit $\mathcal{A} = \mathcal{O}_{\beta-1}(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$ l'algèbre de Banach des fonctions dont les coefficients de Fourier sont dans $O(1/|n|^{\beta-1})$ pour $n \in \mathbb{Z}$. On a $\frac{\widetilde{R}(z,t)-\widetilde{R}(1,t)}{z-1} = \frac{\widetilde{R}(z)-\widetilde{R}(1)}{z-1} + O_{\mathcal{A}}(t)$. La preuve du lemme 4.3.3 s'applique dans l'algèbre $\mathcal{A}$ et avec des tildes partout, en utilisant le lemme de Wiener pour vérifier qu'on reste dans $\mathcal{A}$ quand on prend les inverses. Elle montre que $\frac{\widetilde{\lambda}(z,t)-\widetilde{\lambda}(1,t)}{z-1} = \frac{\widetilde{\lambda}(z)-\widetilde{\lambda}(1)}{z-1} + O_{\mathcal{A}}(t)$, i.e. $K(z,t) = O(t/|n|^{\beta-1})$. En multipliant cette estimée et (5.5) et en utilisant le lemme 5.1.3, on obtient

$$I = O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{2}t^2\right)^n\right).$$

Comme $\frac{\widetilde{R}(z)-\widetilde{R}(1)}{z-1} = O\left(\frac{1}{|n|^\beta}\right)$, on montre que $B(z) = O\left(\frac{1}{|n|^\beta}\right)$ en répétant la troisième étape de la preuve du théorème 2.2.5 dans l'algèbre de Banach $\mathcal{O}_\beta(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$. Comme $1-zB(z)/\mu$ s'annule en 1 d'après l'équation (2.12) et est dans $O\left(\frac{1}{|n|^\beta}\right)$, il peut s'écrire sous la forme $(1-z)D(z)$ avec $D(z) = O(1/|n|^{\beta-1})$.
Comme $\Xi(t) = O(t^2)$, l'équation (5.5) montre que $\Xi(t)\frac{1-z}{1-(1-\frac{1}{\mu}\Xi(t))z} = O(t^2 + 1_{n\geqslant 0}t^4(1-dt^2)^n)$. Le terme II est obtenu en multipliant cette expression et $D(z) = O(1/|n|^{\beta-1})$. Le lemme 5.1.3 montre donc que

$$II = O\left(\frac{t^2}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^4\left(1-\frac{d}{2}t^2\right)^n\right). \qquad \square$$

Corollaire 5.1.6. *On a*

$$\left(\frac{I-\widetilde{R}(z,t)}{1-\left(1-\frac{1}{\mu}\Xi(t)\right)z}\right)^{-1} = \left(\frac{I-\widetilde{R}(z)}{1-z}\right)^{-1} + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{256}t^2\right)^n\right).$$

Démonstration. Soient $\widetilde{P}(z,t)$ le projecteur spectral associé à la valeur propre $\widetilde{\lambda}(z,t)$ de $\widetilde{R}(z,t)$ et $\widetilde{Q}(z,t) = I - \widetilde{P}(z,t)$. Soit $\mathcal{A} = \mathcal{O}_{\beta-1}(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$. Alors $\widetilde{R}(z,t) = \widetilde{R}(z) + O_{\mathcal{A}}(t)$ puisque $R_n(t) = R_n + O(t/n^{\beta-1})$ par hypothèse. Comme $\mathcal{A}$ vérifie le lemme de Wiener, l'expression intégrale du projecteur spectral $\widetilde{P}(z,t)$ montre que $\widetilde{P}(z,t) = \widetilde{P}(z) + O_{\mathcal{A}}(t)$. De plus, $I - \widetilde{R}(z,t)\widetilde{Q}(z,t) = I - \widetilde{R}(z)\widetilde{Q}(z) + O_{\mathcal{A}}(t)$, donc $(I-\widetilde{R}(z,t)\widetilde{Q}(z,t))^{-1} = (I-\widetilde{R}(z)\widetilde{Q}(z))^{-1} + O_{\mathcal{A}}(t)$, puisque $I-\widetilde{R}(z)\widetilde{Q}(z)$ est inversible dans $\mathcal{A}$ et l'inversion est lipschitzienne.
Comme

$$I - \widetilde{R}(z,t) = (1-\widetilde{\lambda}(z,t))\widetilde{P}(z,t) + (I-\widetilde{R}(z,t)\widetilde{Q}(z,t))\widetilde{Q}(z,t),$$

les lemmes 5.1.4 et 5.1.5 impliquent que

$$\begin{aligned}
&\left(\frac{I-\widetilde{R}(z,t)}{1-(1-\frac{1}{\mu}\Xi(t))z}\right)^{-1} \\
&= \frac{1-(1-\frac{1}{\mu}\Xi(t))z}{1-\widetilde{\lambda}(z,t)}\widetilde{P}(z,t) + \left(1-\left(1-\frac{1}{\mu}\Xi(t)\right)z\right)(I-\widetilde{R}(z,t)\widetilde{Q}(z,t))^{-1}\widetilde{Q}(z,t) \\
&= \left[\frac{1-z}{1-\widetilde{\lambda}(z)} + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{128}t^2\right)^n\right)\right]\left[\widetilde{P}(z) + O\left(\frac{t}{|n|^{\beta-1}}\right)\right] \\
&\quad + \left[(1-z) + O(1_{n=0}t^2)\right]\left[(I-\widetilde{R}(z)\widetilde{Q}(z))^{-1} + O\left(\frac{t}{|n|^{\beta-1}}\right)\right]\left[\widetilde{Q}(z) + O\left(\frac{t}{|n|^{\beta-1}}\right)\right] \\
&= \frac{1-z}{1-\widetilde{\lambda}(z)}\widetilde{P}(z) + (1-z)(I-\widetilde{R}(z)\widetilde{Q}(z))^{-1}\widetilde{Q}(z) + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{256}t^2\right)^n\right)
\end{aligned}$$

$$= \left(\frac{I-\widetilde{R}(z)}{1-z}\right)^{-1} + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{256}t^2\right)^n\right). \qquad \square$$

Corollaire 5.1.7. *On a*

$$\left(\frac{I-R(z,t)}{1-\left(1-\frac{1}{\mu}\Xi(t)\right)z}\right)^{-1} = \left(\frac{I-R(z)}{1-z}\right)^{-1} + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{512}t^2\right)^n\right).$$

Démonstration. Soit χ_1, χ_2 une partition de l'unité C^∞ de S^1 telle que $R(z) = \widetilde{R}(z)$ sur le support de χ_1. Comme χ_1 est C^∞, $\chi_1(z) = O(1/|n|^{\beta-1})$. Si on note $A(z,t) = \left(\frac{I-R(z,t)}{1-\left(1-\frac{1}{\mu}\Xi(t)\right)z}\right)^{-1}$, le corollaire 5.1.6 implique donc que

$$\chi_1(z)A(z,t) = \chi_1(z)A(z,0) + O\left(\frac{t}{|n|^{\beta-1}} + 1_{n\geqslant 0}t^3\left(1-\frac{d}{512}t^2\right)^n\right).$$

En ce qui concerne χ_2, on peut modifier R hors du support de χ_2 de sorte que $I-R(z)$ soit partout inversible sur S^1. Soit $\mathcal{A} = \mathcal{O}_{\beta-1}(\mathrm{Hom}(\mathcal{L},\mathcal{L}))$. Comme $I-R(z,t) = I-R(z)+O_{\mathcal{A}}(t)$, le lemme de Wiener implique que $(I-R(z,t))^{-1} = (I-R(z))^{-1} + O_{\mathcal{A}}(t)$, donc

$$\chi_2(z)A(z,t) = \chi_2(z)A(z,0) + O\left(\frac{t}{|n|^{\beta-1}}\right). \qquad \square$$

Les fonctions apparaissant dans l'énoncé de ce corollaire sont en fait définies sur tout le disque $\overline{\mathbb{D}}$, i.e. leurs coefficients s'annulent pour $n<0$. Cependant, il était important dans la preuve de travailler dans un contexte moins rigide, entre autres pour introduire des partitions de l'unité.

Démonstration du théorème 5.1.1. Soit

$$F_t(z) = \left(\frac{I-R(z,t)}{1-\left(1-\frac{1}{\mu}\Xi(t)\right)z}\right)^{-1} - \left(\frac{I-R(z)}{1-z}\right)^{-1}.$$

D'après le théorème 2.2.20, il existe $A(z) = O\left(\frac{1}{n^{\beta-1}}\right)$ tel que

$$\left(\frac{I-R(z)}{1-z}\right)^{-1} = \frac{1}{\mu}P + (1-z)A(z).$$

Alors

$$\begin{aligned}\sum T_{n,t}z^n &= (I-R(z,t))^{-1}\\ &= \frac{1}{1-(1-\frac{1}{\mu}\Xi(t))z}\frac{1}{\mu}P + \frac{1-z}{1-(1-\frac{1}{\mu}\Xi(t))z}A(z) + \frac{1}{1-(1-\frac{1}{\mu}\Xi(t))z}F_t(z)\\ &= I+II+III.\end{aligned}$$

Le coefficient de z^n dans I est $\frac{1}{\mu}\left(1-\frac{1}{\mu}\Xi(t)\right)^n P$. On doit donc estimer les coefficients de II et III pour conclure.

Par (5.5) et le lemme 5.1.3, les coefficients de II sont de la forme

$$\left[1+O\left(t^2(1-dt^2)^n\right)\right]\star O\left(\frac{1}{n^{\beta-1}}\right)=O\left(\frac{1}{n^{\beta-1}}+t^2\left(1-\frac{d}{2}t^2\right)^n\right).$$

De plus, le corollaire 5.1.7 montre que le coefficient de z^n dans III est de la forme

$$O\left((1-dt^2)^n\right)\star O\left(\frac{t}{n^{\beta-1}}+t^3\left(1-\frac{d}{512}t^2\right)^n\right).$$

La convolution entre $(1-dt^2)^n$ et $t^3(1-\frac{d}{512}t^2)^n$ est bornée par la convolution entre $(1-\frac{d}{512}t^2)^n$ et $t^3(1-\frac{d}{512}t^2)^n$, qui donne $nt^3(1-\frac{d}{512}t^2)^n$. Ce terme est majoré par $Ct(1-\frac{d}{1024}t^2)^n$ puisque, pour tout $c>0$,

$$nt^2(1-ct^2)^n\leqslant 2t^2\sum_{k=0}^{n/2}(1-ct^2)^k(1-ct^2)^{n/2}\leqslant\frac{2}{c}(1-ct^2)^{n/2}.$$

Comme $t(1-\frac{d}{1024}t^2)^n\leqslant t(1-\frac{d}{1024}t^2)^n\star\left(\frac{1}{n^{\beta-1}}\right)$, on obtient bien une borne de la forme annoncée dans l'énoncé du théorème. □

5.2 L'estimée clé

Dans cette partie, on suppose que $T : X\to X$ est une tour localement Gibbs-Markov mélangeante de base Y, qui préserve une mesure de probabilité m. On supposera également que le temps de retour φ_Y satisfait $m[\varphi_Y>n]=O(1/n^\beta)$ pour un certain $\beta>2$.

Soit $f\in\mathcal{L}_\tau(X)$ avec $\int f\,\mathrm{d}m=0$, on va estimer très précisément la fonction caractéristique de S_nf dans le théorème 5.2.1. Notons que la fonction $f_Y(y)=\sum_{k=0}^{\varphi_Y(y)-1}f(T^ky)$ satisfait

$$\begin{aligned}\int_Y|f_Y|^2&\leqslant C\sum n^2m[\varphi_Y=n]=C\sum n^2\big(m(\varphi_Y>n-1)-m(\varphi_Y>n)\big)\\&\leqslant C\sum nm(\varphi_Y>n)<\infty,\end{aligned}$$

i.e. $f_Y\in L^2(Y)$, et

$$\sum_{a\in\alpha_0}m[a]Df_Y(a)\leqslant C\sum_{a\in\alpha_0}m[a]\varphi_Y(a),$$

qui est fini par la formule de Kac. Ainsi, f_Y satisfait les hypothèses du théorème 4.1.4. La valeur propre $\lambda(t)$ de $\widehat{T}_{Y,t}=\widehat{T}_Y(e^{itf_Y}\cdot)$ peut donc s'écrire

$$\lambda(t)=1-\frac{\sigma^2}{2m(Y)}\Lambda(t)\tag{5.6}$$

où $\Lambda(t) = t^2 + o(t^2)$. De plus, $\frac{1}{\sqrt{n}} S_n f \to \mathcal{N}(0, \sigma^2)$ d'après le théorème 4.4.1.
Le but de cette partie est de démontrer le résultat suivant :

Théorème 5.2.1. *Sous les hypothèses ci-dessus, supposons que $\sigma^2 > 0$. Alors il existe $C > 0, \alpha > 0$ et $d > 0$ tels que, pour toutes fonctions $u \in \mathcal{L}_\tau(X)$ et $v \in L^\infty(X)$, pour tout $n \in \mathbb{N}^*$, pour tout $t \in [-\alpha, \alpha]$,*

$$\left| \int_X e^{itS_n f} \cdot u \cdot v \circ T^n \, \mathrm{d}m - \left(1 - \frac{\sigma^2}{2}\Lambda(t)\right)^n \int_X u \, \mathrm{d}m \int_X v \, \mathrm{d}m \right| \leqslant C \left[\frac{1}{n^{\beta-1}} + |t| \left(\frac{1}{n^{\beta-1}} \right) \star (1 - dt^2)^n \right] \|u\|_{\mathcal{L}_\tau} \|v\|_\infty .$$

En prenant $t = 0$, on obtient que les corrélations de u et $v \circ T^n$ décroissent en $O(1/n^{\beta-1})$, i.e. on retrouve l'estimée (optimale) du théorème 2.3.6 dans le cas $u_n = 1/n^\beta$.
Dans la suite de cette partie, les hypothèses seront toujours celles du théorème 5.2.1.

5.2.1 Perturbation des opérateurs de transfert

Dans cette partie, on va utiliser tous les opérateurs de transfert définis dans le paragraphe 2.1.1, à savoir R_n, T_n, A_n, B_n et C_n. On va ensuite les perturber, en posant, pour $U = R, T, A, B, C$, $U_n(t) = U_n(e^{itS_n f} \cdot)$. On a déjà utilisé les perturbations $R_n(t)$ et $T_{n,t}$ de R_n et T_n dans le paragraphe 4.4.2, et on a vu qu'elles vérifiaient l'équation du renouvellement

$$T_{n,t} = \sum_{k_1 + \ldots + k_l = n} R_{k_1}(t) \cdots R_{k_l}(t). \tag{5.7}$$

De la même façon, l'équation (2.4) subsiste après perturbation, i.e.

$$\widehat{T}_t^n = C_n(t) + \sum_{a+k+b=n} A_a(t) T_{k,t} B_b(t). \tag{5.8}$$

Comme $f_Y = S_n f$ sur $\{y \in Y \mid \varphi_Y(y) = n\}$, on a $R_n(t)(u) = R_n(e^{itf_Y} u)$. Soient $R(z) = \sum R_n z^n$ et

$$R(z,t) = \sum_{n=0}^{\infty} R_n(t) z^n. \tag{5.9}$$

Cela correspond à considérer toutes les préimages d'un point de Y par T_Y, donc $R(1)$ est l'opérateur de transfert $\widehat{T}_Y$ associé à T_Y et

$$R(1,t)(u) = \widehat{T}_Y(e^{itf_Y} u). \tag{5.10}$$

On va récapituler dans le lemme suivant toutes les propriétés des opérateurs $R_n(t)$ qu'on a déjà plus ou moins établies dans les chapitres précédents et qui nous serviront dans la suite. Pour des raisons techniques, on ne va pas faire agir ces opérateurs sur $\mathcal{L}_\tau$, mais plutôt sur $\mathcal{L}_{\tau^\eta}$ avec $\eta = 1/2$.

Lemme 5.2.2. *Les opérateurs R_n et $R_n(t)$ agissant sur $\mathcal{L}_{\tau^\eta}$ avec $\eta = 1/2$ satisfont les propriétés suivantes :*

1. $\sum_{k=n+1}^{\infty} \|R_k\| = O(1/n^\beta)$.
2. *L'opérateur $R(z,t)$ satisfait une inégalité de Doeblin-Fortet*

$$\|R(z,t)^n u\| \leqslant C(1+|t|)(\tau^\eta|z|)^n \|u\| + C|z|^n \|u\|_1 .$$

   Ainsi, son rayon spectral est $\leqslant |z|$ et son rayon spectral essentiel est $\leqslant \tau^\eta|z|$. En particulier, si $I - R(z,t)$ n'est pas inversible, alors 1 est une valeur propre de $R(z,t)$, et cela ne peut se produire que pour $|z| = 1$.
3. *$R(1)$ a une valeur propre simple en 1, les fonctions propres correspondantes étant les fonctions constantes, et $I - R(z)$ est inversible pour $z \neq 1$.*
4. *Il existe $C > 0$ tel que, pour tout t assez petit, pour tout $n \in \mathbb{N}^*$,*

$$\sum_{k=n+1}^{\infty} \|R_k(t) - R_k(0)\| \leqslant C\frac{|t|}{n^{\beta-1}}.$$

5. *Pour tout réel t, il existe une constante C qui dépend de t telle que, pour tout réel t', pour tout $n \in \mathbb{N}^*$, $\|R_n(t') - R_n(t)\| \leqslant C\frac{|t'-t|^\eta}{n^{\beta-1}}$.*

Démonstration. La première assertion découle du lemme 2.1.2 qui donne $\|R_n\| = O(m[\varphi_Y = n])$.

La deuxième assertion découle des preuves du lemme 2.1.4 et du corollaire 4.1.3.

La troisième assertion est une reformulation des lemmes 2.1.3 et 2.1.4.

Pour la quatrième assertion, on applique le lemme 4.1.2 dans l'espace $\mathcal{L}_{\tau^\eta}$, i.e., les quantités τ, η, s et t du lemme 4.1.2 sont prises égales respectivement à τ^η, 1, t et 0 dans le contexte actuel. On somme les estimées données par ce lemme sur les éléments $a \in \alpha_0$ avec $\varphi_Y(a) = k$, et on obtient

$$\|R_k(t) - R_k\| \leqslant C\left(\sum_{\varphi_Y(a)=k} m(a)|t| D_{\tau^\eta} f_Y(a)\right) + C\int_{\{\varphi_Y=k\}} |e^{itf_Y} - 1|.$$

Mais $D_{\tau^\eta} f_Y(a) \leqslant C\varphi_Y(a) = Ck$ et $|e^{itf_Y} - 1| \leqslant |t||f_Y| \leqslant C|t|\varphi_Y = C|t|k$. Ainsi,

$$\|R_k(t) - R_k\| \leqslant C|t|k m(\varphi_Y = k).$$

Finalement,

$$\sum_{k=n+1}^{\infty} \|R_k(t) - R_k\| \leqslant C|t| \sum_{k=n+1}^{\infty} km(\varphi_Y = k)$$
$$= C|t| \left((n+1)m(\varphi_Y > n) + \sum_{k=n+1}^{\infty} m(\varphi_Y > k) \right),$$

ce qui permet de conclure puisque $m(\varphi_Y > n) = O(1/n^\beta)$.

Finalement, pour la cinquième assertion, on somme encore les inégalités données par le lemme 4.1.2, mais cette fois sur l'espace $\mathcal{L}_\tau$, et on majore $1 + |t|^\eta D_\tau f(a)^\eta$ par $C\varphi_Y(a)^\eta$. On obtient

$$\|R_n(t+s) - R_n(t)\| \leqslant C \sum_{\varphi_Y(a)=n} \left(m(a)|s|^\eta D_\tau f_Y(a)^\eta \varphi_Y(a)^\eta + \int_a |e^{isf} - 1| \varphi_Y(a)^\eta \right).$$

On majore $|e^{isf} - 1|$ par $C|sf|^\eta \leqslant C|s|^\eta \varphi_Y^\eta$, et on obtient au final une majoration en $C|s|^\eta (n^\eta)^2 m(\varphi_Y = n) = C|s|^\eta n m(\varphi_Y = n)$. Cela conclut la preuve car $m(\varphi_Y = n) \leqslant m(\varphi_Y > n-1) = O(1/n^\beta)$. □

On aura aussi besoin des estimées suivantes sur $A_a(t), B_b(t)$ et $C_n(t)$; elles étendent au cas perturbatif les lemmes 2.3.3, 2.3.4 et 2.3.5.

Lemme 5.2.3. *Soient $u, v \in L^\infty(X)$. Alors $\int A_a(t)(u) \cdot v = \int A_a(u) \cdot v + O\left(\frac{|t|}{a^{\beta-1}}\right)$. De plus, $\int A_a(u) \cdot v = O(1/a^\beta)$ et*

$$\sum_{a \in \mathbb{N}} \int A_a(1_Y) \cdot v = \int v. \tag{5.11}$$

Ici, le $O\left(\frac{|t|}{a^{\beta-1}}\right)$ est uniforme sur $\mathbb{R}$: il existe $C > 0$ tel que $\forall t \in \mathbb{R}$, $\forall a \in \mathbb{N}^*$,

$$\left| \int A_a(t)(u) \cdot v - \int A_a(u) \cdot v \right| \leqslant \frac{C|t|}{a^{\beta-1}} \|u\|_\infty \|v\|_\infty .$$

Démonstration. Un changement de variables donne $\int A_a(t)(u) \cdot v = \int_{\{\varphi_Y > a\}} e^{itS_a f} uv'$, où v' est défini par $v'(y,0) = v(y,a)$ si $\varphi_Y(y) > a$ et $v'(y,0) = 0$ sinon. Ainsi,

$$\left| \int A_a(t)(u) \cdot v - \int A_a(u) \cdot v \right| \leqslant \int_{\{\varphi_Y > a\}} |t||S_a f| \|u\|_\infty \|v\|_\infty$$
$$\leqslant |t| m[\varphi_Y > a] a \|f\|_\infty \|u\|_\infty \|v\|_\infty .$$

Comme $m[\varphi_Y > a] = O(1/a^\beta)$, cela conclut la preuve du premier point. De plus, comme $\sum A_a(1_Y) = 1_X$, (5.11) est immédiat. □

Lemme 5.2.4. *Pour $t \in \mathbb{R}$, $B_b(t) = B_b + O\left(\frac{|t|}{b^{\beta-1}}\right)$ où $\|B_b\| = O(1/b^\beta)$. De plus, $\forall u \in \mathcal{L}(X)$, on a*

$$\sum_{b=0}^{\infty} \int_Y B_b u \, \mathrm{d}m = \int_X u \, \mathrm{d}m.$$

Démonstration. Soit Λ_b l'ensemble des points qui arrivent dans la base après exactement b itérations, si bien que $B_b(t)(u)$ prend en compte les valeurs de u sur Λ_b. On vérifie en utilisant le lemme 4.1.2 que $\|B_b(t) - B_b\| \leqslant C|t| m(\Lambda_b) DS_b f(\Lambda_b)$. Comme $DS_b f \leqslant Cb$ et $m(\Lambda_b) = m[\varphi_Y > b] = O(1/b^\beta)$, on obtient que $\|B_b(t) - B_b\| \leqslant C\frac{|t|}{b^{\beta-1}}$.
De plus, $\|B_b\| = O(m(\Lambda_b)) = O(1/b^\beta)$ d'après le lemme 2.3.4. Comme $\int_Y B_b u \, \mathrm{d}m = \int_{\Lambda_b} u \, \mathrm{d}m$, on obtient finalement $\sum_b \int_Y B_b u \, \mathrm{d}m = \int_X u \, \mathrm{d}m$. □

Lemme 5.2.5. *Il existe une constante $C > 0$ telle que, pour tout $n \in \mathbb{N}^*$ et pour tout $t \in \mathbb{R}$, $\|C_n(t)(u)\|_1 \leqslant \frac{C}{n^{\beta-1}} \|u\|_\infty$.*

Démonstration. Pour tout $x \in X$, $|C_n(t)(u)(x)| \leqslant |C_n(u)(x)|$. Ainsi, le lemme est une conséquence directe du lemme 2.3.5. □

5.2.2 Démonstration du théorème 5.2.1

On veut appliquer le théorème 5.1.1 à $R(z,t) = \sum R_n(t) z^n$. On sait déjà que la valeur propre proche de 1 de $R(1,t)$ se met sous la forme $1 - \frac{\sigma^2}{2m(Y)}\Lambda(t)$ avec $\Lambda(t) \sim t^2$, d'après le théorème 4.1.4. De plus, le lemme 5.2.2 implique les estimées requises sur les normes de R_n et $R_n(t)$. Finalement, le projecteur spectral P de $R(1,0)$ associé à sa valeur propre 1 est la projection sur les fonctions constantes sur Y, donnée par $Pu = \frac{\int_Y u}{m(Y)}$. Il satisfait $PR'(1)P = \frac{1}{m(Y)}P$ d'après le lemme 2.1.5.

Ainsi, on peut appliquer le théorème 5.1.1 avec $\Xi(t) = \frac{\sigma^2}{2m(Y)}\Lambda(t)$ et $\mu = \frac{1}{m(Y)}$. Comme $\sum T_{n,t} z^n = (I - R(z,t))^{-1}$ par (5.7), on en déduit qu'il existe un terme d'erreur $E(n,t)$ tel que $T_{n,t} = m(Y)\left(1 - \frac{\sigma^2}{2}\Lambda(t)\right)^n P + E(n,t)$, avec

$$\|E(n,t)\| \leqslant C\left[\frac{1}{n^{\beta-1}} + |t|\left(\frac{1}{n^{\beta-1}}\right) \star \left(1 - dt^2\right)^n\right] =: e(n,t). \tag{5.12}$$

Dans la suite, on notera C et d des constantes génériques qui pourront varier un nombre fini de fois. En particulier, on pourra écrire des inégalités comme $10e(n,t) \leqslant e(n,t)$. Avec cette convention, le lemme 5.1.3 implique que

$$\left(\frac{1}{n^{\beta-1}}\right) \star e(n,t) = O(e(n,t)). \tag{5.13}$$

Soient $u \in \mathcal{L}(X)$ et $v \in L^\infty(X)$. Pour simplifier les expressions, on supposera que ces deux fonctions sont de norme au plus 1. Alors (5.8) implique que

$$\begin{aligned}\int_X e^{itS_nf} \cdot u \cdot v \circ T^n &= \int_X \widehat{T}^n(e^{itS_nf}u)v = \int_X \widehat{T}_t^n(u)v \\ &= \sum_{a+k+b=n} \int_X A_a(t)T_{k,t}B_b(t)(u)v + \int_X C_n(t)(u)v \\ &= \sum_{a+k+b=n} \int_X A_a(t)m(Y)\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k PB_b(t)(u)v \\ &\quad + \sum_{a+k+b=n} \int_X A_a(t)E(k,t)B_b(t)(u)v + \int_X C_n(t)(u)v.\end{aligned} \tag{5.14}$$

D'après le lemme 5.2.5, $\left|\int_X C_n(t)(u)v\right| \leqslant \frac{C}{n^{\beta-1}} \|u\|_\infty \|v\|_\infty$.

Bornons ensuite d'abord la deuxième somme. Si $h \in \mathcal{L}(X)$, la fonction $A_a(t)h$ est supportée par $T^a\big(\{y \in Y \mid \varphi_Y(y) > a\}\big)$, dont la mesure est $m[\varphi_Y > a] = O(1/a^\beta)$. Ainsi, $\left|\int_X A_a(t)(h)v\right| \leqslant \frac{C}{a^\beta} \|h\|_\infty \|v\|_\infty \leqslant \frac{C}{a^{\beta-1}} \|h\|$. De plus, $\|E(k,t)B_b(t)(u)\| \leqslant e(k,t)\,\|B_b(t)\| \leqslant e(k,t)\frac{C}{b^{\beta-1}}$ d'après le lemme 5.2.4. Ainsi,

$$\begin{aligned}\left|\sum_{a+k+b=n} \int_X A_a(t)E(k,t)B_b(t)(u)v\right| &\leqslant \sum_{a+k+b=n} \frac{C}{a^{\beta-1}} e(k,t) \frac{C}{b^{\beta-1}} \\ &\leqslant C\left(\frac{1}{n^{\beta-1}}\right) \star e(n,t) \star \left(\frac{1}{n^{\beta-1}}\right) \leqslant e(n,t)\end{aligned} \tag{5.15}$$

d'après (5.13).

On peut écrire la première somme de (5.14) comme

$$I = \sum_{a+k+b=n} \left(\int_X A_a(t)(1_Y)v\right)\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k \left(\int_Y B_b(t)(u)\right).$$

En utilisant les lemmes 5.2.3 et 5.2.4 et en convolant, on obtient une suite w_c telle que $w_c = O(1/c^\beta)$, $\sum w_c = \left(\int u\right)\left(\int v\right)$, et

$$I = \left(w_c + O\left(\frac{|t|}{c^{\beta-1}}\right)\right) \star \left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k.$$

Comme $\Lambda(t) \sim t^2$ quant $t \to 0$, le terme venant de $O\left(\frac{|t|}{c^{\beta-1}}\right) \star \left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k$ est borné par $|t|\left(\frac{1}{n^{\beta-1}}\right) \star (1-dt^2)^n \leqslant e(n,t)$. De plus, pour $x, y \in \mathbb{R}$,

$$|e^x - e^y| \leqslant |x-y|e^{\max(x,y)}. \tag{5.16}$$

Ainsi,

$$\begin{aligned}\left|\sum_{k=0}^{n} w_{n-k}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k - \sum_{k=0}^{n} w_{n-k}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n\right| \\ \leqslant \sum_{k=0}^{n/2}\frac{C}{(n-k)^\beta} + \sum_{k=n/2}^{n}\frac{C}{(n-k)^\beta}(n-k)\left|\ln\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)\right|\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^k \\ \leqslant \frac{C}{n^{\beta-1}} + \sum_{k=n/2}^{n}\frac{C}{(n-k)^{\beta-1}}t^2(1-dt^2)^{n/2} \leqslant e(n,t).\end{aligned}$$

Ainsi, à $O(e(n,t))$ près, l'intégrale $\int e^{itS_nf}\cdot u\cdot v\circ T^n$ est égale à $\sum_{k=0}^{n} w_{n-k}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n$. Finalement,

$$\begin{aligned}\left|\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n\left(\int_X u\right)\left(\int_X v\right) - \sum_{k=0}^{n} w_{n-k}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n\right| \\ = \left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n\left|\sum_{k=n+1}^{\infty} w_k\right| \leqslant C\sum_{k=n+1}^{\infty}\frac{1}{k^\beta} \leqslant \frac{C}{n^{\beta-1}} \leqslant e(n,t).\end{aligned}$$

Cela conclut la preuve. □

5.3 Le théorème de la limite locale

5.3.1 Problèmes de périodicité

Ce paragraphe est à rapprocher de la fin de la partie 3 de [AD01b].

Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante préservant une mesure de probabilité m, de base Y, et $f \in \mathcal{L}_\tau(X)$. Pour $t \in \mathbb{R}$, on dit que e^{itf} est *cohomologue* à une constante $\lambda \in S^1$ s'il existe $\omega : X \to S^1$ mesurable telle que $e^{itf} = \lambda\frac{\omega\circ T}{\omega}$ presque partout.

Proposition 5.3.1. *Soient $t \in \mathbb{R}$ et $z \in S^1$. Les assertions suivantes sont équivalentes :*

1. *e^{itf} est cohomologue à z^{-1}.*
2. *1 est valeur propre de l'opérateur $R(z,t)$ agissant sur $\mathcal{L}_\tau(Y)$.*

Démonstration. Supposons tout d'abord qu'il existe $\omega \in \mathcal{L}_\tau(Y)$ non nul tel que $R(z,t)\omega = \omega$. Comme $R(z,t)(v) = \widehat{T}_Y(z^{\varphi_Y}e^{itf_Y}v)$, l'opérateur adjoint dans L^2 de $R(z,t)$ est $W(v) = z^{-\varphi_Y}e^{-itf_Y}v\circ T_Y$. Alors $\|W\omega-\omega\|_2^2 = \|W\omega\|_2^2 - \|\omega\|_2^2$. Comme T_Y préserve la mesure, on a $\|\omega\|_2 = \|W\omega\|_2$, donc $\|W\omega-\omega\|_2 = 0$. Par conséquent, $\omega = z^{-\varphi_Y}e^{-itf_Y}\omega\circ T_Y$ presque

partout. En prenant le module, on obtient $|\omega| = |\omega| \circ T_Y$, et l'ergodicité de T_Y implique que $|\omega|$ est presque partout constant. Comme ω n'est pas la fonction nulle, on peut supposer que $|\omega| = 1$. On étend ω à X en posant

$$\omega(x,k) = \omega(x,0)e^{itf(x,0)} \cdots e^{itf(x,k-1)} z^k.$$

Alors, pour $k < \varphi_Y(x) - 1$, on a par construction $\omega \circ T(x,k)/\omega(x,k) = ze^{itf(x,k)}$. De plus, pour $k = \varphi_Y(x) - 1$,

$$\omega \circ T(x,k)/\omega(x,k) = \omega \circ T_Y(x,0)/\omega(x,k) = e^{itf_Y(x,0)} z^{k+1} \omega(x,0)/\omega(x,k) = e^{itf(x,k)} z.$$

Ainsi, $e^{itf} = z^{-1}\omega \circ T/\omega$ presque partout.

Réciproquement, supposons qu'une fonction mesurable ω sur X satisfasse $e^{itf} = z^{-1}\omega \circ T/\omega$. Le calcul précédent montre qu'en ce cas $e^{itf_Y} = z^{-\varphi_Y}\omega \circ T_Y/\omega$. L'opérateur $R(z,t) = \widehat{T}_Y(z^{\varphi_Y} e^{itf_Y} \cdot)$ agit sur $L^1(Y)$ et

$$R(z,t)(\omega) = \widehat{T}_Y\left(z^{\varphi_Y} e^{itf_Y}\omega\right) = \widehat{T}_Y\left(\omega \circ T_Y\right) = \omega.$$

Mais, dans le point 2 du lemme 5.2.2, on a vu que $R(z,t)$ satisfait une inégalité de Doeblin-Fortet entre les espaces $\mathcal{L}_\tau(Y)$ et $L^1(Y)$. D'après la proposition 1.1.12, cela implique que les fonctions propres de $R(z,t)$ dans $L^1(Y)$ sont en fait dans $\mathcal{L}_\tau(Y)$. Ainsi, $\omega \in \mathcal{L}_\tau(Y)$. □

Corollaire 5.3.2. *L'ensemble $\mathfrak{A} := \{t \in \mathbb{R} \mid e^{itf}$ est cohomologue à une constante$\}$ est un sous-groupe fermé de $\mathbb{R}$. De plus, pour tout $t \in \mathfrak{A}$, il existe un unique $z(t) \in S^1$ tel que e^{itf} soit cohomologue à $z(t)$. Finalement, l'application $t \mapsto z(t)$ est un morphisme continu de $\mathfrak{A}$ dans S^1.*

Démonstration. L'ensemble $\mathfrak{A}$ est clairement un sous-groupe de $\mathbb{R}$. Si e^{itf} est cohomologue simultanément à z et z', alors $z' = \frac{e^{ih\circ T}}{e^{ih}} z$ pour une certaine fonction $h : X \to \mathbb{R}$. Comme T est mélangeante d'après le théorème 1.3.15, toute constante s satisfaisant $g \circ T = sg$ pour une fonction non nulle g est nécessairement égale à 1. Cela implique que $z = z'$. L'application $t \mapsto z(t)$ est donc bien définie, et c'est un morphisme de groupe.

Il reste à vérifier que $\mathfrak{A}$ est fermé et que $z(t)$ est continue. Soit t_n une suite de $\mathfrak{A}$ qui converge vers $T \in \mathbb{R}$. Soit Z une valeur d'adhérence de la suite $z(t_n)$. Par le point 5 du lemme 5.2.2, $(z,t) \mapsto R(z,t)$ est une application continue à valeurs dans $\mathrm{Hom}(\mathcal{L}_{\tau^{1/2}}(Y), \mathcal{L}_{\tau^{1/2}}(Y))$. Si $I - R(Z^{-1}, T)$ était inversible, alors $I - R(z_n^{-1}, t_n)$ serait également inversible pour n assez grand, ce qui est absurde d'après la proposition précédente (appliquée dans $\mathcal{L}_{\tau^{1/2}}$). Ainsi, $I - R(Z^{-1}, T)$ n'est pas inversible, donc 1 est une valeur propre de $R(Z^{-1}, T)$ par quasi-compacité. Cela implique que $T \in \mathfrak{A}$ et que $Z = z(T)$, encore une fois par la proposition précédente. □

Par conséquent, il y a trois possibilités : $\mathfrak{A}$ est soit $\mathbb{R}$, soit $\{0\}$ (auquel cas on dit que f est apériodique), soit un sous-groupe discret de $\mathbb{R}$. Si c'est $\mathbb{R}$, [MS80] implique que f peut être

écrite sous la forme $\chi - \chi \circ T$, donc $\sigma^2 = 0$ dans le théorème central limite et il n'y a rien à démontrer. Si $\mathfrak{A} = \{0\}$, alors f ne peut pas s'écrire sous la forme $\rho + g - g \circ T + \lambda q$ où $\rho \in \mathbb{R}$, $g : X \to \mathbb{R}$ est mesurable, $\lambda > 0$ et $q : X \to \mathbb{Z}$. On verra dans le théorème 5.3.3 comment obtenir un théorème de la limite locale dans ce cas. Finalement, si $\mathfrak{A}$ est un sous-groupe discret non trivial de $\mathbb{R}$, par exemple $\mathfrak{A} = 2\pi\mathbb{Z}$, alors on peut écrire $f = \rho + g - g \circ T + q$ où q prend des valeurs entières, mais il n'existe pas de telle décomposition de f où q prendrait ses valeurs dans $n\mathbb{Z}$ avec $n \geqslant 2$. Ce cas est traité dans le théorème 5.3.4.

5.3.2 Le cas apériodique

Lorsque la fonction f est apériodique, on peut démontrer le théorème de la limite locale suivant :

Théorème 5.3.3 (Théorème de la limite locale, cas apériodique)**.** *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante de base Y, qui préserve une mesure de probabilité m. On suppose que $m[\varphi_Y > n] = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f \in \mathcal{L}_\tau(X)$ d'intégrale nulle, et soit σ^2 la variance dans le théorème central limite pour f, donnée par le théorème 4.4.1.*

On suppose que f est apériodique, c'est-à-dire qu'elle ne s'écrit pas sous la forme $\rho + g - g \circ T + \lambda q$ où $\rho \in \mathbb{R}$, $g : X \to \mathbb{R}$ est mesurable, $\lambda > 0$ et $q : X \to \mathbb{Z}$. Cela implique en particulier que $\sigma^2 > 0$.

Alors, pour tout intervalle borné $J \subset \mathbb{R}$, pour toute suite k_n avec $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$, pour toutes fonctions $u \in \mathcal{L}_\tau(X)$ et $v : X \to \mathbb{R}$ mesurable,

$$\sqrt{n}\, m\{x \in X \mid S_n f(x) \in J + k_n + u(x) + v(T^n x)\} \to |J| \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

La fonction à droite est la densité de $\mathcal{N}(0, \sigma^2)$: lorsque $u = v = 0$ et $k_n = \kappa\sqrt{n}$, ce théorème signifie que

$$m\left\{\frac{1}{\sqrt{n}} S_n f \in \kappa + \frac{J}{\sqrt{n}}\right\} \sim P\left(\mathcal{N}(0, \sigma^2) \in \kappa + \frac{J}{\sqrt{n}}\right).$$

Autrement dit, $\frac{1}{\sqrt{n}} S_n f$ se comporte non seulement globalement comme $\mathcal{N}(0, \sigma^2)$ comme l'affirme le théorème central limite, mais également localement.

L'hypothèse d'apériodicité est importante. Autrement, f pourrait par exemple ne prendre que des valeurs entières, et le théorème ne pourrait pas être valide par exemple pour $k_n = 0$, $u = v = 0$ et $J = [1/3, 2/3]$.

Démonstration. Pour tout $(z, t) \in (\overline{\mathbb{D}} \times \mathbb{R}) \backslash \{(1, 0)\}$, $I - R(z, t)$ est inversible : pour $|z| < 1$, le rayon spectral de $R(z, t)$ est au plus $|z| < 1$ d'après le point 2 du lemme 5.2.2, et pour $|z| = 1$ cela vient de la proposition 5.3.1 et du corollaire 5.3.2 puisque $\mathfrak{A} = \{0\}$.

Soit $\alpha > 0$ donné par le théorème 5.2.1. Ce théorème permet de contrôler le comportement des fonctions caractéristiques de $S_n f$ pour $|t| \leqslant \alpha$. Soit $K > 0$. On va montrer l'affirmation suivante : *il existe* $C > 0$ *tel que, pour tout* $|t| \in [\alpha, K]$, *pour tout* $n \in \mathbb{N}^*$, *pour toutes fonctions* $u \in \mathcal{L}_\tau(X)$ *et* $v \in L^\infty(X)$,

$$\left| \int_X e^{itS_n f} \cdot u \cdot v \circ T^n \right| \leqslant \frac{C}{n^{\beta-1}} \|u\|_{\mathcal{L}} \|v\|_\infty . \tag{5.17}$$

Soit $\mathcal{A} = \mathcal{O}^+_{\beta-1}(\mathrm{Hom}(\mathcal{L}_{\tau^{1/2}}(Y), \mathcal{L}_{\tau^{1/2}}(Y)))$ l'algèbre de Banach des séries $\sum_0^\infty A_n z^n$ avec $\|A_n\| = O(1/n^{\beta-1})$, muni de la norme $\|\sum A_n z^n\| = \sup(n+1)^{\beta-1} \|A_n\|$. L'application $t \mapsto R(z,t)$ est continue de $[-K, -\alpha] \cup [\alpha, K]$ dans $\mathcal{A}$ d'après le point 5 du lemme 5.2.2. De plus, $I - R(z,t)$ est inversible sur $\overline{\mathbb{D}}$ pour t dans ces intervalles. Le lemme de Wiener montre donc que $(I - R(z,t))^{-1} \in \mathcal{A}$, et la continuité de l'inversion implique que $t \mapsto (I - R(z,t))^{-1}$ est continue. Par compacité, il existe C tel que $\|(I - R(z,t))^{-1}\|_{\mathcal{A}} \leqslant C$ pour $|t| \in [\alpha, K]$. Comme $(I - R(z,t))^{-1} = \sum T_{n,t} z^n$, cela implique que $\|T_{n,t}\| \leqslant \frac{C}{n^{\beta-1}}$, uniformément en n et t.

On a

$$\int_X e^{itS_n f} \cdot u \cdot v \circ T^n = \int_X C_n(u)v + \sum_{a+k+b=n} \int_X A_a(t) T_{k,t} B_b(t)(u) \cdot v.$$

D'après le lemme 5.2.5, $\left|\int_X C_n(u)v\right| \leqslant \frac{C}{n^{\beta-1}} \|u\|_\infty \|v\|_\infty$. D'après les lemmes 5.2.3 et 5.2.4,

$$\begin{aligned}\left| \int A_a(t) T_{k,t} B_b(t)(u) v \right| &\leqslant \frac{C}{a^{\beta-1}} \|T_{k,t} B_b(t)(u)\|_\infty \|v\|_\infty \leqslant \frac{C}{a^{\beta-1}} \|T_{k,t}\| \, \|B_b(t)\| \, \|u\| \, \|v\|_\infty \\ &\leqslant \frac{C}{a^{\beta-1}} \frac{C}{k^{\beta-1}} \frac{C}{b^{\beta-1}} \|u\| \, \|v\|_\infty .\end{aligned}$$

Ainsi,

$$\left| \int_X e^{itS_n f} \cdot u \cdot v \circ T^n \right| \leqslant C \left(\frac{1}{n^{\beta-1}}\right) \star \left(\frac{1}{n^{\beta-1}}\right) \star \left(\frac{1}{n^{\beta-1}}\right) \|u\| \, \|v\|_\infty \leqslant \frac{C}{n^{\beta-1}} \|u\| \, \|v\|_\infty ,$$

et l'équation (5.17) est démontrée.

On prouve maintenant le théorème de la limite locale en utilisant la méthode de Breiman ([Bre68]). Soient u, v et k_n comme dans les hypothèses du théorème. Soit $\psi \in L^1(\mathbb{R})$ dont la transformée de Fourier $\widehat{\psi}$ est continue et à support inclus dans $[-K, K]$. Alors $\psi(x) = \frac{1}{2\pi} \int_{-K}^K \widehat{\psi}(t) e^{itx} \, \mathrm{d}t$ donc

$$\begin{aligned}\sqrt{n} E(\psi(S_n f - k_n - u - v \circ T^n)) &= \frac{\sqrt{n}}{2\pi} \int_{-K}^K \widehat{\psi}(t) E\left(e^{it(S_n f - k_n - u - v\circ T^n)}\right) \, \mathrm{d}t \\ &= \frac{\sqrt{n}}{2\pi} \int_{-\alpha}^{\alpha} \widehat{\psi}(t) e^{-itk_n} E(e^{itS_n f} e^{-itu} e^{-itv\circ T^n}) \, \mathrm{d}t \qquad (5.18) \\ &\quad + \frac{\sqrt{n}}{2\pi} \int_{\alpha \leqslant |t| \leqslant K} \widehat{\psi}(t) e^{-itk_n} E(e^{itS_n f} e^{-itu} e^{-itv\circ T^n}) \, \mathrm{d}t.\end{aligned}$$

Pour $\alpha \leqslant |t| \leqslant K$, les normes $\|e^{-itu}\|$ et $\|e^{-itv}\|_\infty$ restent bornées donc l'équation (5.17) montre que $\left|E(e^{itS_nf}e^{-itu}e^{-itv\circ T^n})\right| \leqslant \frac{C}{n^{\beta-1}}$. Par conséquent, la deuxième intégrale dans (5.18) tend vers 0. Pour la première intégrale, le théorème 5.2.1 montre qu'on peut approcher $E(e^{itS_nf}e^{-itu}e^{-itv\circ T^n})$ par $\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n \int e^{-itu}\int e^{-itv}$. L'erreur est bornée par

$$C\sqrt{n}\int_{-\alpha}^{\alpha}\left(\frac{1}{n^{\beta-1}}+|t|\left(\frac{1}{n^{\beta-1}}\right)\star(1-dt^2)^n\right)\mathrm{d}t.$$

Montrons que cette intégrale tend vers 0. C'est clair pour le premier terme. Pour le deuxième, on coupe l'intégrale en deux morceaux. Pour $|t| \leqslant \frac{1}{\sqrt{n}}$, la convolution est bornée puisque $\frac{1}{n^{\beta-1}}$ est sommable et $(1-dt^2)^n \leqslant 1$, donc l'intégrale est $\leqslant C\sqrt{n}\int_{|t|\leqslant 1/\sqrt{n}}|t| \to 0$. Pour $|t| \geqslant \frac{1}{\sqrt{n}}$, le lemme 5.1.3 montre que la convolution est majorée par $\frac{1}{|t|n^{\beta-1}}+|t|(1-dt^2)^n$, donc l'intégrale vaut au plus

$$C\sqrt{n}\int_{1/\sqrt{n}\leqslant|t|\leqslant\alpha}\frac{1}{|t|n^{\beta-1}}+|t|(1-dt^2)^n \leqslant C\sqrt{n}\frac{\ln n}{n^{\beta-1}}+C\sqrt{n}\left[\frac{(1-dt^2)^{n+1}}{-2d(n+1)}\right]_{1/\sqrt{n}}^{\alpha}\to 0.$$

Finalement, on a montré que

$$\sqrt{n}E(\psi(S_nf-k_n-u-v\circ T^n)) = \frac{\sqrt{n}}{2\pi}\int_{-\alpha}^{\alpha}\widehat{\psi}(t)e^{-itk_n}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n E(e^{-itu})E(e^{-itv})\,\mathrm{d}t+o(1).$$

Mais

$$\begin{aligned}\frac{\sqrt{n}}{2\pi}\int_{-\alpha}^{\alpha}\widehat{\psi}(t)e^{-itk_n}\left(1-\frac{\sigma^2}{2}\Lambda(t)\right)^n E(e^{-itu})E(e^{-itv})\,\mathrm{d}t&\\ = \frac{1}{2\pi}\int_{-\alpha\sqrt{n}}^{\alpha\sqrt{n}}\widehat{\psi}\left(\frac{t}{\sqrt{n}}\right)e^{-itk_n/\sqrt{n}}\left(1-\frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n E(e^{-i\frac{t}{\sqrt{n}}u})E(e^{-i\frac{t}{\sqrt{n}}v})\,\mathrm{d}t&\\ \to \frac{1}{2\pi}\int_{\mathbb{R}}\widehat{\psi}(0)e^{-it\kappa}e^{-\frac{\sigma^2}{2}t^2}\,\mathrm{d}t,&\end{aligned}$$

par convergence dominée. On a utilisé le fait que $\Lambda(t)\sim t^2$ près de 0, si bien qu'en particulier $\left(1-\frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n \leqslant \left(1-\frac{\sigma^2t^2}{4n}\right)^n \leqslant e^{-\frac{\sigma^2t^2}{4}}$ si α est assez petit, ce qui donne la domination.

Soit $\chi(\kappa) = \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} = \frac{1}{2\pi}\int_{\mathbb{R}}e^{-it\kappa}e^{-\frac{\sigma^2}{2}t^2}$. On a prouvé que, pour toute fonction $\psi \in L^1$ dont la transformée de Fourier est continue à support compact,

$$\sqrt{n}E(\psi(S_nf-k_n-u-v\circ T^n)) \to \chi(\kappa)\int_{\mathbb{R}}\psi(x)\,\mathrm{d}x. \tag{5.19}$$

L'équation (5.19) peut alors être étendue à une classe plus large de fonctions par des arguments de densité (voir [Bre68]), et cette classe plus large contient en particulier les fonctions caractéristiques d'intervalles bornés. Cela conclut la preuve. □

5.3.3 Le cas périodique

Le théorème suivant donne le théorème de la limite locale quand le groupe $\mathfrak{A}$ du paragraphe 5.3.1 est un sous-groupe discret de $\mathbb{R}$, par exemple $2\pi\mathbb{Z}$.

Théorème 5.3.4 (Théorème de la limite locale, cas périodique)**.** *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante de base Y, qui préserve une mesure de probabilité m. On suppose que $m[\varphi_Y > n] = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f \in \mathcal{L}_\tau(X)$ d'intégrale nulle, et soit σ^2 la variance dans le théorème central limite pour f donnée par le théorème 4.4.1.*

Supposons que $f = \rho + q$ où q est à valeurs entières et $\rho \in \mathbb{R}$, et que f ne peut pas être écrite sous la forme $f = \rho' + g - g \circ T + \lambda q'$ où $\lambda \in \mathbb{N}\backslash\{0, 1\}$ et $q' : X \to \mathbb{Z}$. Cela implique en particulier que $\sigma > 0$. Alors, pour toute suite k_n avec $k_n - n\rho \in \mathbb{Z}$ telle que $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$,

$$\sqrt{n}\, m\{x \in X \mid S_n f(x) = k_n\} \to \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

Démonstration. La preuve est essentiellement la même que celle du théorème 5.3.3, sauf qu'on va utiliser la transformée de Fourier sur $\mathbb{Z}$, c'est-à-dire les séries de Fourier, au lieu de la transformée de Fourier sur $\mathbb{R}$.

Si k et l sont deux entiers,

$$1_{k=l} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{it(l-k)}\, \mathrm{d}t.$$

En appliquant cette équation à $k_n - n\rho$ et $S_n f(x) - n\rho$ et en intégrant, on obtient

$$m\{x \in X \mid S_n f(x) = k_n\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk_n} E(e^{itS_n f})\, \mathrm{d}t.$$

C'est un analogue de (5.18). À partir de là, la preuve du théorème 5.3.3 s'applique. Le seul problème est de montrer que, sur $[-\pi, -\alpha] \cup [\alpha, \pi]$, $I - R(z, t)$ est inversible pour $z \in \overline{\mathbb{D}}$. Cela vient des hypothèses sur f qui assurent que $I - R(z, t)$ est inversible dès que $t \notin 2\pi\mathbb{Z}$ d'après la proposition 5.3.1 et le corollaire 5.3.2, puisque $\mathfrak{A} = 2\pi\mathbb{Z}$. □

5.4 Le théorème de Berry-Esseen

Théorème 5.4.1 (Vitesse dans le théorème central limite)**.** *Soit $T : X \to X$ une tour localement Gibbs-Markov mélangeante de base Y, qui préserve une mesure de probabilité m. On suppose que $m[\varphi_Y > n] = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f \in \mathcal{L}_\tau(X)$ d'intégrale nulle, et soit σ^2 la variance dans le théorème central limite pour f donnée par le théorème 4.4.1.*

Supposons que $\sigma^2 > 0$ *et qu'il existe* $0 < \delta \leqslant 1$ *tel que* $E(|f_Y|^2 1_{|f_Y|>z}) = O(z^{-\delta})$ *quand* $z \to +\infty$. *Si* $\delta = 1$, *supposons de plus que* $E(f_Y^3 1_{|f_Y|\leqslant z}) = O(1)$. *Alors il existe* $C > 0$ *tel que* $\forall n \in \mathbb{N}^*, \forall a \in \mathbb{R}$,

$$\left| m\left\{ x \mid \frac{1}{\sqrt{n}} S_n f(x) \leqslant a \right\} - P\big(\mathcal{N}(0,\sigma^2) \leqslant a\big) \right| \leqslant \frac{C}{n^{\delta/2}}.$$

Lorsque $f_Y \in L^p$ pour un certain $2 < p \leqslant 3$, alors les conditions du théorème sont satisfaites pour $\delta = p - 2$. En particulier, quand $f_Y \in L^3$, on obtient une convergence en $O(1/\sqrt{n})$, ce qui est le théorème de Berry-Esseen habituel. On a donné un résultat un peu plus précis dont les hypothèses correspondent aux conditions nécessaires et suffisantes pour obtenir une vitesse en $O(n^{-\delta/2})$ dans le théorème central limite pour des variables aléatoires indépendantes identiquement distribuées, d'après [IL71, théorème 3.4.1]. En effet, dans les applications, une hypothèse $f_Y \in L^p$ ne serait pas toujours suffisante, par exemple dans le cas des applications LSV dans le paragraphe 5.5.4. Notons aussi que, pour toute fonction $f \in \mathcal{L}_\tau(X)$, les hypothèses du théorème sont satisfaites pour $\delta = \beta - 2$ si $2 < \beta < 3$ et pour $\delta = 1$ si $\beta > 3$.

Démonstration du théorème 5.4.1. L'inégalité de Berry-Esseen ([Fel66, lemme XVI.3.2]) implique que le résultat sera prouvé si l'on montre que, pour un certain $c > 0$,

$$\int_{-c\sqrt{n}}^{c\sqrt{n}} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - e^{-\frac{\sigma^2}{2}t^2} \right| = O\left(\frac{1}{n^{\delta/2}}\right).$$

On estime tout d'abord l'intégrale entre $-1/n$ et $1/n$:

$$\begin{aligned}\int_{-1/n}^{1/n} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - e^{-\frac{\sigma^2}{2}t^2} \right| &\leqslant \int_{-1/n}^{1/n} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - 1 \right| + \int_{-1/n}^{1/n} \frac{1}{|t|} \left| e^{-\frac{\sigma^2}{2}t^2} - 1 \right| \\ &\leqslant \int_{-1/n}^{1/n} \frac{1}{\sqrt{n}} E(|S_n f|) + \int_{-1/n}^{1/n} \frac{\sigma^2}{2} |t|.\end{aligned}$$

Mais $E(|S_n f|) \leqslant nE(|f|)$, donc on obtient $O(1/\sqrt{n})$ pour ce terme.

Soit $\Lambda(t)$ comme dans le théorème 5.2.1. Alors, si c est assez petit,

$$\begin{aligned}\int_{1/n\leqslant|t|\leqslant c\sqrt{n}} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - e^{-\frac{\sigma^2}{2}t^2} \right| &\leqslant \int_{1/n\leqslant|t|\leqslant c\sqrt{n}} \frac{1}{|t|} \left| \left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n - e^{-\frac{\sigma^2}{2}t^2} \right| \\ &\quad + \int_{1/n\leqslant|t|\leqslant c\sqrt{n}} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - \left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n \right|.\end{aligned}$$

Montrons que le deuxième terme satisfait

$$\int_{1/n\leqslant|t|\leqslant c\sqrt{n}} \frac{1}{|t|} \left| E(e^{i\frac{t}{\sqrt{n}} S_n f}) - \left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n \right| = O\left(\frac{1}{\sqrt{n}}\right).$$

Soit $e(n,t) = C\left[\frac{1}{n^{\beta-1}} + |t|\left(\frac{1}{n^{\beta-1}}\right) \star (1-dt^2)^n\right]$. Par le théorème 5.2.1,

$$\left|E(e^{i\frac{t}{\sqrt{n}}S_n f}) - \left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n\right| \leqslant e(n, t/\sqrt{n}).$$

Ainsi, il suffit de démontrer que

$$\int_{1/n\leqslant|t|\leqslant c\sqrt{n}} \frac{e(n,t/\sqrt{n})}{|t|} = O\left(\frac{1}{\sqrt{n}}\right).$$

Pour le terme $\frac{1}{n^{\beta-1}}$, l'intégrale vaut $\frac{\ln n}{n^{\beta-1}}$, qui est en $O(1/\sqrt{n})$.
Pour l'autre terme $c(n,t) = |t|\left(\frac{1}{n^{\beta-1}}\right) \star (1-dt^2)^n$, on coupe l'intégrale en deux. Pour $|t| \leqslant 1$, la convolution est bornée puisque $\frac{1}{n^{\beta-1}}$ est sommable et $(1-d\frac{t^2}{n})^n \leqslant 1$. Il reste $\int \frac{1}{|t|}\frac{|t|}{\sqrt{n}} = \frac{1}{\sqrt{n}}$.
Pour $|t| \geqslant 1$, on utilise le lemme 5.1.3, qui donne que $c(n,t) \leqslant \frac{1}{|t|n^{\beta-1}} + |t|\left(1-\frac{d}{2}t^2\right)^n$. Ainsi,

$$\int_1^{\sqrt{n}} \frac{1}{|t|}\left|c\left(n, \frac{t}{\sqrt{n}}\right)\right| \mathrm{d}t \leqslant \int_1^{\sqrt{n}} \frac{\sqrt{n}}{t^2 n^{\beta-1}} + \frac{1}{\sqrt{n}}\int_1^{\sqrt{n}} e^{-dt^2/2}\,\mathrm{d}t = O\left(\frac{1}{\sqrt{n}}\right).$$

On a donc démontré que

$$\int_{-c\sqrt{n}}^{c\sqrt{n}} \frac{1}{|t|}\left|E(e^{i\frac{t}{\sqrt{n}}S_n f}) - e^{-\frac{\sigma^2}{2}t^2}\right| \leqslant \int_{|t|\leqslant c\sqrt{n}} \frac{1}{|t|}\left|\left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n - e^{-\frac{\sigma^2}{2}t^2}\right| + O\left(\frac{1}{\sqrt{n}}\right). \tag{5.20}$$

On s'est ramené à l'étude des puissances d'une fonction, et on va donc pouvoir appliquer les mêmes méthodes que celles utilisées en probabilités. Plus précisément, l'étude de la vitesse dans le théorème central limite dans [IL71, théorème 3.4.1] repose sur les deux faits suivants :

1. Si une variable aléatoire Z vérifie $E(|Z|^2 1_{|Z|>z}) =_{+\infty} O(z^{-\delta})$ avec $0 < \delta \leqslant 1$ et, dans le cas $\delta = 1$, $E(Z^3 1_{|Z|\leqslant z}) = O(1)$, alors il existe une constante $\lambda^2 \geqslant 0$ telle que $E(e^{itZ}) = 1 - \frac{\lambda^2}{2}t^2(1+\gamma(t))$, avec $\int_{-x}^x t^2|\gamma(t)| = O(x^{3+\delta})$ quand $x \to 0$.
2. Si une fonction $\widetilde{\gamma}(t)$ vérifie $\int_{-x}^x t^2|\widetilde{\gamma}(t)| = O(x^{3+\delta})$ avec $0 < \delta \leqslant 1$, et $\sigma^2 \geqslant 0$, alors $\int_{|t|\leqslant c\sqrt{n}} \frac{1}{|t|}\left|\left(1 - \frac{\sigma^2 t^2}{2n}(1+\widetilde{\gamma}(t/\sqrt{n}))\right)^n - e^{-\frac{\sigma^2}{2}t^2}\right| = O(n^{-\delta/2})$.

Comme $m[\varphi_Y > n] = O(1/n^\beta)$ avec $\beta > 2$, on vérifie aisément que f_Y vérifie (4.6), si bien que le lemme 4.1.6 s'applique. Ainsi, la valeur propre $\lambda(t)$ de $\widehat{T}_{Y,t}$ s'écrit $E_Y(e^{itf_Y}) + \alpha t^2 + O(t^3)$ pour une certaine constante α. De plus, le fait 1 rappelé ci-dessus, appliqué à la variable aléatoire $f_Y : Y \to \mathbb{R}$, montre que $E_Y(e^{itf_Y}) = 1 - \frac{\lambda^2}{2}t^2(1+\gamma(t))$ pour une certaine fonction $\gamma(t)$ avec $\int_{-x}^x t^2|\gamma(t)| = O(x^{3+\delta})$. Comme $\lambda(t) = 1 - \frac{\sigma^2}{2m(Y)}\Lambda(t)$, on en déduit que $\Lambda(t) = t^2\big(1+m(Y)\frac{\lambda^2}{\sigma^2}\gamma(t)+O(t)\big)$. Ainsi, on peut écrire $\Lambda(t) = t^2(1+\widetilde{\gamma}(t))$ avec $\int_{-x}^x t^2|\widetilde{\gamma}(t)| = O(x^{3+\delta})$, et le fait 2 montre alors que $\int_{|t|\leqslant c\sqrt{n}} \frac{1}{|t|}\left|\left(1 - \frac{\sigma^2}{2}\Lambda\left(\frac{t}{\sqrt{n}}\right)\right)^n - e^{-\frac{\sigma^2}{2}t^2}\right| = O(n^{-\delta/2})$.

D'après (5.20), on a donc $\int_{-c\sqrt{n}}^{c\sqrt{n}} \frac{1}{|t|}\left|E(e^{i\frac{t}{\sqrt{n}}S_n f}) - e^{-\frac{\sigma^2}{2}t^2}\right| = O(n^{-\delta/2})$, ce qui conclut la preuve. □

5.5 Applications

Dans cette partie, on montre comment les théorèmes de la limite locale et de Berry-Esseen, démontrés dans le cas des tours localement Gibbs-Markov mélangeantes, peuvent en fait s'appliquer à toutes les applications qui admettent des modèles localement Gibbs-Markov (définies dans le paragraphe 1.3.4). La réduction pose des problèmes essentiellement au niveau des hypothèses d'apériodicité, car il n'est pas évident qu'une fonction apériodique pour une application $T : X \to X$ reste apériodique dans une extension de X. En général, la période peut effectivement varier d'un facteur multiplicatif fini, comme on le verra dans l'exemple 5.5.4. Pour cette raison, on se restreindra pour le théorème de la limite locale au cas apériodique.

5.5.1 Le cas Markov-transitif

Les preuves des théorèmes des deux parties précédentes ont reposé de façon importante sur le fait que les opérateurs de transfert de premier retour forment une suite d'opérateurs de renouvellement, et donc sur le fait que la tour était Markov-mélangeante. On peut en fait affaiblir cette hypothèse, de manière analogue au paragraphe 4.4.1 :

Proposition 5.5.1. *Les théorèmes 5.3.3 et 5.4.1 restent vrais si on y remplace l'hypothèse "Markov-mélangeant" par "Markov-transitif".*

Démonstration. Soit $T : X \to X$ une tour localement Gibbs-Markov transitive. Comme dans la preuve de la proposition 4.4.3, on décompose $X = X_0 \cup \ldots \cup X_{d-1}$ où T envoie X_i dans X_{i+1} pour $0 \leqslant i \leqslant d-1$, et $T^d : X_i \to X_i$ est une tour localement Gibbs-Markov mélangeante.

Soit $f : X \to \mathbb{R}$ appartenant à $\mathcal{L}_\tau$, apériodique. Pour $0 \leqslant i \leqslant d-1$, soit $f_i = (S_d f)_{|X_i}$. Montrons qu'elle est apériodique pour T^d. Sinon, $f_i = \rho + g - g \circ T^d + \lambda q$, où g et q sont définies sur X_i. On étend g et q à tout X en posant, pour $x \in X_i$ et $0 < r < d$, $g(T^r x) = g(x) - \sum_0^{r-1} \left(f(T^j x) - \frac{\rho}{d}\right) + \lambda q(x)$ et $q(T^r x) = 0$. Alors $f = \frac{\rho}{d} + g - g \circ T + \lambda q$, ce qui est absurde. Ainsi, toutes les fonctions f_i sont apériodiques. On peut leur appliquer le théorème 5.3.3 dans la tour mélangeante (X_i, T^d) et on obtient que, si J est un intervalle borné de $\mathbb{R}$ et $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$,

$$\sqrt{dn}\, m\left\{x \in X_i \mid S_{dn} f(x) \in J + k_{dn} + u(x) + v(T^{dn}x)\right\} \to m(X_i)|J| \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}},$$

où la variance σ^2 est indépendante de i, comme on l'a vu dans la preuve de la proposition 4.4.3. En sommant sur i, on obtient la conclusion du théorème 5.3.3 pour f, et pour les temps de la forme dn. Pour des temps $dn+r$ avec $0 < r < d$, on utilise le même résultat pour $f \circ T^r$,

$u - S_r f$, $v \circ T^r$ et la suite k_{dn+r}, et on obtient que

$$\sqrt{dn+r}\, m\left\{x \in X \mid S_{dn} f \circ T^r(x) \in J + k_{dn+r} + u(x) - S_r f(x) + v(T^{dn+r}x)\right\} \to |J| \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

Comme $S_{dn} f \circ T^r(x) + S_r f(x) = S_{dn+r} f(x)$, cela conclut la preuve du théorème 5.3.3 dans le cas Markov-transitif.

De la même façon, le théorème central limite avec vitesse se déduit du même résultat sur chaque X_i, pour des temps de la forme dn. On étend ensuite le résultat à des temps arbitraires : pour un temps $dn + r$ avec $r < d$, on a $\frac{1}{\sqrt{dn}}|S_{dn+r} f - S_{dn} f| \leqslant \frac{r\|f\|_\infty}{\sqrt{dn}}$, donc cela introduit une erreur d'au plus $O(1/\sqrt{n})$. □

5.5.2 Apériodicité et extensions

Dans tout ce paragraphe, $T' : X' \to X'$ sera une application préservant une mesure de probabilité m' sur un espace X'. Soient également (X, m) un espace probabilisé standard, $T : X \to X$ une application ergodique qui préserve m et $\pi : X' \to X$ une projection à fibres dénombrables telle que $m = \pi_*(m')$ et $T \circ \pi = \pi \circ T'$.

On va alors relier les propriétés d'apériodicité d'une fonction $f : X \to \mathbb{R}$ à celles de $f' = f \circ \pi : X' \to \mathbb{R}$.

Proposition 5.5.2. *La fonction f est un cobord pour T, i.e. elle s'écrit sous la forme $f = \chi - \chi \circ T$, si et seulement si f' est également un cobord pour T'.*

Démonstration. Si $f = \chi - \chi \circ T$ alors, en posant $\chi' = \chi \circ \pi$, on a $f' = \chi' - \chi' \circ T'$. En revanche, la réciproque n'est pas immédiate car, si $f' = \chi' - \chi' \circ T'$, il n'est pas évident que χ' soit constante sur les fibres de π, ce qui impliquerait qu'elle se factoriserait en $\chi \circ \pi$.

On utilise la caractérisation suivante des cobords : *soit T un endomorphisme d'un espace probabilisé (X, m). Alors une fonction mesurable f sur X peut s'écrire comme $\chi - \chi \circ T$ presque partout si et seulement si*

$$\forall \varepsilon > 0, \exists C > 0, \forall n \geqslant 1, m\{x \in X \mid |S_n f(x)| \geqslant C\} \leqslant \varepsilon. \tag{5.21}$$

Cette caractérisation, due à Schmidt, est prouvée par exemple dans [AW00].

Comme (5.21) ne fait intervenir que les distributions des sommes de Birkhoff et que $S_n f$ et $S_n f'$ ont la même distribution respectivement pour m et m', la proposition en découle immédiatement. □

L'hypothèse de dénombrabilité des fibres ne sert pas dans la proposition 5.5.2. En revanche, elle sera essentielle dans le théorème qui suit :

Théorème 5.5.3. *La fonction f est apériodique sur X pour T (i.e. elle ne s'écrit pas sous la forme $\rho + g - g \circ T + \lambda q$ où $\rho \in \mathbb{R}$, $g : X \to \mathbb{R}$ est mesurable, $\lambda > 0$ et $q : X \to \mathbb{Z}$) si et seulement si f' est apériodique sur X' pour T'.*

Démonstration. Si f est périodique sur X, i.e. $f = \rho+g-g\circ T+\lambda q$, alors $f' = \rho+g'-g'\circ T'+\lambda q'$ où $g' = g \circ \pi$ et $q' = q \circ \pi$. Ainsi, f' est périodique. La réciproque est plus compliquée. Supposons que f' est périodique, on va montrer que f l'est également en utilisant des idées de [AW00]. On peut supposer par exemple que $\lambda = 2\pi$. Quitte à remplacer m' par une de ses composantes ergodiques, on peut la supposer ergodique.

Comme la projection π est à fibres dénombrables, il existe une partie mesurable A de X' telle que π est un isomorphisme entre A et X, avec $m'(A) > 0$. On définit une fonction $\widetilde{g}$ sur X par $\widetilde{g}(x) = g'(x')$ où x' est l'unique préimage de x dans A. Quitte à remplacer f par $f - \widetilde{g} + \widetilde{g} \circ T - \rho'$ et g' par $g' - \widetilde{g} \circ \pi$, on peut supposer sans perte de généralité que $g' = 0$ sur A et que $\rho' = 0$.

Pour $x \in X$, soit $W_n(x)$ la mesure sur S^1 donnée par

$$W_n(x) = \frac{1}{n} \sum_{k=1}^{n} \delta(e^{iS_k f(x)})$$

où $\delta(y)$ est la masse de Dirac en y. Pour $u \in C^0(S^1)$, on peut trouver, par compacité, une sous-suite n_k telle que

$$\int_{S^1} u \, \mathrm{d}W_{n_k}(x) \to L(u)(x) \text{ en topologie faible} * \text{ dans } L^\infty(X).$$

Par procédé diagonal, on peut garantir cette convergence pour un ensemble dénombrable dense de fonctions de $C^0(S^1)$. Quitte à prendre encore une sous-suite, on peut même demander que $\frac{1}{n}\sum_{k=1}^{n} \int_{S^1} u \, \mathrm{d}W_{n_k}(x) \to L(u)(x)$ ponctuellement sur un ensemble $Y \subset X$ avec $m(Y) = 1$, par le théorème de Komlos ([Kom67]). Par densité, on obtient la même convergence pour tout $u \in C^0(S^1)$.

Pour $x \in Y$, l'application $u \in C^0(S^1) \mapsto L(u)(x) \in \mathbb{R}$ est une forme linéaire continue positive envoyant 1 sur 1, et est donc donnée par une mesure de probabilité P_x. De plus, ces mesures satisfont $P_{Tx}(S) = P_x(e^{if(x)}S)$ pour tout borélien S de S^1, car $W_n(Tx)(S) = W_n(x)(e^{if(x)}S) \pm \frac{2}{n}$.

Pour un certain $\varepsilon > 0$, on va montrer que

$$m\{x \mid P_x(\{1\}) \geqslant \varepsilon\} \geqslant \varepsilon. \tag{5.22}$$

Si $x' \in A \cap T'^{-k}(A)$, alors $e^{iS_k f \circ \pi(x')} = e^{i(g'(x') - g' \circ T'^k(x'))} = 1$, i.e. $S_k f \circ \pi(x') \in 2\pi\mathbb{Z}$. Ainsi,

$$\begin{aligned}\int_X W_n(x)(\{1\})\,\mathrm{d}m(x) &= \frac{1}{n}\sum_{k=1}^{n}\int_X 1(S_k f(x) \in 2\pi\mathbb{Z})\,\mathrm{d}m(x)\\ &\geqslant \frac{1}{n}\sum_{k=1}^{n}\int_{X'} 1(A \cap T'^{-k}A)\,\mathrm{d}m'(x')\\ &= \int_{X'} 1_A \cdot \left(\frac{1}{n}\sum_{k=1}^{n} 1_A \circ T'^k\right)\,\mathrm{d}m'(x') \to m'(A)^2 > 0\end{aligned}$$

par le théorème de Birkhoff. Ainsi, si n est assez grand, $\int_X W_n(x)(\{1\}) \geqslant 3\varepsilon > 0$. Par conséquent, $\int \left(\frac{1}{n}\sum_{k=1}^{n} W_{n_k}(x)\right)(\{1\}) \geqslant 2\varepsilon$ pour n assez grand. Comme $\left(\frac{1}{n}\sum_{k=1}^{n} W_{n_k}(x)\right)(\{1\}) \leqslant 1$, on obtient

$$m\left\{x \mid \left(\frac{1}{n}\sum_{k=1}^{n} W_{n_k}(x)\right)(\{1\}) \geqslant \varepsilon\right\} \geqslant \varepsilon.$$

Ainsi, l'ensemble $C = \left\{x \mid \limsup\left(\frac{1}{n}\sum_{k=1}^{n} W_{n_k}(x)\right)(\{1\}) \geqslant \varepsilon\right\}$ satisfait $m(C) \geqslant \varepsilon$. Finalement, $P_x(\{1\}) \geqslant \varepsilon$ sur C, et cela démontre (5.22).

On définit une mesure μ sur $X \times S^1$ par $\mu(U \times V) = \int_U P_x(V)\,\mathrm{d}m(x)$. Alors μ est invariante sous l'action de $T_f : (x, y) \mapsto (T(x), e^{-if(x)}y)$. [AW00] montre que, pour presque toute composante ergodique P de μ, il existe un sous-groupe compact H de S^1 et une application $\omega : X \to S^1$ telle que, si m_H est la mesure de Haar sur H,

$$\forall U \times V \subset X \times S^1,\ P(U \times V) = \int_U m_H(\omega(x)V)\,\mathrm{d}m(x).$$

De plus, il existe alors une fonction ψ à valeurs dans H telle que $e^{if(x)} = \frac{\omega(Tx)}{\omega(x)}\psi(x)$.

Si, pour toute composante P de μ, on avait $H = S^1$, alors $P = m \otimes \mathrm{Leb}$ pour tout P, donc $\mu = m \otimes \mathrm{Leb}$. C'est impossible puisque $\mu(C \times \{1\}) > 0$ tandis que $(m \otimes \mathrm{Leb})(C \times \{1\}) = 0$. Ainsi, pour un certain choix de P, $H = \mathbb{Z}/k\mathbb{Z}$, donc $\psi(x)^k = 1$, et $e^{ikf(x)} = \frac{\omega(Tx)^k}{\omega(x)^k}$. Ainsi, f est périodique sur X. □

La preuve montre seulement que la période de f sur X divise la période de $f \circ \pi$ sur X', et pas qu'elles sont égales. L'exemple qui suit montre que les périodes peuvent effectivement différer, même lorsque X' est une tour de Young.

Exemple 5.5.4. Construction d'un exemple où les périodes diffèrent : on prend $X = \{0, 1\}^{\mathbb{N}}$ muni de la mesure de Bernoulli qui donne un même poids $1/2$ à 0 et 1, avec $f(x) = x_0$ et T le décalage à gauche.

Fait : f est de période exactement 1.

Démonstration. Supposons que $f = \rho + g - g \circ T + q$ où q est à valeurs dans $N\mathbb{Z}$, $N \geqslant 2$. Alors $e^{2i\pi f/N} = \lambda e^{2i\pi g/N} e^{-2i\pi g\circ T/N}$. La fonction $e^{2i\pi f/N}$ est α-mesurable, où α est la partition de X en cylindres $[0], [1]$. Ainsi, [AD01b, théorème 3.1] assure que $e^{2i\pi g/N}$ est α_*-mesurable, où α_* est la partition engendrée par les images des éléments de α. Autrement dit, $e^{2i\pi g/N}$ est constante, puis $e^{2i\pi f/N} = \lambda$ est également constante, ce qui est absurde puisque f prend les valeurs 0 et 1. ◊

Fait : il existe une tour de Young X' au-dessus de X dans laquelle f' est de période au moins 2.

Démonstration. L'idée est de partir d'un produit gauche $X' = X \times \{-1, 1\}$, avec $T'(x, \varepsilon) = (T(x), e^{i\pi f(x)}\varepsilon)$. Si $g(x, \varepsilon) = \varepsilon$ et h est telle que $g = e^{i\pi h}$, alors $f' = h - h \circ T' + l$ où l est à valeurs dans $2\mathbb{Z}$.

X' est muni d'une structure markovienne avec 4 éléments de partition : les $[i] \times \{\varepsilon\}$ pour $i = 0, 1$ et $\varepsilon = \pm 1$. Cette structure est transitive et même mélangeante puisque $[0] \times 1$ est envoyé sur lui-même. On construit une tour X'' de base $[0] \times 1$ en considérant les retours à $[0] \times 1$. Comme la base de X'' se projette injectivement dans X, c'est en particulier une tour au-dessus de $[0]$ dans X, dans laquelle f n'est plus de période 1. ◊

5.5.3 Applications ayant un modèle localement Gibbs-Markov

Dans ce paragraphe, T sera une application sur un espace métrique mesuré (X, d, m) admettant un modèle localement Gibbs-Markov, comme dans la définition 1.3.17 : il existe un nombre fini de parties bornées deux à deux disjointes $U_1, \dots, U_N$ de X, une partition (modulo 0) finie ou dénombrable $(W_j)_{j\in J}$ de $\bigcup U_i$ dont chaque élément est inclus dans l'un des U_i, et des temps $\varphi_j \in \mathbb{N}^*$ tels que

– Pour tout $j \in J$, T^{φ_j} est un isomorphisme mesurable entre W_j et l'un des U_i, et sa distorsion $g(x) = \frac{\mathrm{d}m}{\mathrm{d}(m\circ T^{\varphi_j})}(x)$ vérifie : pour presque tous $x, y \in W_j$,

$$\left|1 - \frac{g(x)}{g(y)}\right| \leqslant C d(T^{\varphi_j}x, T^{\varphi_j}y)$$

où C est une constante indépendante de j.

– Il existe $C > 0$ tel que, pour tout $j \in J$, pour presque tous $x, y \in W_j$, pour tout $k < \varphi_j$,

$$d(T^k x, T^k y) \leqslant C d(T^{\varphi_j}x, T^{\varphi_j}y).$$

– Il existe $\lambda > 1$ tel que, pour tout $j \in J$, pour presque tous $x, y \in W_j$,

$$d(T^{\varphi_j}x, T^{\varphi_j}y) \geqslant \lambda d(x, y).$$

– $\sum \varphi_j m(W_j) < \infty$.

D'après la proposition 1.3.18, T admet alors un nombre fini de mesures invariantes absolument continues ergodiques $m_1, \dots, m_k$ sur $\Lambda = \bigcup_{n\geqslant 0} \bigcup_{i=1}^{N} T^n(U_i)$.

Dans ce cadre, on obtient les théorèmes suivants :

Théorème 5.5.5. *Supposons que $\sum_{\varphi_j>n} m(W_j) = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f : X \to \mathbb{R}$ une fonction hölderienne, avec $\int f \,\mathrm{d}m_i = 0$ pour un certain i.*

Alors il existe $\sigma^2 \geqslant 0$ tel que $\frac{1}{\sqrt{n}} S_n f \to \mathcal{N}(0, \sigma^2)$ pour la probabilité m_i. De plus, $\sigma^2 = 0$ si et seulement si il existe $\chi : X \to \mathbb{R}$ telle que $f = \chi - \chi \circ T$ sur un ensemble de mesure pleine pour m_i.

Théorème 5.5.6. *Supposons que $\sum_{\varphi_j>n} m(W_j) = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f : X \to \mathbb{R}$ une fonction hölderienne, avec $\int f \,\mathrm{d}m_i = 0$ pour un certain i. Supposons qu'il n'existe pas $\rho \in \mathbb{R}$, $g : X \to \mathbb{R}$, $\lambda > 0$ et $q : X \to \mathbb{Z}$ tels que $f = \rho + g - g \circ T + \lambda q$ sur un ensemble de mesure pleine pour m_i.*

Soit $\sigma^2 > 0$ donné par le théorème 5.5.5. Soient $J \subset \mathbb{R}$ un intervalle borné, $u, v : X \to \mathbb{R}$ deux fonctions respectivement hölderienne et mesurable, k_n une suite telle que $k_n/\sqrt{n} \to \kappa \in \mathbb{R}$. Alors

$$\sqrt{n}\, m_i \{x \in X \mid S_n f(x) \in J + k_n + u(x) + v(T^n x)\} \to |J| \frac{e^{-\frac{\kappa^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}.$$

Théorème 5.5.7. *Supposons que $\sum_{\varphi_j>n} m(W_j) = O(1/n^\beta)$ pour un certain $\beta > 2$. Soit $f : X \to \mathbb{R}$ une fonction hölderienne, avec $\int f \,\mathrm{d}m_i = 0$ pour un certain i. Soit $\sigma^2 \geqslant 0$ donné par le théorème 5.5.5, et supposons que $\sigma^2 > 0$.*

Pour $x \in W_j$, soit $f_Y(x) = \sum_0^{\varphi_j - 1} f(T^k x)$. Soit U la réunion des ensembles U_k sur lesquels m_i est équivalente à m. Supposons que $\int_U |f_Y|^2 1_{|f_Y|>z} \,\mathrm{d}m = O(z^{-\delta})$ quand $z \to +\infty$, pour un certain $0 < \delta \leqslant 1$. Supposons de plus que $\int_U |f_Y|^3 \,\mathrm{d}m < +\infty$ si $\delta = 1$.

Alors il existe $C > 0$ tel que $\forall n \in \mathbb{N}^*, \forall a \in \mathbb{R}$,

$$\left| m_i \left\{ x \mid \frac{1}{\sqrt{n}} S_n f(x) \leqslant a \right\} - P\big(\mathcal{N}(0, \sigma^2) \leqslant a\big) \right| \leqslant \frac{C}{n^{\delta/2}}.$$

Les hypothèses de ce dernier théorème sur f_Y sont automatiquement vérifiées si $f_Y \in L^p$, pour $\delta = p - 2$. De plus, elles sont vérifiées pour $\delta = \beta - 2$ si $2 < \beta < 3$, et pour $\delta = 1$ si $\beta > 3$.

Démonstration. Ces trois théorèmes sont des conséquences immédiates des théorèmes équivalents dans l'extension X' de X en forme de tour localement Gibbs-Markov construite dans le paragraphe 1.3.4. Pour montrer dans le théorème 5.5.5 que $\sigma^2 = 0$ si et seulement si f est un cobord, on utilise la proposition 5.5.2. Enfin, pour montrer que l'apériodicité de f dans X implique celle de f' dans X' (pour qu'on puisse appliquer le théorème 5.3.3), on utilise le théorème 5.5.3. □

5.5.4 Exemples

L'application de Viana (3.7) admet un modèle localement Gibbs-Markov qui satisfait de plus $\sum_{\varphi_j>n} m(W_j) = O(e^{-c\sqrt{n}})$ pour un certain $c > 0$, d'après le chapitre 3. Ainsi, les théorèmes du paragraphe 5.5.3 s'appliquent : les fonctions hölderiennes satisfont un théorème de la limite locale, et un théorème central limite avec vitesse en $1/\sqrt{n}$.

De la même façon, les applications LSV (1.25) de paramètre $\alpha \in]0, 1/2[$ sont localement Gibbs-Markov, et elles vérifient les hypothèses du paragraphe précédent pour $\beta = 1/\alpha > 2$. Ainsi, on obtient un théorème de la limite locale pour ces applications. Il faut noter que l'hypothèse d'apériodicité peut parfois être vérifiée de façon concrète en utilisant les informations sur les points périodiques du système, car le théorème 1.3.19 assure que les cobords sont nécessairement assez réguliers. Par exemple, le corollaire 1.4.4 montre que $\log |T'|$ est apériodique pour l'application LSV, si bien que cette fonction vérifie le théorème de la limite locale.

On obtient également une vitesse dans le théorème central limite en $1/\sqrt{n}$ pour toutes les fonctions hölderiennes lorsque $\alpha < 1/3$, et en $1/n^{\frac{1}{2\alpha}-1}$ lorsque $1/3 < \alpha < 1/2$, d'après le théorème 5.5.7 et les commentaires qui le suivent. De plus, lorsque $f(0) = 0$, la fonction f_Y gagne en intégrabilité, si bien que la vitesse de convergence augmente. Par exemple, même lorsque $1/3 < \alpha < 1/2$, une fonction lipschitzienne de moyenne nulle et avec $f(0) = 0$ donne lieu à une convergence en $1/\sqrt{n}$ dans le théorème central limite.

Finalement, ces résultats s'appliquent encore aux applications de Holland, ou au produit gauche étudié dans la partie 1.5 lorsque $\alpha_{\min} < 1/2$, d'après la proposition 1.5.18.

Bibliographie

[AA03] José Ferreira Alves and Vítor Araújo, *Random perturbations of nonuniformly expanding maps*, Astérisque (2003), no. 286, xvii, 25–62, Geometric methods in dynamics. I. MR2052296. Cité page 116.

[AA04] ———, *Hyperbolic times : frequency versus integrability*, Ergodic Theory Dynam. Systems **24** (2004), no. 2, 329–346. MR2054046. Cité page 117.

[Aar97] Jon Aaronson, *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997. MR1450400. Cité pages 6, 12, 16, 19, 29, 32, 40, 41, 46, 47, 57 et 155.

[ABV00] José Ferreira Alves, Christian Bonatti, and Marcelo Viana, *SRB measures for partially hyperbolic systems whose central direction is mostly expanding*, Invent. Math. **140** (2000), no. 2, 351–398. MR1757000. Cité pages 114, 116, 119, 120, 121, 122 et 127.

[AD01a] Jon Aaronson and Manfred 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), Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2001, pp. 215–223. MR1890328. Cité page 153.

[AD01b] ———, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn. **1** (2001), no. 2, 193–237. MR1840194. Cité pages 33, 152, 153, 203 et 215.

[ALP05] José Ferreira Alves, Stefano Luzzatto, and Vilton Pinheiro, *Markov structures and decay of correlations for non-uniformly expanding dynamical systems*, Ann. Inst. H. Poincaré Anal. Non Linéaire **22** (2005), no. 6, 817–839. MR2172861. Cité pages 114, 115, 116, 117, 119, 121, 122 et 123.

[Alv00] José Ferreira Alves, *SRB measures for non-hyperbolic systems with multidimensional expansion*, Ann. Sci. École Norm. Sup. (4) **33** (2000), no. 1, 1–32. MR1743717. Cité pages 114 et 119.

[AV02] José Ferreira Alves and Marcelo Viana, *Statistical stability for robust classes of maps with non-uniform expansion*, Ergodic Theory Dynam. Systems **22** (2002), no. 1, 1–32. MR1889563. Cité page 116.

[AW00] Jon Aaronson and Benjamin Weiss, *Remarks on the tightness of cocycles*, Colloq. Math. **84/85** (2000), 363–376, Dedicated to the memory of Anzelm Iwanik. MR1784202. Cité pages 212, 213 et 214.

[Bal00] Viviane Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, vol. 16, World Scientific Publishing Co. Inc., River Edge, NJ, 2000. MR1793194.

[BG03] Viviane Baladi and Sébastien Gouëzel, *A note on stretched exponential decay of correlations for the Viana-Alves map*, `arXiv.org/math.DS/0311189`, 2003. Cité page 117.

[BGT87] Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987. MR898871. Cité pages 54, 55, 58, 103 et 161.

[BP42] Salomon Bochner and Ralph S. Phillips, *Absolutely convergent Fourier expansions for non-commutative normed rings*, Ann. of Math. (2) **43** (1942), 409–418. MR0007939. Cité pages 83 et 93.

[Bre68] Leo Breiman, *Probability*, Addison-Wesley Publishing Company, Reading, Mass., 1968. MR0229267. Cité pages 206 et 207.

[Bro96] Anne Broise, *Transformations dilatantes de l'intervalle et théorèmes limites*, Astérisque **238** (1996), 1–109. MR1634271. Cité pages 51, 142, 149 et 190.

[BST03] Jérôme Buzzi, Olivier Sester, and Masato Tsujii, *Weakly expanding skew-products of quadratic maps*, Ergodic Theory Dynam. Systems **23** (2003), no. 5, 1401–1414. MR2018605. Cité page 117.

[dMvS93] Welington de Melo and Sebastian van Strien, *One-dimensional dynamics*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 25, Springer-Verlag, Berlin, 1993. MR1239171. Cité page 64.

[DS58] Nelson Dunford and Jacob T. Schwartz, *Linear Operators. I. General Theory*, With the assistance of W. G. Bade and R. G. Bartle. Pure and Applied Mathematics, Vol. 7, Interscience Publishers, Inc., New York, 1958. MR0117523. Cité pages 21, 82, 84 et 171.

[Eag76] G. K. Eagleson, *Some simple conditions for limit theorems to be mixing*, Teor. Verojatnost. i Primenen. **21** (1976), no. 3, 653–660. MR0428388. Cité page 155.

[Fel66] William Feller, *An introduction to probability theory and its applications. Vol. II*, John Wiley & Sons Inc., New York, 1966. MR0210154. Cité pages 147, 151 et 209.

[FHV03] Pierre Ferrero, Nicolai T.A. Haydn, and Sandro Vaienti, *Entropy fluctuations for parabolic maps*, Nonlinearity **16** (2003), 1203–1218. MR1986291. Cité page 53.

[FL01] Albert M. Fisher and Artur Lopes, *Exact bounds for the polynomial decay of correlation, $1/f$ noise and the CLT for the equilibrium state of a non-Hölder potential*, Nonlinearity **14** (2001), no. 5, 1071–1104. MR1862813.

[Fre87] J.B.G. Frenk, *On Banach algebras, renewal measures and regenerative processes*, CWI Tract, vol. 38, Centrum voor Wiskunde en Informatica, Amsterdam, 1987. MR906870. Cité page 93.

[GH88] Yves Guivarc'h and Jean Hardy, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov*, Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 1, 73–98. MR937957. Cité page 11.

[GLJ93] Yves Guivarc'h and Yves Le Jan, *Asymptotic winding of the geodesic flow on modular surfaces and continued fractions*, Ann. Sci. École Norm. Sup. (4) **26** (1993), 23–50. MR1209912.

[Gou04a] Sébastien Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields **128** (2004), no. 1, 82–122. MR2027296. Cité pages 13, 142 et 182.

[Gou04b] ———, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math. **139** (2004), 29–65. MR2041223. Cité pages 13 et 77.

[Gou05] ———, *Berry-Esseen theorem and local limit theorem for non uniformly expanding maps*, Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 6, 997–1024. MR2172207. Cité pages 13 et 190.

[Gou06] ———, *Decay of correlations for nonuniformly expanding systems*, Bull. Soc. Math. France **134** (2006), no. 1, 1–31. MR2233699. Cité pages 13 et 114.

[Gou07] ———, *Statistical properties of a skew product with a curve of neutral points*, Ergodic Theory Dynam. Systems **27** (2007), no. 1, 123–151. MR2297091. Cité pages 12, 16 et 142.

[Hen93] Hubert Hennion, *Sur un théorème spectral et son application aux noyaux lipchitziens*, Proc. Amer. Math. Soc. **118** (1993), no. 2, 627–634. MR1129880. Cité page 23.

[Hol05] Mark Holland, *Slowly mixing systems and intermittency maps*, Ergodic Theory Dynam. Systems **25** (2005), no. 1, 133–159. MR2122916. Cité page 54.

[Hu01] Huyi Hu, *Statistical properties of some almost hyperbolic systems*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 367–384. MR1858539. Cité page 58.

[Hu04] ———, *Decay of correlations for piecewise smooth maps with indifferent fixed points*, Ergodic Theory Dynam. Systems **24** (2004), no. 2, 495–524. MR2054191.

[IL71] Ildar A. Ibragimov and Yuri V. Linnik, *Independent and stationary sequences of random variables*, Wolters-Noordhoff Publishing, Groningen, 1971, With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. MR0322926. Cité pages 209 et 210.

[ITM50] Cassius T. Ionescu Tulcea and Gheorghe Marinescu, *Théorie ergodique pour des classes d'opérations non complètement continues*, Ann. of Math. (2) **52** (1950), 140–147. MR0037469. Cité page 23.

[Kac96] Alexander G. Kachurovskiĭ, *Rates of convergence in ergodic theorems*, Russian Math. Surveys **51** (1996), 653–703. MR1422228. Cité pages 154 et 162.

[Kah70] Jean-Pierre Kahane, *Séries de Fourier absolument convergentes*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 50, Springer-Verlag, Berlin, 1970. MR0275043. Cité pages 172 et 173.

[Kel80] Gerhard Keller, *Un théorème de la limite centrale pour une classe de transformations monotones par morceaux*, C. R. Acad. Sci. Paris **291** (1980), A155–A158. MR605005.

[Kom67] János Komlós, *A generalization of a problem of Steinhaus*, Acta Math. Acad. Sci. Hungar. **18** (1967), 217–229. MR0210177. Cité pages 155 et 213.

[Liv96] Carlangelo Liverani, *Central limit theorem for deterministic systems*, International Conference on Dynamical Systems (Montevideo, 1995), Pitman Res. Notes Math. Ser., vol. 362, Longman, Harlow, 1996, pp. 56–75. MR1460797. Cité page 7.

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685. MR1695915. Cité pages 9, 52 et 65.

[Man52] Szolem Mandelbrojt, *Séries adhérentes, régularisation des suites, applications*, Gauthier-Villars, Paris, 1952. MR0051893. Cité page 91.

[MS80] Calvin C. Moore and Klaus Schmidt, *Coboundaries and homomorphisms for non-singular actions and a problem of H. Helson*, Proc. London Math. Soc. (3) **40** (1980), no. 3, 443–475. MR572015. Cité page 204.

[MT04] Ian Melbourne and Andrei Török, *Statistical limit theorems for suspension flows*, Israel J. Math. **144** (2004), 191–209. MR2121540. Cité page 154.

[Pèn02] Françoise Pène, *Averaging method for differential equations perturbed by dynamical systems*, ESAIM Probab. Statist. **6** (2002), 33–88 (electronic). MR1905767. Cité pages 59 et 69.

[PY01] Mark Pollicott and Michiko Yuri, *Statistical properties of maps with indifferent periodic points*, Comm. Math. Phys. **217** (2001), no. 3, 503–520. MR1822105. Cité page 58.

[RE83] Jacques Rousseau-Egele, *Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux*, Ann. Probab. **11** (1983), no. 3, 772–788. MR704569.

[Rud91] Walter Rudin, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR1157815. Cité pages 83 et 93.

[Sar02] Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653. MR1946554. Cité pages 9, 78, 87 et 101.

[Sar03] ———, *Existence of Gibbs measures for countable Markov shifts*, Proc. Amer. Math. Soc. **131** (2003), no. 6, 1751–1758. MR1955261. Cité page 40.

[Spi64] Frank Spitzer, *Principles of random walk*, The University Series in Higher Mathematics, D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London, 1964. MR0171290. Cité page 35.

[Via97] Marcelo Viana, *Multidimensional nonhyperbolic attractors*, Inst. Hautes Études Sci. Publ. Math. (1997), no. 85, 63–96. MR1471866. Cité pages 58, 60, 114 et 116.

[You98] Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650. MR1637655. Cité pages 50 et 123.

[You99] ———, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188. MR1750438. Cité pages 9, 49, 78, 100, 114, 123, 137, 138, 139 et 140.

Résumé

Dans ce travail, nous nous intéressons aux propriétés de mélange des applications non uniformément dilatantes, par exemple les applications en dimension 1 ayant un point fixe neutre. Après un premier chapitre introductif rappelant des résultats bien connus sur les applications Gibbs-Markov, nous établissons au chapitre deux des résultats très précis sur la vitesse de décorrélation des applications localement Gibbs-Markov (bornes inférieures et supérieures), en appliquant des techniques d'algèbres de Banach aux opérateurs de transfert de premier retour. Dans le chapitre trois, nous montrons en utilisant des techniques de temps hyperboliques qu'une large classe d'applications admet un modèle localement Gibbs-Markov, ce qui permet de leur appliquer les résultats du deuxième chapitre.

Les deux derniers chapitres sont consacrés à des théorèmes de type probabiliste pour les sommes de Birkhoff de fonctions hölderiennes, dans le cadre des applications localement Gibbs-Markov. En étendant à un cadre perturbatif les techniques spectrales du chapitre deux, nous montrons d'une part la convergence des sommes de Birkhoff vers des lois normales ou des lois stables (suivant le système considéré), et d'autre part des théorèmes limites fins comme le théorème de Berry-Esseen (avec vitesse optimale) ou le théorème de la limite locale, sous des hypothèses légèrement plus fortes.

Abstract

In this thesis, we study the mixing properties of non uniformly expanding maps, for example maps in dimension 1 with a neutral fixed point. After an introductory chapter recalling well known facts on Gibbs-Markov maps, we establish in chapter 2 very precise results on the speed of decay of correlations of locally Gibbs-Markov maps (lower and upper bounds), by applying Banach algebra techniques to the first return transfer operators. In chapter 3, we show – using hyperbolic times techniques – that a large class of maps admit locally Gibbs-Markov models, which makes it possible to apply the results of the previous chapter to these maps.

The last two chapters are devoted to probabilistic limit theorems for the Birkhoff sums of Hölder functions, in the setting of locally Gibbs-Markov maps. We extend to a perturbative setting the spectral results of chapter 2, and prove on the one hand the convergence of the Birkhoff sums to normal laws or stable laws (depending on the system we are considering), and on the other hand fine limit theorems, the Berry-Esseen theorem (with optimal speed) and the local limit theorem, under slightly stronger assumptions.