Université de Rennes 1

Institut de Recherche Mathématique de Rennes

École doctorale Matisse

Comportement quantitatif de certains systèmes dynamiques. Exemples et applications

Sébastien Gouëzel

Document de synthèse présenté le

22 novembre 2012

en vue de l'obtention du diplôme

d'HABILITATION À DIRIGER DES RECHERCHES EN MATHÉMATIQUES

Pascal Hubert	Aix-Marseille Université	(rapporteur)
François Ledrappier	University of Notre Dame	(rapporteur)
Ian Melbourne	University of Warwick	(rapporteur)
Bachir Bekka	Université de Rennes 1	(examinateur)
Jérôme Buzzi	Université Paris Sud 11	(examinateur)
Pierre Collet	École Polytechnique	(examinateur)
Anton Zorich	Université Paris 7	(examinateur)

Table des matières

Remerciements 3

Introduction 5

Chapitre I. Quelques exemples et résultats significatifs 7
1. Vitesse de mélange et théorèmes limites 8
2. Vitesse de mélange pour des applications de l'intervalle 10
3. Résonances de Ruelle 14
4. La méthode spectrale de Nagaev-Guivarc'h 16
5. Le billard stade 18

Chapitre II. Caractérisation de la convergence des sommes de Birkhoff pour les applications Gibbs-Markov 21
1. Théorèmes limites pour des sommes de variables aléatoires i.i.d. 21
2. Théorèmes limites pour les applications Gibbs-Markov 22
3. Stratégie de preuve 24
4. Utilisation des développements caractéristiques précis 26
5. Développement caractéristique pour les applications Gibbs-Markov 28

Chapitre III. Principe d'invariance presque sûr par la méthode spectrale 31
1. Résultat principal probabiliste 31
2. Applications en dynamique 34
3. Démonstration du théorème III.1.3 37

Chapitre IV. Étude statistique fine des applications Anosov lisses 43
1. Du cas dilatant au cas hyperbolique 43
2. Différents espaces candidats dans le cas hyperbolique 46
3. Un espace de Banach convenable en général 53

Chapitre V. Spectre du laplacien sur l'espace des surfaces plates 57
1. Définitions et résultat principal 57
2. Espaces de Banach distributionnels pour le flot de Teichmüller 60
3. Relation entre informations géométriques et algébriques sur l'action de $\mathrm{SL}(2, \mathbb{R})$ 66

Bibliographie personnelle 71

Bibliographie générale 73

Remerciements

Pascal HUBERT et François LEDRAPPIER ont accepté de consacrer du temps à la lecture de mon mémoire, et d'écrire un rapport à son sujet. Je les en remercie vivement. Ian MELBOURNE s'est attelé à la même tâche, mais il a en plus été confronté à la barrière de la langue, particulièrement importante dans un rapport qui se veut plus didactique que rempli de formules. Je l'en remercie donc d'autant plus.

Mes remerciements vont également à Bachir BEKKA, Jérôme BUZZI, Pierre COLLET et Anton ZORICH, qui m'ont fait l'honneur d'être membres de mon jury d'habilitation à diriger des recherches.

La vie quotidienne à l'IRMAR doit beaucoup à la présence de Florent, Jürgen, Ismaël, François, Serge (à mi-temps), Nizar, Ludovic, Stéphane, qu'ils en soient remerciés.

À Agathe, Gaétan, Quentin et Rémi, qui remplissent ma vie de bonheur.

Introduction

PLUTÔT qu'une compilation exhaustive de résultats, cette habilitation est une synthèse sélective de mes travaux : j'ai choisi quelques résultats, et je les ai présentés relativement en détail, avec au moins des ébauches de preuve (mais en évitant soigneusement les parties trop techniques). J'ai choisi des résultats suffisamment variés pour que les arguments utilisés d'un chapitre à l'autre soient vraiment différents, tout en conservant un certain fil directeur : en un mot, il s'agit d'établir des propriétés quantitatives de certains systèmes dynamiques, parfois en vue d'en déduire des comportements de type probabiliste.

Dans le chapitre II, je considère des systèmes dynamiques très dilatants (plus précisément, des systèmes Gibbs-Markov), et je me demande dans quels cas les sommes de Birkhoff $S_n f$ d'une fonction f ayant un peu de régularité peuvent vérifier un théorème limite : si $(S_n f - A_n)/B_n$ converge en distribution vers une loi limite non triviale, que peut-on dire sur celle-ci ? La difficulté est qu'on ne met aucune hypothèse d'intégrabilité sur f, ni aucune hypothèse sur l'asymptotique de B_n, pour éventuellement détecter tous les phénomènes qui pourraient être nouveaux dans le cas dynamique. Il se trouve qu'il n'y a en fait rien de nouveau par rapport au cas indépendant : la loi limite est nécessairement gaussienne ou stable, et les normalisations sont analogues au cas indépendant.

Dans le chapitre III, je m'intéresse à un raffinement fort du théorème central limite, appelé principe d'invariance presque sûr : l'enjeu est d'approcher presque sûrement des sommes de Birkhoff par des trajectoires browniennes, avec un contrôle assez précis. Je démontre que c'est possible dès que la fonction caractéristique du processus considéré vérifie des propriétés d'indépendance asymptotique assez fortes (ce qui est bien connu dans nombre de systèmes dynamiques grâce à des arguments spectraux). La preuve est principalement probabiliste, faisant intervenir des arguments de couplage et des techniques classiques de « gros blocs, petits blocs ».

Dans le chapitre IV, j'explique comment établir des asymptotiques précises des corrélations dans des systèmes uniformément hyperboliques très lisses. Il faut pour cela éviter les techniques de codage, qui perdent presque toute l'information sur la régularité des systèmes. L'idée est d'introduire certains espaces de Banach sur lesquels la dynamique agit avec un rayon spectral essentiel arbitrairement petit. Après la présentation de cas particuliers où l'on peut utiliser des espaces de distributions, j'explique pourquoi, dans le cas général, l'espace le mieux adapté à ces questions est un espace de courants généralisés. En plus de techniques analytiques et spectrales, on devra donc faire intervenir un peu de géométrie.

Finalement, le chapitre V est consacré à l'étude du spectre du laplacien sur l'espace des modules des surfaces de genre $\geqslant 2$. Il s'agit de démontrer un analogue dans un contexte non homogène du fait que le spectre du laplacien est discret dans $[0, 1/4]$ sur les surfaces hyperboliques de volume fini. Le laplacien (feuilleté) est difficile à étudier directement car ce n'est pas un opérateur elliptique, la preuve passe plutôt par la compréhension d'un flot appelé flot de Teichmüller. Comme celui-ci a des propriétés hyperboliques, on peut faire appel aux techniques du chapitre IV. Pour relier les propriétés spectrales du flot au spectre du laplacien, on a aussi besoin d'utiliser des propriétés fines des représentations unitaires de $\mathrm{SL}(2, \mathbb{R})$; les techniques de ce chapitre sont donc à la fois géométriques et algébriques.

Ces quatre chapitres sont précédés d'un chapitre introductif qui présente des résultats extrêmement classiques du domaine, qu'il faut connaître pour pouvoir lire la suite. Dans celui-ci, j'ai essayé autant que possible de présenter des preuves complètes de certains résultats élémentaires qui servent à la fois d'introduction et de motivation aux énoncés des chapitres suivants.

Il ressort du plan précédent que je vais ignorer certains aspects de mes travaux. En particulier, je ne parlerai pas d'application intermittentes, même si j'y ai consacré plusieurs articles (voir [Gouëzel 2004a,b, 2005, 2007a,b]). Je ne parlerai pas non plus d'applications hyperboliques par morceaux [Baladi et Gouëzel 2009, 2010], des travaux qui s'inscrivent dans la continuité des résultats du chapitre IV. Je ne mentionnerai pas non plus certains résultats sur les théorèmes limites comme [Chazottes et Gouëzel 2007, 2012 ; Dedecker, Gouëzel et Merlevède 2010, 2012]. Enfin, je ne parlerai ni de mes résultats sur le théorème local pour les produits gauches [Gouëzel 2009b] car ils me semblent trop délicats à exposer de manière intelligible en quelques pages, ni de marches aléatoires sur les groupes hyperboliques [Gouëzel et Lalley 2011 ; Gouëzel 2012] car c'est un sujet un peu à l'écart de mes thématiques habituelles. Une liste complète de mes travaux est donnée page 71.

CHAPITRE I

Quelques exemples et résultats significatifs

Sommaire

1. **Vitesse de mélange et théorèmes limites** 8
2. **Vitesse de mélange pour des applications de l'intervalle** 10
3. **Résonances de Ruelle** 14
4. **La méthode spectrale de Nagaev-Guivarc'h** 16
5. **Le billard stade** 18

Un des résultats d'universalité les plus frappants en mathématiques est donné par le théorème central limite : quelle que soit la suite $(Z_n)_{n\in\mathbb{N}}$ de variables aléatoires centrées indépendantes identiquement distribuées dans L^2, la suite $(\sum_{k=0}^{n-1} Z_k)/\sqrt{n}$ converge en distribution vers une limite gaussienne. Si les Z_n ne sont pas identiquement distribués mais sont encore indépendants, un résultat analogue reste valide si la contribution individuelle de chaque Z_i est asymptotiquement négligeable. Ce résultat est utilisé pour justifier l'apparition de gaussiennes en présence de petites perturbations indépendantes. Cependant, cette hypothèse d'indépendance n'est pas si facile à justifier dans notre monde dont les lois physiques sont, au moins à une échelle macroscopique, complètement déterministes. Il est donc naturel de chercher à l'affaiblir. Différentes hypothèses de dépendance faible ont été introduites dans cette optique dans la littérature probabiliste.

On s'intéressera dans ce texte principalement à des suites produites par l'itération d'une transformation déterministe, comme suit. Soient (X, μ) un espace muni d'une mesure de probabilité μ, et $T : X \to X$ une transformation préservant μ, ergodique. Pour toute fonction $f : X \to \mathbb{R}$, les variables aléatoires $Z_n = f \circ T^n$ sont identiquement distribuées (mieux, elles forment un processus stationnaire, i.e., la loi de $(Z_{n+k})_{n\in\mathbb{N}}$ est la même pour tout $k \in \mathbb{N}$), mais elles ne sont pas indépendantes sauf cas très particulier. D'après [Volný 1990], il existe toujours des fonctions f bornées dont les sommes de Birkhoff $S_n f = \sum_{k=0}^{n-1} f \circ T^k$ ont un comportement pathologique, il faut donc se restreindre à des fonctions f ayant des propriétés supplémentaires pour espérer des théorèmes limites.

Partant de cette problématique, on décrira dans ce chapitre introductif quelques résultats généraux et quelques exemples simples (ou moins simples), classiques. Ce parcours rapide sera surtout l'occasion d'introduire différents sujets dignes d'intérêt qui seront développés dans les chapitres suivants.

Dans la section 1, on énoncera un théorème de GORDIN qui montre que les sommes de Birkhoff satisfont un théorème central limite dès que la vitesse de mélange est assez grande. Dans la section 2, on décrira comment estimer la vitesse de mélange, en particulier pour des applications dilatantes de l'intervalle, ce qui impliquera un théorème central limite d'après ce qui précède. Dans la section 3, on évoquera brièvement la notion de résonance de RUELLE. Dans la section 4, on expliquera une autre méthode pour obtenir des théorèmes limites pour des systèmes dynamiques, la méthode spectrale de NAGAEV-GUIVARC'H. Enfin, dans la section 5, on décrira un exemple dynamiquement plus complexe, le *billard stade*, dans lequel le mélange est assez lent, et dans lequel le théorème central limite n'a pas lieu avec une normalisation $1/\sqrt{n}$.

1. Vitesse de mélange et théorèmes limites

Si f et $g \circ T^k$ étaient indépendantes, les corrélations $\int f \cdot g \circ T^k - \int f \cdot \int g$ seraient nulles. L'indépendance asymptotique de f et $g \circ T^k$ peut donc se décrire par la vitesse de décroissance des corrélations, aussi appelée vitesse de mélange. Elle joue souvent un rôle dans l'existence de théorèmes limites, comme dans le théorème suivant :

THÉORÈME 1.1 (GORDIN-HANNAN). *Soit $T : X \to X$ une transformation mesurable préservant une mesure de probabilité μ ergodique. Soit $f : X \to \mathbb{R}$ une fonction de carré intégrable de moyenne nulle. Supposons qu'il existe une suite ρ_n sommable telle que, pour toute fonction $g \in L^2$ et pour tout $n \in \mathbb{N}$, on ait*

$$\left| \int f \cdot g \circ T^n \right| \leqslant \rho_n \|g\|_{L^2} . \tag{1.1}$$

Alors les sommes de Birkhoff de f satisfont un théorème central limite : $S_n f/\sqrt{n}$ converge en distribution vers une gaussienne $\mathcal{N}(0, \sigma^2)$, où $\sigma^2 = \int f^2 + 2\sum_{k>0} \int f \cdot f \circ T^k \geqslant 0$.

De plus, $\sigma^2 = 0$ si et seulement si f est un cobord dans L^2, i.e., s'il existe $u \in L^2$ telle que $f = u - u \circ T$ dans L^2.

La convergence de la série définissant σ^2 est une conséquence de la sommabilité des ρ_n. De plus, la variance de $S_n f$ est $n \int f^2 + 2 \sum_{k=1}^{n} (n-k) \int f \cdot f \circ T^k$, ce qui montre que σ^2 est la limite de la variance de $S_n f/\sqrt{n}$, et est donc positif ou nul.

Notons que ce théorème est sans intérêt si T est inversible, puisque l'hypothèse n'est jamais satisfaite : en prenant $g = f \circ T^{-n}$, on voit que $\rho_n \geqslant \|f\|_{L^2}$ pour tout n, ce qui n'est pas sommable si f est non nulle. Du point de vue des théorèmes limites, les applications non inversibles sont ainsi souvent plus faciles à analyser que les applications inversibles. Une technique efficace et très classique pour traiter le cas inversible est alors de se ramener au cas non inversible, en considérant une sous-tribu pour laquelle T est mesurable mais non inversible (par exemple en quotientant par les variétés stables locales).

La preuve du théorème 1.1 est très instructive, puisqu'elle est l'occasion d'introduire un objet crucial dans la suite, l'*opérateur de transfert*, et d'illustrer la puissance de certains outils probabilistes (ici, une méthode de martingales).

Démonstration. Soient $\mathcal{A}$ la tribu des ensembles mesurables, et $\mathcal{A}_n = T^{-n}\mathcal{A}$: c'est une suite décroissante de tribus. Une fonction est mesurable pour $\mathcal{A}_n$ si et seulement si elle est constante sur les préimages de presque tout point par T^n, i.e., si elle s'écrit sous la forme $g \circ T^n$. L'équation (1.1) signifie donc que $\|E(f \mid \mathcal{A}_n)\|_{L^2} \leqslant \rho_n$.

Supposons tout d'abord que $E(f \mid \mathcal{A}_1) = 0$, i.e., la moyenne de f sur les préimages de presque tout point est nulle. Comme $f \circ T^k$ est mesurable par rapport à $\mathcal{A}_1$ pour tout $k \geqslant 1$, on a donc une forme faible d'indépendance entre f et les $f \circ T^k$ pour $k \geqslant 1$. En composant par T^n, on obtient que $E(f \circ T^n \mid \mathcal{A}_{n+1}) = 0$ tandis que les $f \circ T^k$ sont mesurables par rapport à $\mathcal{A}_{n+1}$ pour $k > n$. On dit que les $f \circ T^n$ forment une *suite de différences de martingales* pour la suite (décroissante) de tribus $\mathcal{A}_n$. Des résultats classiques de probabilités impliquent directement que $S_n f / \operatorname{Var}(S_n f)$ converge en distribution vers $\mathcal{N}(0,1)$ si $\operatorname{Var}(S_n f) \to \infty$. On vérifie aisément que $\operatorname{Var}(S_n f) = n\sigma^2 + O(1)$, où σ^2 est comme dans l'énoncé du théorème. Ainsi, $\sigma^2 \geqslant 0$, et on obtient un théorème central limite si $\sigma^2 > 0$. Si $\sigma^2 = 0$, $S_n f$ est borné dans L^2, ce qui implique que f est un cobord (théorème de Leonov, qui se démontre en regardant une limite faible u de $n^{-1} \sum_{k=0}^{n-1} S_k f$ et en remarquant que $f = u - u \circ T$).

Plus généralement, si on peut écrire $f = g + u - u \circ T$ où $E(g \mid \mathcal{A}_1) = 0$ et $u \in L^2$, alors $S_n f = S_n g + u - u \circ T^n$. Comme $(u - u \circ T^n)/\sqrt{n}$ tend en probabilité vers 0, on en déduit comme dans le paragraphe précédent que $S_n f / \sqrt{n}$ vérifie un théorème central limite.

Sous les hypothèses du théorème, j'affirme qu'une telle décomposition $f = g + u - u \circ T$ existe toujours. Pour la construire, il est utile de reformuler les espérances conditionnelles par rapport aux $\mathcal{A}_n$ comme les itérés d'un seul opérateur, comme suit. Pour toute fonction g, la fonction $E(g \mid \mathcal{A}_1)$ s'écrit sous la forme $h \circ T$, il est donc naturel de définir un opérateur $\hat{T}$ par $\hat{T} g = h$. On a alors, pour toute fonction u,

$$\int g \cdot u \circ T = \int E(g \mid \mathcal{A}_1) \cdot u \circ T = \int (\hat{T} g) \circ T \cdot u \circ T = \int \hat{T} g \cdot u.$$

Autrement dit, $\hat{T}$ est l'adjoint de l'opérateur de composition par T. C'est l'*opérateur de transfert*. Par construction, il vérifie

$$\hat{T}(g \circ T) = g\ ; \quad (\hat{T} g) \circ T = E(g \mid \mathcal{A}_1)\ ; \quad (\hat{T}^n g) \circ T^n = E(g \mid \mathcal{A}_n).$$

S'il existe une décomposition $f = g + u - u \circ T$ avec $E(g \mid \mathcal{A}_1) = 0$, on obtient $\hat{T} f = \hat{T} u - u$, puis $\hat{T}^n f = \hat{T}^n u - \hat{T}^{n-1} u$. Par conséquent, $\sum_{k=1}^n \hat{T}^k f = \hat{T}^n u - u$. Si $\hat{T}^n u$ tend vers 0 (ce qu'on peut attendre si u est de moyenne nulle), on aura donc $u = -\sum_{k>0} \hat{T}^k f$.

Posons donc $u = -\sum_{k>0} \hat{T}^k f$. Comme $\left\| \hat{T}^k f \right\|_{L^2} \leqslant \rho_k$ est sommable, cette fonction est bien définie dans L^2. Soit $g = f - u + u \circ T$, alors

$$\hat{T} g = \hat{T} f + \sum_{k>0} \hat{T}^{k+1} f - \sum_{k>0} \hat{T}^k f = 0.$$

Ainsi, la décomposition $f = g + u - u \circ T$ satisfait les propriétés requises. Cela montre que les sommes de Birkhoff de f vérifient un théorème central limite. □

Il existe de nombreuses variantes du théorème précédent, remplaçant par exemple L^2 par L^1 ou affaiblissant l'hypothèse de sommabilité. Je citerai juste le résultat suivant, également dû à GORDIN et HANNAN, utile dans le cas inversible. Il pourrait se formuler avec des hypothèses sur les corrélations, mais il est plus naturel de l'écrire en termes d'espérances conditionnelles.

THÉORÈME 1.2. *Soit $T : X \to X$ une transformation mesurable inversible préservant une mesure de probabilité μ ergodique. Soit $\tilde{\mathcal{A}}$ une sous-tribu pour laquelle T est mesurable, i.e., $T^{-1}\tilde{\mathcal{A}} \subset \tilde{\mathcal{A}}$. Soit $f : X \to \mathbb{R}$ une fonction de carré intégrable de moyenne nulle telle que*

$$\sum_{n>0} \left\| E(f \mid T^{-n}\tilde{\mathcal{A}}) \right\|_{L^2} < \infty \quad \text{et} \quad \sum_{n<0} \left\| f - E(f \mid T^{-n}\tilde{\mathcal{A}}) \right\|_{L^2} < \infty. \tag{1.2}$$

Alors les sommes de Birkhoff de f satisfont un théorème central limite : $S_n f/\sqrt{n}$ converge en distribution vers une gaussienne $\mathcal{N}(0, \sigma^2)$, où $\sigma^2 = \int f^2 + 2\sum_{k>0} \int f \cdot f \circ T^k \geqslant 0$.

De plus, $\sigma^2 = 0$ si et seulement si f est un cobord dans L^2, i.e., s'il existe $u \in L^2$ telle que $f = u - u \circ T$ dans L^2.

La preuve de ce résultat est analogue à celle du théorème 1.1, utilisant une décomposition de f en somme de martingale et cobord. Lorsque f est mesurable par rapport à $\tilde{\mathcal{A}}$, ce théorème se réduit au précédent. L'intuition à avoir est que la tribu $T^{-n}\tilde{\mathcal{A}}$ converge vers la tribu triviale $\{\emptyset, X\}$ lorsque n tend vers $+\infty$, et vers la tribu des boréliens lorsque n tend vers $-\infty$. Lorsque cette convergence est assez rapide, les sommes (1.2) ont une chance d'être finies.

2. Vitesse de mélange pour des applications de l'intervalle

Dans ce paragraphe, nous décrivons une classe d'exemples pour lesquels il est assez aisé d'estimer la vitesse de décorrélation, si bien que les résultats du paragraphe précédent s'appliquent. Il s'agit des applications dilatantes de l'intervalle avec branches complètes.

Le premier exemple d'une telle application, et assurément le plus simple, est l'application doublement sur l'intervalle $[0, 1]$, donnée par $T(x) = 2x \mod 1$. Elle préserve la mesure de probabilité $\mu = \mathrm{Leb}$.

PROPOSITION 2.1. *Soit f une fonction lipschitzienne, de moyenne nulle et de meilleure constante de Lipschitz $\mathrm{Lip}(f)$. Pour toute fonction g dans L^1 et pour tout $n \in \mathbb{N}$, on a*

$$\left| \int f \cdot g \circ T^n \right| \leqslant 2^{-n} \,\mathrm{Lip}(f) \, \|g\|_{L^1} .$$

D'après le théorème 1.1, il s'ensuit que les sommes de Birkhoff de f satisfont un théorème central limite.

DÉMONSTRATION. Comme T est un morphisme de groupes sur $\mathbb{R}/\mathbb{Z}$, on peut aisément faire le calcul en utilisant des séries de Fourier. Cependant, cet argument ne se généralise pas bien, contrairement à celui qui suit, qui repose sur l'utilisation de l'opérateur de transfert.

Définissons momentanément un opérateur $\mathcal{L}$, agissant sur les fonctions bornées par $\mathcal{L}u(x) = \frac{1}{2}\sum_{T(y)=x} u(y)$. On a alors

$$\begin{aligned}\int \mathcal{L}u \cdot v &= \frac{1}{2}\left[\int u(x/2)v(x)\,\mathrm{d}x + \int u(x/2+1/2)v(x)\,\mathrm{d}x\right]\\ &= \int_{[0,1/2]} u(y)v(2y)\,\mathrm{d}y + \int_{[1/2,1]} u(y)v(2y-1)\,\mathrm{d}y\\ &= \int_{[0,1]} u \cdot v \circ T.\end{aligned}$$

Ainsi, $\mathcal{L}$ est l'opérateur de transfert $\hat{T}$ associé à T.

Si u est une fonction lipschitzienne, alors $x \mapsto u(x/2)$ est lipschitzienne de constante de Lipschitz au plus $\mathrm{Lip}(u)/2$, et il en va de même de $x \mapsto u(x/2+1/2)$. Ainsi,

$$\mathrm{Lip}(\hat{T}u) \leqslant \mathrm{Lip}(u)/2. \tag{2.1}$$

En particulier, $\mathrm{Lip}(\hat{T}^n f) \leqslant 2^{-n}\,\mathrm{Lip}(f)$. Comme $\hat{T}^n f$ est de moyenne nulle, elle prend des valeurs positives et négatives. Ainsi, $\left\|\hat{T}^n f\right\|_{L^\infty} \leqslant \mathrm{Lip}(\hat{T}^n f) \leqslant 2^{-n}\,\mathrm{Lip}(f)$. Finalement, pour toute fonction g dans L^1,

$$\left|\int f \cdot g \circ T^n\right| = \left|\int \hat{T}^n f \cdot g\right| \leqslant \left\|\hat{T}^n f\right\|_{L^\infty} \|g\|_{L^1} \leqslant 2^{-n}\,\mathrm{Lip}(f)\,\|g\|_{L^1}. \qquad \square$$

La preuve précédente s'adapte facilement au cas des fonctions hölderiennes. Il est cependant plus rapide (et, dans d'autres cas, nettement plus efficace) d'utiliser un argument d'interpolation (voir [Bergh et Löfström 1976]) comme suit. L'espace des fonctions hölderiennes C^α (pour $0 < \alpha < 1$) est un espace d'interpolation (de paramètre α) entre C^0 et C^1. Ainsi, il existe une constante C_α telle que toute forme linéaire $U : C^0 \to \mathbb{C}$ satisfait

$$\|U\|_{C^\alpha \to \mathbb{C}} \leqslant C_\alpha \|U\|_{C^0\to\mathbb{C}}^{1-\alpha} \|U\|_{C^1\to\mathbb{C}}^{\alpha}.$$

Fixons $n \in \mathbb{N}$ et g une fonction dans L^1, et considérons $U : f \mapsto \int f \cdot g \circ T^n - \int f \cdot \int g$. Cette forme linéaire est bornée par $\|g\|_{L^1}$ sur C^0, et par $2^{-n}\|g\|_{L^1}$ sur C^1, par la proposition précédente. Elle est donc bornée par $C_\alpha 2^{-\alpha n}$ sur C^α, i.e.,

$$\left|\int f \cdot g \circ T^n - \int f \cdot \int g\right| \leqslant C_\alpha 2^{-\alpha n} \|f\|_{C^\alpha} \|g\|_{L^1}.$$

Le théorème 1.1 implique donc que les sommes de Birkhoff de fonctions hölderiennes satisfont également un théorème central limite.

Remarque 2.2. L'argument d'interpolation précédent pour passer d'estimées dans un espace fort comme C^1 à des espaces plus faibles n'a rien de spécifique à l'application de doublement. Cet argument est dû dans un contexte dynamique à Dolgopyat et Dinh–Sibony.

Passons maintenant à des exemples un peu plus complexes.

Définition 2.3. *Une transformation* $T : [0,1] \to [0,1]$ *est dilatante par morceaux* C^r *à branches surjectives s'il existe une partition finie de* $[0,1]$ *en intervalles* $I_1, \dots, I_k$ *tels que la restriction de* T *à* I_k *admette une extension de classe* C^r *à* $\overline{I_k}$*, surjective, dont la dérivée est partout strictement plus grande que* 1 *en valeur absolue.*

Théorème 2.4. *Soit* $T : [0,1] \to [0,1]$ *une application dilatante par morceaux* C^2 *à branches surjectives. Elle admet une unique mesure de probabilité* μ *invariante absolument continue par rapport à la mesure de Lebesgue. De plus, il existe* $C > 0$ *et* $\rho < 1$ *tels que, pour toute fonction* f *lipschitzienne de moyenne nulle pour* μ *et pour toute fonction* g *dans* L^1,

$$\left| \int f \cdot g \circ T^n \, \mathrm{d}\mu \right| \leqslant C \rho^n \operatorname{Lip}(f) \, \|g\|_{L^1} .$$

L'estimée du théorème s'étend comme ci-dessus aux fonctions hölderiennes. Le théorème 1.1 implique donc que les sommes de Birkhoff de fonctions hölderiennes satisfont un théorème central limite.

Remarque 2.5. Dans le théorème précédent, il n'est pas crucial que les branches de T soient surjectives : dans le cas contraire, on peut imiter la preuve qui suit en remplaçant les fonctions lipschitziennes par les fonctions à variation bornée. Il n'est pas non plus crucial que T n'ait qu'un nombre fini de branches, du moment que T satisfait une condition de distorsion adéquate, comme dans l'exemple de l'application de Gauß $x \mapsto \{1/x\}$. En revanche, il est essentiel que la dérivée de T (ou d'un de ses itérés) soit partout strictement plus grande que 1 en valeur absolue : en présence d'un point fixe où la dérivée de T vaut 1, les orbites sont chaotiques loin du point fixe, mais dès qu'elles sont proches du point fixe elles y passent beaucoup de temps. Ce type de comportement est qualifié d'intermittent, et peut donner lieu à des vitesses de mélange arbitrairement lentes.

Démonstration du théorème 2.4. Soit $\hat{T}$ l'opérateur de transfert de T pour la mesure de Lebesgue (qui n'est en général pas invariante). Avec la formule du changement de variables, on vérifie que

$$\hat{T}u(x) = \sum_{T(y)=x} \frac{1}{|T'(y)|} u(y).$$

Plus généralement, $\hat{T}^n u(x) = \sum_{T^n(y)=x} u(y)/|(T^n)'(y)|$. Si on note ψ_j les branches inverses de T^n, on peut aussi écrire $\hat{T}^n u(x) = \sum \psi_j'(x) u(\psi_j x)$.

Contrairement au cas de l'application doublement, le jacobien n'est en général pas constant. Ainsi, $\operatorname{Lip}(\hat{T}^n u)$ ne va pas tendre vers 0 à cause de la contribution du jacobien dans la formule ci-dessus (ce n'est pas surprenant, puisque $\hat{T}^n u$ devrait tendre vers la densité de la mesure invariante, qui n'a pas de raison d'être constante). En revanche, l'influence des irrégularités initiales de u devient de moins en moins grande, puisque

$\operatorname{Lip}(u \circ \psi_j) \leqslant \lambda^{-n} \operatorname{Lip}(u)$, où $\lambda > 1$ est le coefficient minimal de dilatation de T. On obtient une inégalité du type

$$\left\| \hat{T}^n u \right\|_{\mathrm{Lip}} \leqslant C\lambda^{-n} \|u\|_{\mathrm{Lip}} + C \|u\|_{C^0} . \tag{2.2}$$

C'est un affaiblissement de l'inégalité (2.1), avec un terme d'erreur supplémentaire contrôlé par le sup de u. Notons toutefois que l'inclusion de Lip dans C^0 est compacte. Cela nous permettra d'exploiter l'inégalité précédente, moyennant un peu de théorie spectrale que nous allons maintenant exposer.

Intermède : Quelques éléments de théorie spectrale

Soient $(\mathcal{B}, \|\cdot\|)$ un espace de Banach sur $\mathbb{C}$ et $\mathcal{L}$ un opérateur linéaire continu sur $\mathcal{B}$. Le *spectre* $\sigma(\mathcal{L})$ de $\mathcal{L}$ est l'ensemble des nombres complexes z tels que $zI - \mathcal{L}$ ne soit pas inversible.

Si z est un point isolé du spectre de $\mathcal{L}$, on peut définir le projecteur spectral correspondant $\Pi_z \coloneqq \frac{1}{2i\pi} \int_C (wI - \mathcal{L})^{-1} \,\mathrm{d}w$, où C est un petit cercle autour de z. Cette définition est indépendante du choix de C, par holomorphie de $w \mapsto (wI - \mathcal{L})^{-1}$ hors de $\sigma(\mathcal{L})$. Lorsque $\mathcal{B}$ est de dimension finie, on vérifie (en mettant $\mathcal{L}$ sous forme triangulaire et en utilisant la formule de Cauchy) que Π_z est le projecteur sur le sous-espace caractéristique pour la valeur propre z, de noyau la somme directe des autres sous-espaces caractéristiques. En dimension infinie, un résultat analogue reste vrai : l'opérateur Π_z est une projection, son image et son noyau sont invariants par $\mathcal{L}$, et le spectre de la restriction de $\mathcal{L}$ à l'image est $\{z\}$, tandis que le spectre de la restriction de $\mathcal{L}$ au noyau est $\sigma(\mathcal{L}) - \{z\}$.

On dit que z est une valeur propre isolée de multiplicité finie de $\mathcal{L}$ si z est un point isolé de $\sigma(\mathcal{L})$, et si l'image de Π_z est de dimension finie. On note $\sigma_{\mathrm{ess}}(\mathcal{L})$ le spectre essentiel de $\mathcal{L}$, c'est-à-dire l'ensemble des points de $\sigma(\mathcal{L})$ qui ne sont pas des valeurs propres isolées de multiplicité finie.

On définit le rayon spectral $r(\mathcal{L})$ de $\mathcal{L}$ comme $\sup\{|z| \,:\, z \in \sigma(\mathcal{L})\}$, et le rayon spectral essentiel $r_{\mathrm{ess}}(\mathcal{L})$ comme $\sup\{|z| \,:\, z \in \sigma_{\mathrm{ess}}(\mathcal{L})\}$. Ces quantités peuvent se calculer comme suit : $r(\mathcal{L}) = \inf_{n\in\mathbb{N}} \|\mathcal{L}^n\|^{1/n}$, et

$$r_{\mathrm{ess}}(\mathcal{L}) = \inf \|\mathcal{L}^n - K\|^{1/n} , \tag{2.3}$$

où l'infimum est pris sur les entiers n et sur les opérateurs compacts K. En particulier, on obtient que le rayon spectral essentiel d'un opérateur compact est nul, ce qui correspond au fait bien connu que le spectre d'un opérateur compact est une suite de valeurs propres de multiplicité finie qui tend vers 0. La formule précédente montre également que le rayon spectral essentiel n'est pas modifié par l'ajout d'un opérateur compact : intuitivement, ajouter un opérateur compact revient simplement à ajouter (ou perturber) des valeurs propres de multiplicité finie.

Pour estimer le rayon spectral essentiel, la formule (2.3) n'est pas commode car elle suppose d'exhiber de bons opérateurs compacts. Il est plus pratique d'utiliser une technique reposant sur des inégalités de DOEBLIN-FORTET, aussi nommées inégalités de Lasota-Yorke, et que nous appellerons inégalités DFLY dans la suite. Cette technique est formalisée dans le lemme suivant essentiellement dû à [Hennion 1993] (la formulation qui suit se trouve dans [Bardet, Gouëzel et Keller 2007]).

Lemme 2.6. *Supposons que, pour un certain $n > 0$ et pour tout $x \in \mathcal{B}$, on ait*

$$\|\mathcal{L}^n x\| \leqslant M^n \|x\| + \|x\|_w \,, \tag{2.4}$$

où $\|\cdot\|_w$ est une seminorme sur $\mathcal{B}$ telle que la boule unité de $\mathcal{B}$ (pour $\|\cdot\|$) soit relativement compacte pour $\|\cdot\|_w$. Alors $r_{\mathrm{ess}}(\mathcal{L}) \leqslant M$.

Si l'on pouvait décomposer $\mathcal{L}$ en somme de deux opérateurs $\mathcal{L}_1 + \mathcal{L}_2$ bornés respectivement par le premier et le deuxième terme de l'inégalité (2.4), l'opérateur $\mathcal{L}_1$ aurait un rayon spectral (et donc un rayon spectral essentiel) borné par M, tandis que $\mathcal{L}_2$ serait compact. Il suivrait donc de (2.3) que le rayon spectral essentiel de $\mathcal{L}$ serait au plus M, ce qui est le résultat annoncé par le lemme. Cet argument heuristique suffit à motiver l'énoncé du lemme, mais sa preuve rigoureuse est totalement différente.

Fin de la démonstration du théorème 2.4. L'inégalité (2.2) est une inégalité DFLY puisque l'inclusion des fonctions lipschitziennes dans les fonctions continues est compacte par le théorème d'Ascoli. Ainsi, le rayon spectral essentiel de l'opérateur de transfert $\hat{T}$ sur Lip est $\leqslant \lambda^{-1} < 1$. En particulier, $\hat{T}$ a un nombre fini de valeurs propres de module $\geqslant 1$. En utilisant la régularité des fonctions propres correspondantes, on vérifie qu'il y a en fait une unique valeur propre en 1, simple. Si Π_1 désigne le projecteur spectral correspondant, on montre qu'il s'écrit sous la forme $\Pi_1(u) = (\int uh \,\mathrm{dLeb})h$, où h est la densité d'une mesure de probabilité invariante μ. De plus, $\left\|\hat{T}^n - \Pi_1\right\|$ décroît exponentiellement vite (la vitesse est donnée par le module maximal d'une valeur propre différente de 1, ou par λ^{-1} en l'absence d'une telle valeur propre ; en particulier, cette vitesse n'est pas explicite). Cela implique la décroissance exponentielle des corrélations recherchée. □

3. Résonances de Ruelle

Le preuve du théorème 2.4 donne plus que le simple mélange exponentiel : comme le rayon spectral essentiel de l'opérateur de transfert $\hat{T}$ sur Lip est $\leqslant \lambda^{-1}$, l'opérateur $\hat{T}^n$ se comporte pour tout $\rho > \lambda^{-1}$ comme les puissances d'une matrice, en dehors d'un terme qui décroît au moins comme ρ^n. De manière équivalente, l'opérateur de transfert associé à la mesure invariante μ (qui est conjugué au précédent) vérifie la même propriété. Ainsi, il existe des nombres complexes $(\lambda_j)_{1 \leqslant j \leqslant M(\rho)}$ (qu'on peut supposer ordonnés par module décroissant), des entiers $m_j \geqslant 0$ (la taille du bloc de Jordan éventuel pour la valeur propre λ_j) et des formes bilinéaires $F_{j,m}$ sur $\mathrm{Lip} \times L^1$ pour $0 \leqslant m \leqslant m_j$ tels que, pour toutes fonctions $u \in \mathrm{Lip}$ et $v \in L^1$,

$$\int u \cdot v \circ T^n \, \mathrm{d}\mu = \sum_{j=1}^{M(\rho)} \sum_{m=0}^{m_j} F_{j,m}(u,v) n^m \lambda_j^n + O(\rho^n).$$

Le terme d'erreur $O(\rho^n)$ est uniforme en u et v, i.e., il est de la forme $C\rho^n \|u\|_{\mathrm{Lip}} \|v\|_{L^1}$. Notons d'ailleurs que l'uniformité (qui découle ici de la preuve) est automatique dans tous les cas analogues par un argument de type Banach-Steinhaus [Chazottes, Collet et Schmitt 2005, Theorem B.1].

On dira dans ce cas que les corrélations admettent une *description matricielle* jusqu'au terme d'erreur ρ^n. Les coefficients $F_{j,m}(u,v)$ qui apparaissent dans l'asymptotique sont parfois appelés *distributions de Ruelle* (dans ce cadre, ce ne sont de vraies distributions que sur la fonction u, puisque les formes linéaires sur L^1 sont simplement données par intégration contre des fonctions bornées), et les λ_j les *résonances de Ruelle*.

On peut être plus précis en ce qui concerne la description des corrélations :

Proposition 3.1. *Si la transformation T est dilatante par morceaux à branches C^∞ surjectives, alors les corrélations de fonctions C^∞ admettent une description matricielle jusqu'à un terme d'erreur ε^n, pour tout $\varepsilon > 0$.*

En effet, en reproduisant la démonstration du théorème 2.4 mais en remplaçant l'espace Lip par un espace de fonctions C^k, on démontre une inégalité DFLY de la forme

$$\left\| \hat{T}^n u \right\|_{C^k} \leqslant C\lambda^{-kn} \|u\|_{C^k} + C \|u\|_{C^{k-1}}.$$

L'injection de C^k dans C^{k-1} étant compacte, on peut à nouveau faire appel au lemme 2.6 de Hennion pour obtenir que le rayon spectral essentiel de $\hat{T}$ sur C^k est borné par λ^{-k}. La description matricielle des corrélations avec un terme d'erreur λ^{-kn} suit.

En général, les distributions de Ruelle correspondant à des résonances de plus en plus proches de 0 ont des ordres de plus en plus élevés. En revanche, lorsqu'on augmente k, il n'apparaît pas de nouvelles résonances dans les zones qui étaient déjà hors du spectre essentiel de $\hat{T}$ sur les espaces de différentiabilité plus faible (cela suit directement du fait que les résonances doivent apparaître dans l'asymptotique matricielle des corrélations de fonctions C^∞, et ne sont donc pas liées à l'espace qu'on choisit de considérer). Ce phénomène est beaucoup plus général : en un sens, les valeurs propres isolées de multiplicité finie ne dépendent que de l'opérateur, et pas de l'espace sur lequel il agit. Ce fait est rendu précis par l'énoncé suivant [Baladi et Tsujii 2008, Lemma A.1] :

Proposition 3.2. *Considérons deux opérateurs linéaires continus $M_1 : F_1 \to F_1$ et $M_2 : F_2 \to F_2$ sur deux espaces de Banach. Supposons qu'il existe un espaces vectoriel E et un espace vectoriel topologique G, des injections d'image dense $i_k : E \to F_k$ et des injections continues $j_k : F_k \to G$ (pour $k \in \{1,2\}$) telles que, pour tout $n \in \mathbb{N}$, $j_1 \circ M_1^n \circ i_1 = j_2 \circ M_2^n \circ i_2$. Notons A l'ensemble des nombres complexes qui sont valeurs propres isolées de multiplicité finie pour M_1 et M_2 (éventuellement de multiplicité nulle, i.e., on inclut les points qui ne sont pas dans le spectre de l'opérateur), et A_∞ la composante connexe de l'infini dans A. Alors, pour tout $z \in A_\infty$, les multiplicités des* *espaces caractéristiques associés à z pour M_1 et M_2 coïncident, et les actions de M_1 et M_2 sur ces deux espaces sont conjuguées.*

Dans les applications, il faut penser à F_1 et F_2 comme deux espaces de fonctions ou de distributions, tandis que E est souvent l'ensemble des fonctions C^∞ et G est l'ensemble de toutes les distributions. L'énoncé de la proposition est une manière de donner un sens précis au fait de considérer un même opérateur sur des espaces différents.

La preuve de la proposition 3.2 est très simple une fois que l'on dispose du bon énoncé :

Démonstration. Dans A_∞, les points qui appartiennent au spectre de M_1 ou de M_2 forment un ensemble discret. Notons A'_∞ leur complémentaire. Sur celui-ci, les fonctions $z \mapsto j_k \circ (z - M_k)^{-1} \circ i_k$ (pour $k \in \{1, 2\}$) sont holomorphes. Pour $|z|$ grand, leurs développements en série de Taylor tronqués à un ordre arbitraire N coïncident par hypothèse. Par continuité de j_1 et j_2, en passant à la limite en N, on obtient que ces fonctions sont égales. Par holomorphie, elles coïncident même sur A'_∞. L'expression intégrale des projecteurs spectraux $\Pi_1(z)$ et $\Pi_2(z)$ implique alors que, pour tout $z \in A_\infty$,

$$j_1 \circ \Pi_1(z) \circ i_1 = j_2 \circ \Pi_2(z) \circ i_2. \tag{3.1}$$

Comme l'image de i_1 est dense dans F_1, l'image de $\Pi_1(z) \circ i_1$ est dense dans l'image de $\Pi_1(z)$, qui est de dimension finie. Par conséquent, ces deux espaces coïncident. On en déduit que la dimension de l'image de $\Pi_1(z)$ est la même que celle de l'image de $j_1 \circ \Pi_1(z) \circ i_1$. Il suit donc de (3.1) que les images de $\Pi_1(z)$ et $\Pi_2(z)$ ont même dimension.

Finalement, $j_2^{-1} \circ j_1$ est une bijection entre les espaces de dimension finie $\operatorname{Im} \Pi_1(z)$ et $\operatorname{Im} \Pi_2(z)$, qui y conjugue les actions de M_1 et M_2. □

4. La méthode spectrale de Nagaev-Guivarc'h

Les techniques de martingales exposées dans la section 1 sont très efficaces pour démontrer le théorème central limite (ou des raffinements de celui-ci). Cependant, certains résultats ne sont pas à leur portée (mentionnons par exemple la convergence vers des lois stables, ou le théorème limite local). Dans ce paragraphe, nous exposons une autre méthode, la méthode spectrale de Nagaev-Guivarc'h, qui requiert des hypothèses plus fortes de trou spectral, mais qui permet d'étudier une gamme de problèmes plus étendue. L'idée centrale de cette méthode est d'étendre les méthodes de fonctions caractéristiques, très puissantes pour des variables indépendantes, au cas dépendant.

Pour illustrer cette méthode, nous allons redémontrer le théorème central limite pour les applications de l'intervalle :

Théorème 4.1. *Soit $T : [0, 1] \to [0, 1]$ une application dilatante par morceaux C^2 à branches surjectives. Soit $f : [0, 1] \to \mathbb{R}$ une fonction lipschitzienne de moyenne nulle pour la mesure invariante absolument continue μ. Les sommes de Birkhoff $S_n f$ vérifient un théorème central limite : $S_n f/\sqrt{n}$ converge vers une gaussienne $\mathcal{N}(0, \sigma^2)$.*

Nous avons déjà démontré ce théorème (il suit des théorèmes 2.4 et 1.1), c'est plutôt la technique de preuve qui est intéressante.

DÉMONSTRATION. Notons $\hat{T}$ l'opérateur de transfert associé à T et à la mesure de Lebesgue, et définissons un opérateur perturbé $\hat{T}_t(u) = \hat{T}(e^{\mathbf{i}tf} u)$. Il vérifie $\int u e^{\mathbf{i}tS_n f} v \circ T^n = \int \hat{T}_t^n(u) v$. En particulier, si on note h la densité de μ, alors la fonction caractéristique de $S_n f$ s'écrit

$$E(e^{\mathbf{i}tS_n f}) = \int \hat{T}_t^n(h).$$

Pour t petit, l'opérateur $\hat{T}_t$ est une petite perturbation de l'opérateur $\hat{T}$, qui a une valeur propre simple en 1 et un trou spectral. Des théorèmes standard de perturbation (voir [KATO 1966]) assurent que $\hat{T}_t$ a encore une unique valeur propre simple $\lambda(t)$ proche de 1, et un trou spectral. Comme $t \mapsto \hat{T}_t$ est analytique, la fonction $t \mapsto \lambda(t)$ est même analytique. En particulier, il existe $\rho < 1$ et une fonction analytique $c(t)$ avec $c(0) = 1$ tels que, pour tout t assez petit et pour tout $n \in \mathbb{N}$,

$$E(e^{\mathbf{i}tS_n f}) = c(t)\lambda(t)^n + O(\rho^n).$$

C'est un analogue d'une estimée triviale (mais cruciale) sur les fonctions caractéristiques de variables aléatoires indépendantes : si $Z_0, Z_1, \dots$ sont indépendantes et identiquement distribuées, et S_n désigne leur somme partielle au rang n,

$$E(e^{\mathbf{i}tS_n}) = E(e^{\mathbf{i}tZ_0})^n.$$

Cette équation est à la source de nombreuses preuves concernant S_n, qu'on va pouvoir imiter dans le cas dépendant si l'on peut comprendre assez précisément $\lambda(t)$ pour t petit.

Dans notre cas, comme λ est analytique, elle admet un développement $\lambda(t) = 1 + at + bt^2 + o(t^2)$. On vérifie que $a = \int f \,\mathrm{d}\mu = 0$. Par conséquent, pour tout $t \in \mathbb{R}$,

$$E(e^{\mathbf{i}tS_n f/\sqrt{n}}) = c(t/\sqrt{n})\lambda(t/\sqrt{n})^n + o(1) \to e^{bt^2}.$$

Par le théorème de LÉVY, cela assure d'une part que $t \mapsto e^{bt^2}$ est la fonction caractéristique d'une certaine variable aléatoire (et donc que $b \in]-\infty, 0]$), et d'autre part que $S_n f/\sqrt{n}$ converge en distribution vers cette variable aléatoire (gaussienne). □

Mentionnons maintenant une autre application de la méthode spectrale, qui ne peut pas être démontrée par des méthodes de martingales : le théorème de la limite locale.

THÉORÈME 4.2. *Sous les hypothèses du théorème précédent, supposons de plus que f est* apériodique*, i.e., qu'il n'existe pas de fonction mesurable u telle que $f + u - u \circ T$ prenne presque partout ses valeurs dans un réseau translaté $\alpha + \beta\mathbb{Z}$. En ce cas, pour tout intervalle borné $J \subset \mathbb{R}$,*

$$\mu\{x \in [0,1] \,:\, S_n f(x) \in J\} \sim \frac{|J|}{\sigma\sqrt{2\pi n}}. \tag{4.1}$$

Tandis que le théorème central limite affirme que $S_n f$ ressemble à une gaussienne à une échelle $\sqrt{n}$, le théorème local donne le même résultat mais à une échelle $O(1)$, puisque le terme de droite dans (4.1) est essentiellement la probabilité qu'une gaussienne $\mathcal{N}(0, \sigma^2)$ appartienne à l'intervalle $J/\sqrt{n}$. Des perturbations de $S_n f$ à une échelle $O(1)$ ou des phénomènes fins de périodicité suffisent à mettre ce résultat en échec, ce qui explique qu'on ait besoin d'hypothèses fortes et de techniques sophistiquées pour le démontrer (même dans le cas de variables i.i.d. de carré intégrable, ce résultat a été seulement démontré dans [STONE 1965]).

DÉMONSTRATION. Il s'agit de démontrer que $E(\varphi(S_n f)) \sim (\int \varphi)/(\sigma\sqrt{2\pi n})$ pour toute fonction φ indicatrice d'un intervalle borné. Par un argument d'approximation (non trivial !), il suffit de le démontrer pour une autre classe de fonctions φ suffisamment large, les fonctions lisses dont la transformée de Fourier est à support compact. Si φ est une telle fonction, on a $\varphi(\zeta) = \int_{-M}^{M} e^{\mathbf{i}t\zeta}\hat{\varphi}(t)\,\mathrm{d}t$, donc

$$E(\varphi(S_n f)) = \int_{-M}^{M} E(e^{\mathbf{i}tS_n f})\hat{\varphi}(t)\,\mathrm{d}t. \tag{4.2}$$

On peut alors utiliser la description de $E(e^{\mathbf{i}tS_n f})$ provenant de la méthode spectrale, au moins pour t petit. Pour t plus grand, l'apériodicité de f implique que le rayon spectral de $\hat{T}_t$ est strictement plus petit que 1, et donc que $E(e^{\mathbf{i}tS_n f})$ décroît exponentiellement vite en n. Le coefficient exponentiel étant uniforme pour t dans un intervalle compact loin de 0, on en déduit que la contribution de ces t à (4.2) est exponentiellement petite, donc sans importance.

On conclut finalement en utilisant pour t petit l'égalité $E(e^{\mathbf{i}tS_n f}) = c(t)\lambda(t)^n + O(\rho^n)$, le développement asymptotique de $\lambda(t)$, et le changement de variables $t' = \sqrt{n}t$. □

5. Le billard stade

Dans ce paragraphe, nous décrivons un exemple de système non-uniformément hyperbolique, pour lequel un certain nombre de résultats des paragraphes précédents sont mis en défaut. Il s'agit du billard stade : considérons un convexe dont les bords sont deux demi-cercles et deux segments parallèles, partons d'un vecteur sur le bord de ce convexe pointant vers son intérieur, et considérons le nouveau vecteur obtenu après la première réflexion sur le bord du convexe. On obtient ainsi une transformation T sur l'espace $X = S^1 \times \left]-\pi/2, \pi/2\right[$ (le facteur $r \in S^1$ correspond à la position du point sur le bord du stade, tandis que le facteur $\theta \in \left]-\pi/2, \pi/2\right[$ correspond à l'angle fait par le vecteur avec la normale intérieure), qui préserve la mesure $\mathrm{d}\mu = \cos\theta\,\mathrm{d}r\,\mathrm{d}\theta$ (cette formule explicite est due à l'origine hamiltonienne du système).

Si l'on considère le même type de transformation mais dans le tore et avec des obstacles strictement convexes, on obtient un système appelé *billard de* SINAI qui est maintenant bien compris (voir par exemple [CHERNOV et MARKARIAN 2006]) : il est hyperbolique (on peut écrire explicitement des familles de cônes complémentaires strictement contractées et strictement dilatées par la différentielle de l'application), ergodique, et mélange exponentiellement vite [YOUNG 1998]. Techniquement, la présence de discontinuités (où la dérivée est en plus non bornée) complique fortement l'étude, mais l'hyperbolicité finit par l'emporter sur les phénomènes de découpage, ce qui est au cœur des démonstrations des résultats précédents. Du mélange exponentiel, on peut déduire le théorème central limite pour les fonctions hölderiennes grâce au théorème 1.2.

Dans le cas du stade, les familles de cônes évoquées précédemment sont encore contractées et dilatées, mais d'un coefficient dépendant du point. En particulier, deux types d'orbite ont une hyperbolicité très faible : ce sont d'une part les orbites qui rebondissent presque perpendiculairement entre les segments du bord, et d'autre part les orbites presque tangentes aux demi-cercles. Plus précisément, notons B_n l'ensemble des points qui sont sur un des demi-cercles, dont les n prochains itérés rencontrent presque orthogonalement les segments, et qui arrivent ensuite sur l'autre demi-cercle. Ces orbites présentent une hyperbolicité très faible jusqu'au temps n, et $\mu(B_n) \sim c/n^3$ (comme on le vérifie aisément géométriquement). Les orbites rasant le cercle pendant une durée n ayant une mesure $\sim c'/n^4$, nous pourrons négliger ce second ensemble de trajectoires.

On peut utiliser ces informations pour estimer heuristiquement les corrélations. Notons $A_n = \bigcup_{k>2n} \bigcup_{j<k-n} T^j B_k$. Cet ensemble est constitué de points qui restent sur les segments pendant un temps au moins n, et sa mesure est $\asymp 1/n$. Soient f_n la fonction caractéristique de A_n et g une fonction supportée par les demi-cercles d'intégrale non nulle, on obtient $f_n \cdot g \circ T^n = 0$. En particulier,

$$\left| \int f_n \cdot g \circ T^n - \int f_n \int g \right| = \left| \int f_n \int g \right| = C\mu(A_n) \asymp 1/n.$$

Si f_n était une suite de fonctions hölderiennes de norme uniformément bornée, on en déduirait que les corrélations de fonctions hölderiennes ne peuvent pas décroître plus rapidement que $1/n$. Ce n'est pas le cas, mais ce calcul permet cependant de s'attendre à une vitesse de mélange en $1/n$.

[Chernov et Zhang 2008] montre que les corrélations décroissent au plus en $O(1/n)$. Même si la borne inférieure n'est pas démontrée rigoureusement, c'est assurément le bon ordre de grandeur. En particulier, les corrélations ne sont pas sommables, et on ne s'attend pas à avoir un théorème central limite classique. Effectivement, dans [Bálint et Gouëzel 2006], nous avons démontré qu'on a dans cette situation un théorème limite non standard : si f est une fonction hölderienne de moyenne nulle, alors

$$\frac{S_n f}{\sqrt{n \log n}} \to I \sqrt{\frac{4 + 3\log 3}{4 - 3\log 3} \cdot \frac{\ell^2}{4(\pi + \ell)}} \mathcal{N}(0,1), \tag{5.1}$$

où ℓ est la longueur des segments divisée par le rayon des demi-cercles, et I est la moyenne de f sur les trajectoires rebondissant perpendiculairement d'un segment à l'autre. Lorsqu'une trajectoire est coincée pendant un temps k entre les deux segments en y rebondissant presque perpendiculairement, elle va contribuer à un accroissement des sommes de Birkhoff de l'ordre de kI, qui ne sera pas compensé par les autres valeurs négatives de f. L'essence de la convergence en distribution ci-dessus est que ce phénomène est suffisamment fréquent pour forcer une croissance de $S_n f$ plus rapide que le $\sqrt{n}$ naïvement attendu, en $\sqrt{n \log n}$.

Le facteur $(4 + 3\log 3)/(4 - 3\log 3)$ dans (5.1) mérite un commentaire. Si une trajectoire fait k_0 rebonds sur les segments, puis arrive ensuite sur les demi-cercles, elle reviendra rapidement faire à nouveau un nombre important k_1 de rebonds sur les segments, puis k_2, etc. Dans l'asymptotique k_0 grand, la distribution de k_{i+1} n'est pas indépendante de ce qui précède, mais elle ne dépend essentiellement que de k_i. Ainsi, le nombre de sauts successifs sur les segments est dirigé par une chaîne de Markov explicite. Le facteur numérique $(4+3\log 3)/(4-3\log 3)$ provient de la mesure stationnaire de cette chaîne de Markov.

Notons une conséquence simple de (5.1) sur les corrélations : si f est une fonction hölderienne de moyenne nulle avec $I = I(f) \neq 0$, alors $\int f \cdot f \circ T^n$ n'est pas $o(1/n)$ (la majoration $O(1/n)$ de [Chernov et Zhang 2008] est donc optimale au moins le long d'une sous-suite). En effet,

$$\int \left(\sum_{k=0}^{n-1} f \circ T^k \right)^2 = n \int f^2 + 2 \sum_{k=1}^{n-1} (n-k) \int f \cdot f \circ T^k .$$

Si on avait $\int f \cdot f \circ T^n = o(1/n)$, on en déduirait $\int (S_n f)^2 = o(n \log n)$, et $S_n f / \sqrt{n \log n}$ convergerait en probabilité vers 0. C'est absurde au vu de (5.1).

Donnons rapidement une idée de la preuve de (5.1). L'idée principale est d'induire pour se ramener à un système très hyperbolique, de démontrer un théorème limite pour l'application induite, et de remonter ensuite à l'application initiale. Plus précisément, soit $Y \subset X$ un sous-ensemble ne contenant pas de point qui va rebondir plusieurs fois sur les segments ou sur les demi-cercles. Notons $\varphi_Y(x) = \inf\{n \geqslant 1 \,:\, T^n(x) \in Y\}$ le temps de retour en Y (il est défini presque partout sur Y par le théorème de récurrence de Poincaré), et $T_Y = T^{\varphi_Y} : Y \to Y$ l'application induite. Elle est suffisamment hyperbolique pour que la théorie de Young s'applique (en particulier, T_Y mélange exponentiellement vite). Soit $f : X \to \mathbb{R}$ une fonction de moyenne nulle, définissons $f_Y : Y \to \mathbb{R}$ par $f_Y = \sum_0^{\varphi_Y - 1} f \circ T^k$. Par construction, les sommes de Birkhoff de f_Y pour T_Y (notées $S_n^Y f_Y$) forment une sous-suite des sommes de Birkhoff de f pour T. De plus, la distribution de f_Y peut être bien comprise : cette fonction n'est pas dans L^2, mais presque. On peut alors utiliser la méthode spectrale de Nagaev-Guivarc'h pour montrer que $S_n^Y f_Y / \sqrt{n \log n}$ converge en distribution vers une limite gaussienne [Aaronson et Denker 2001a]. Un argument élémentaire permet enfin de remonter à l'application initiale [Melbourne et Török 2004].

CHAPITRE II

Caractérisation de la convergence des sommes de Birkhoff pour les applications Gibbs-Markov

Sommaire

1. Théorèmes limites pour des sommes de variables aléatoires i.i.d. 21
2. Théorèmes limites pour les applications Gibbs-Markov 22
3. Stratégie de preuve 24
4. Utilisation des développements caractéristiques précis 26
5. Développement caractéristique pour les applications Gibbs-Markov 28

1. Théorèmes limites pour des sommes de variables aléatoires i.i.d.

Il est possible de décrire exactement dans quels cas une somme de variables aléatoires indépendantes identiquement distribuées, convenablement renormalisée, converge en distribution vers une loi non triviale (i.e., non concentrée en un point). L'exemple le plus simple de ce type de convergence est le cas L^2 : si $Z_0, Z_1, \dots$ sont indépendantes, identiquement distribuées, de carré intégrable (et non constantes !), alors $(Z_0 + \dots + Z_{n-1} - nE(Z_0))/\sqrt{n}$ converge en distribution vers une gaussienne $\mathcal{N}(0, \sigma^2)$, où $\sigma^2 > 0$ est la variance commune des Z_i.

Ce type de convergence peut également s'obtenir en dehors du cas L^2 : supposons que, lorsque $x \to +\infty$,

$$P(Z_i > x) = (c_1 + o(1))\ell(x)x^{-\beta}, \quad P(Z_i < -x) = (c_2 + o(1))\ell(x)x^{-\beta}, \tag{1.1}$$

avec $\beta \in (0, 2)$, ℓ une fonction à variation lente (i.e., $\ell(\lambda x)/\ell(x)$ converge vers 1 lorsque x tend vers l'infini, pour tout $\lambda > 0$), et $c_1 + c_2 > 0$. Alors $(Z_0 + \dots + Z_{n-1} - A_n)/B_n$ converge en distribution vers une limite non triviale, où A_n est une suite convenable (on peut prendre $A_n = 0$ pour $\beta < 1$ et $A_n = nE(Z_i)$ pour $\beta > 1$, le cas $\beta = 1$ étant plus délicat) et B_n satisfait $n\ell(B_n) \sim B_n^\beta$ (B_n est donc de l'ordre de $n^{1/\beta}$, à un terme à variation lente près). Cet énoncé, dû à P. Lévy, est facile : il suffit de travailler avec les fonctions caractéristiques des Z_i, qu'on peut estimer en utilisant l'information (1.1) sur les queues. Les distributions limites qu'on obtient dans ce cadre sont appelées lois stables (d'indice β).

Il est remarquable qu'on puisse établir une réciproque à ce résultat : si une suite de variables aléatoires indépendantes identiquement distribuées $Z_0, Z_1, \dots$ est telle que

$(Z_0 + \cdots + Z_{n-1} - A_n)/B_n$ converge en distribution vers une variable aléatoire non triviale, pour certaines suites $A_n, B_n \in \mathbb{R}$, alors les Z_i appartiennent à L^2 ou elles vérifient (1.1) pour certains c_1, c_2, ℓ, β (on peut aussi avoir $\beta = 2$, auquel cas il faut modifier légèrement (1.1)). La preuve de cette réciproque, également due à Lévy, est nettement plus délicate que le sens direct, et est de nature quasiment algébrique : il s'agit d'associer à la distribution limite une *mesure de Lévy* qui vérifie automatiquement des propriétés d'autosimilarité, et qui est donc nécessairement une fonction puissance. La variation régulière des queues de Z_i suit ensuite du théorème de Karamata [Bingham, Goldie et Teugels 1987, Theorem 1.6.1].

Il est naturel de se demander si des résultats de rigidité aussi forts existent pour les sommes de Birkhoff $S_n f$ dans certaines classes de systèmes dynamiques. Si l'on n'impose aucune hypothèse de régularité sur f, le comportement asymptotique de $S_n f$ peut être essentiellement arbitraire (voir [Volný 1990]). En revanche, si les conditions de régularité sont trop fortes, on obtiendra toujours le théorème central limite (c'est par exemple le cas si l'on considère f lipschitzienne, et T une application de l'intervalle uniformément dilatante, voir le paragraphe I.2). L'objet de ce chapitre est d'étudier ce qui se passe pour une classe de systèmes fortement hyperboliques (les *applications Gibbs-Markov*), mais sous des hypothèses de régularité extrêmement faibles sur f (et sans aucune hypothèse d'intégrabilité). On verra qu'on peut caractériser exactement les fonctions f pour lesquelles un théorème limite a lieu (théorème 2.2). En particulier, les lois limites sont toujours des gaussiennes ou des lois stables : il n'y a pas de théorème limite exotique dans ce cadre. Ce travail a été publié dans [Gouëzel 2010b].

2. Théorèmes limites pour les applications Gibbs-Markov

Les transformations Gibbs-Markov sont des applications markoviennes sur une infinité de symboles, partageant beaucoup de propriétés des sous-décalages de type fini.

Définition 2.1. *Soit T une transformation sur un espace métrique borné X, préservant une mesure de probabilité μ. On dit que c'est une application Gibbs-Markov par rapport à une partition dénombrable Π de X si elle satisfait les propriétés suivantes :*

- *pour tout $a \in \Pi$, l'ensemble $T(a)$ est une réunion d'éléments de Π, et $T : a \to T(a)$ est une bijection bimesurable ;*
- *la transformation T a la propriété de grande image : $\inf_{a\in\Pi} \mu(Ta) > 0$;*
- *la transformation T est localement uniformément dilatante : il existe $\lambda > 1$ tel que, pour tout $a \in \Pi$ et pour tous $x, y \in a$, on ait $d(Tx, Ty) \geqslant \lambda d(x, y)$;*
- *la distorsion de T est lipschitzienne : notant $J(x)$ l'inverse du jacobien de T en x, on a pour tout $a \in \Pi$ et pour tous $x, y \in a$*

$$\left|\frac{J(x)}{J(y)} - 1\right| \leqslant Cd(Tx, Ty).$$

On peut identifier un point avec la suite des éléments de Π qu'il visite, et donc identifier X avec un sous-ensemble de $\Pi^{\mathbb{N}}$ correspondant aux suites dont toutes les transitions sont admissibles. En particulier, on peut définir sur X une nouvelle distance symbolique $d_\kappa(x, y) = \kappa^{s(x,y)}$, où $\kappa < 1$ est fixé et $s(x, y)$ est le temps de séparation de x

et y, i.e., le premier instant où leurs itérés ne sont pas dans le même élément de partition. On vérifie aisément que, si $\kappa \geqslant \lambda^{-1}$, alors $d(x,y) \leqslant C d_\kappa(x,y)$. Ainsi, T est encore Gibbs-Markov pour la distance d_κ, et celle-ci est dilatée d'un facteur exactement κ^{-1} à chaque itération. Notons qu'une fonction α-hölderienne pour la distance d_κ est lipschitzienne pour la distance d_{κ^α}, il est donc équivalent de traiter de fonctions lipschitziennes ou de fonctions hölderiennes.

Nous cherchons à voir dans quelles conditions les sommes de Birkhoff d'une fonction mesurable f satisfont un théorème limite. Il ne serait pas satisfaisant de mettre *a priori* des hypothèses d'intégrabilité sur f. En revanche, nous aurons besoin d'hypothèses de régularité sur f. Celles-ci ne peuvent pas être globales, car elles forceraient f à être bornée. On pourrait demander à f d'être constante sur chaque élément de Π, et ce serait déjà un résultat non trivial de caractériser les théorèmes limites apparaissant dans ce cadre. Cependant, nous allons imposer une hypothèse nettement plus faible. Pour $a \in \Pi$, notons $Df(a)$ la meilleure constante de Lipschitz de f sur a.

Théorème 2.2. *Soit $T : X \to X$ une application Gibbs-Markov mélangeante. Soit $f : X \to \mathbb{R}$ satisfaisant $\sum_{a\in\Pi} \mu(a) Df(a)^\eta < \infty$, pour un certain $\eta \in]0,1]$.*

Si $f \in L^2$, alors

- *Soit f est la somme d'un cobord mesurable et d'une constante, i.e., il existe une fonction mesurable u et une constante c telles que $f = u - u \circ T + c$ presque partout. Alors u est bornée, et $S_n f - nc$ converge en distribution vers la différence $Z - Z'$, où Z et Z' sont deux variables aléatoires indépendantes distribuées comme u.*
- *Sinon, posons $\sigma^2 = \int \tilde{f}^2 + 2\sum_{k=1}^\infty \int \tilde{f} \cdot \tilde{f} \circ T^k$ où $\tilde{f} = f - \int f \,\mathrm{d}\mu$. Cette série converge absolument, $\sigma^2 > 0$, et $(S_n f - n\int f)/\sqrt{n}$ converge en distribution vers $\mathcal{N}(0,\sigma^2)$.*

Si $f \notin L^2$, considérons une suite $Z_0, Z_1, \dots$ de variables aléatoires indépendantes distribuées comme f. Considérons aussi des suites $A_n \in \mathbb{R}$ et $B_n > 0$, et une variable aléatoire non dégénérée W. Alors $(S_n f - A_n)/B_n$ converge en distribution vers W si et seulement si $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$ converge également vers W.

En particulier, en dehors du cas L^2 où tout est compris (et où la dépendance joue vraiment un rôle, comme le montre la formule pour σ^2), tout se passe comme si on additionnait des variables aléatoires indépendantes identiquement distribuées, si bien que les résultats évoqués dans le paragraphe 1 permettent de caractériser exactement les cas où l'on a convergence. En particulier, on n'a pas besoin de refaire la preuve du théorème de Lévy pour obtenir cette caractérisation, ce qui est heureux car les arguments algébriques basés sur les mesures de Lévy semblent indissociables de l'indépendance et semblent donc extrêmement délicats à faire fonctionner dans un cadre dépendant comme celui des applications Gibbs-Markov.

Notons que le cas L^2 n'est pas difficile, il se traite aisément avec des méthodes de martingales comme le théorème I.1.1. Dans le cas $f \notin L^2$, on a deux directions à démontrer. Supposons tout d'abord que $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$ converge en loi vers une distribution W. En ce cas, d'après le paragraphe 1, on peut décrire explicitement les queues des Z_i, ainsi que les suites A_n et B_n. En particulier, on connaît exactement la classe d'intégrabilité de f. Ces informations sont suffisantes pour faire fonctionner la méthode spectrale de Nagaev-Guivarc'h (exposée dans le paragraphe I.4), et obtenir la convergence de $(S_n f - A_n)/B_n$ vers la même distribution W ; ces résultats sont démontrés dans [Aaronson et Denker 2001a,b] (sous une hypothèse de régularité plus forte $\sup_{a\in\Pi} Df(a) < \infty$). L'autre implication est nettement plus délicate, car on ne dispose d'aucune estimée *a priori* sur f, ni d'aucune information sur les suites A_n et B_n. On va décrire une preuve qui permet en fait de prouver les deux implications simultanément.

3. Stratégie de preuve

La démonstration repose sur la méthode spectrale de Nagaev-Guivarc'h. Autrement dit, si on note $\hat{T}$ l'opérateur de transfert associé à T (il est donné par $\hat{T}u(x) = \sum_{Ty=x} J(y)u(y)$), et $\hat{T}_t$ sa perturbation donnée par $\hat{T}_t(u) = \hat{T}(e^{\mathrm{i}tf}u)$, il s'agit de montrer que, pour t petit, l'opérateur $\hat{T}_t$ a une unique valeur propre proche de 1, et d'obtenir un développement asymptotique assez précis de celle-ci en 0 pour pouvoir en déduire un théorème limite.

La première difficulté dans cette direction est que la condition $\sum \mu(a)Df(a)^\eta < \infty$ n'est pas invariante par $\hat{T}$: on peut même avoir $D(\hat{T}f)(a) = +\infty$ pour un certain $a \in \Pi$.

Exemple 1. Soit $\kappa < 1$ fixé. Considérons le décalage plein markovien sur une infinité de symboles $a_0, a_1, \dots$, muni de la distance d_κ, avec $\mu(a_i) = Ce^{-i^2/2}$ pour un certain $C > 0$, et définissons une fonction f comme suit : f est nulle sur a_i, sauf sur les points dont les i itérés suivants sont dans a_0, auquel cas elle vaut e^{i^2}. En ce cas, $Df(a_i) = \kappa^{-i}e^{i^2}$, si bien que $\sum \mu(a_i)Df(a_i)^\eta < \infty$ pour tout $\eta < 1/2$. Soit $x_i = [a_0, \dots, a_0, a_1, \dots]$ où a_0 est répété i fois, il a une unique préimage y_i dans a_i et $f(y_i) = e^{i^2}$. Par conséquent, $\hat{T}f(x_i) \geqslant \mu(a_i)f(y_i)$, qui tend vers l'infini avec i. Cela montre que $\hat{T}f$ est non bornée sur a_0 ; en particulier, elle ne peut y être lipschitzienne.

Cependant, on doit plutôt considérer $\hat{T}_t(u)$, i.e., $\hat{T}(e^{\mathrm{i}tf}u)$ pour une fonction sympathique u, et on peut espérer que l'exponentielle complexe, qui exclut l'explosion en module, rende les choses plus contrôlées. Effectivement, $\hat{T}_t$ se comporte bien sur l'espace des fonctions hölderiennes d'exposant η (mais en général pas sur l'espace des fonctions lipschitziennes) : il a pour t petit une valeur propre $\lambda(t)$ proche de 1, qu'il s'agit d'estimer précisément. Si $f \in L^p$, une estimée naïve de $\lambda(t)$ ferait intervenir un terme d'erreur en $O(|t|^p)$. L'étape la plus délicate de la preuve est de voir qu'on peut améliorer cette estimée, avec un terme d'erreur $O(|t|^{p+\varepsilon})$ pour un certain $\varepsilon > 0$ uniforme.

Regardons maintenant les conséquences d'une telle estimée de $\lambda(t)$. La fonction caractéristique de $(S_n f - A_n)/B_n$ est asymptotiquement donnée par $\lambda(t/B_n)^n e^{-\mathrm{i}tA_n/B_n}$ (si B_n tend vers l'infini, ce qu'il faudra également démontrer). Pour pouvoir avancer, il faut savoir que les termes d'erreur dans $\lambda(t)$ (liés à l'intégrabilité de f) sont asymptotiquement négligeables, ce qui sera le cas si B_n croît approximativement comme $n^{1/p}$

(comme on s'y attend dans le cas indépendant). La difficulté est le manque d'information *a priori* sur B_n : si celui-ci croissait trop lentement, le terme d'erreur dans $\lambda(t)$ serait complètement incontrôlable ! La preuve va consister a obtenir des estimées *a priori* (reliées uniquement à l'intégrabilité de f) sur un certain ensemble de valeurs de t ayant 0 comme point d'accumulation, en déduire que B_n est à variation régulière (i.e., B_n croît essentiellement comme un polynôme multiplié par des fonctions à croissance lente) grâce à des propriétés de rigidité de ce type de suites, et obtenir *a posteriori* que le terme d'erreur est négligeable pour toutes les valeurs de t.

La preuve se divise donc en deux étapes complètement indépendantes : d'une part, obtenir un développement asymptotique assez précis de la fonction caractéristique de $S_n f$, et d'autre part en déduire des résultats sur les théorèmes limites. Pour formaliser ce raisonnement, introduisons la définition suivante.

DÉFINITION 3.1. *Soient $T : X \to X$ une transformation mélangeante préservant une mesure de probabilité, et $f : X \to \mathbb{R}$ une fonction mesurable. On dira que f admet un développement caractéristique pour T s'il existe deux fonctions c, λ définies sur un voisinage de 0 avec $c(0) = \lambda(0) = 1$ et une suite ε_n qui tend vers 0 telles que, pour tout t assez petit et pour tout $n \in \mathbb{N}$,*

$$|E(e^{\mathbf{i}tS_n f}) - c(t)\lambda(t)^n| \leqslant \varepsilon_n.$$

On dit que ce développement caractéristique est précis si l'une des conditions suivantes est satisfaite :

(1) *Soit il existe $q \leqslant 2$ et $\varepsilon > 0$ tels que $f \notin L^q$ et*

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O(|t|^{q+\varepsilon}) + O(t^2) + o\left(\int |e^{\mathbf{i}tf} - 1|^2\right) + O\left(\int |e^{\mathbf{i}tf} - 1|\right)^2.$$

(2) *Soit $f \in L^2$ et il existe $b \in \mathbb{C}$ tel que*

$$\lambda(t) = 1 + \mathbf{i}tE(f) - bt^2/2 + o(t^2).$$

Le théorème 2.2 est une conséquence des deux résultats suivants :

THÉORÈME 3.2. *Soient T une application Gibbs-Markov mélangeante, et $f : X \to \mathbb{R}$ une fonction satisfaisant $\sum_{a\in\Pi} \mu(a) Df(a)^\eta < \infty$, pour un certain $\eta > 0$. Alors f admet un développement caractéristique précis pour T.*

THÉORÈME 3.3. *Soient $T : X \to X$ une transformation mélangeante préservant une mesure de probabilité, et $f : X \to \mathbb{R}$ une fonction mesurable admettant un développement caractéristique précis.*

Si $f \in L^2$, alors $S_n f/\sqrt{n}$ converge en distribution vers une gaussienne.

Si $f \notin L^2$, considérons $Z_0, Z_1, \dots$ une suite de variables aléatoires indépendantes distribuées comme f, ainsi que $A_n \in \mathbb{R}$, $B_n > 0$ et W une variable aléatoire non dégénérée. Alors $(S_n f - A_n)/B_n$ converge en distribution vers W si et seulement si $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$ converge également vers W.

Pour obtenir le théorème 2.2 complet, il faut en plus discuter un peu le cas où la variance est nulle dans le théorème central limite, ce qui n'est pas difficile. Les preuves de ces deux résultats sont complètement indépendantes. Cependant, les énoncés eux-mêmes ne le sont pas réellement, car ils reposent tous deux sur la définition 3.1, qui est un peu arbitraire : on aurait pu autoriser d'autres termes d'erreur dans le développement asymptotique de $\lambda(t)$. Si on avait choisi des termes d'erreur plus restrictifs, le théorème 3.2 aurait été plus délicat à démontrer, mais la preuve du théorème 3.3 aurait inversement été plus aisée. La définition 3.1 donnée ci-dessus est exactement à la limite de ce que je sais faire : en général, je ne sais rien démontrer de plus précis sur $\lambda(t)$ que ce qui est énoncé dans cette définition, et cela suffit tout juste pour obtenir le théorème 3.3 (au prix d'une preuve assez détournée). Dans les deux sections qui suivent, je vais donner des idées de preuve de ces deux résultats, en commençant par le second.

4. Utilisation des développements caractéristiques précis

Dans cette section, on décrit la démonstration du théorème 3.3. Donnons tout d'abord la preuve (facile) dans le cas L^2. Dans ce cas, il suit de la définition d'un développement caractéristique précis que

$$e^{-\mathrm{i}t\sqrt{n}E(f)}\lambda\left(\frac{t}{\sqrt{n}}\right)^n \to e^{-(b-E(f)^2)t^2/2}.$$

Ainsi, $E(e^{\mathrm{i}t(S_nf-nE(f))/\sqrt{n}})$ converge vers $e^{-(b-E(f)^2)t^2/2}$. En particulier, cette dernière fonction est la fonction caractéristique d'une variable aléatoire W, et $(S_nf-nE(f))/\sqrt{n}$ converge en distribution vers celle-ci. Cela montre que $b-E(f)^2\in[0,+\infty[$, et que W est gaussienne de variance correspondante.

Dans le cas non L^2, on a *a priori* deux implications à démontrer. On va les établir simultanément, par l'intermédiaire du résultat suivant :

THÉORÈME 4.1. *Soient T et $\tilde{T}$ deux applications mélangeantes sur deux espaces probabilisés X et $\tilde{X}$, et $f : X \to \mathbb{R}$ et $\tilde{f} : \tilde{X} \to \mathbb{R}$ deux fonctions avec la même distribution n'appartenant pas à L^2. On suppose que (f,T) et $(\tilde{f},\tilde{T})$ admettent tous deux un développement caractéristique précis. Si $(\sum_{k=0}^{n-1} f\circ T^k - A_n)/B_n$ converge en distribution vers une variable aléatoire W non dégénérée, alors $(\sum_{k=0}^{n-1} \tilde{f}\circ\tilde{T}^k - A_n)/B_n$ converge également vers W.*

Montrons comment cet énoncé, plus symétrique, implique le théorème 3.3. Prenons pour $\tilde{T}$ le décalage à gauche sur $\mathbb{R}^{\mathbb{N}}$, muni de la mesure produit de telle sorte que les fonctions coordonnées soient indépendantes et distribuées comme f. Prenons pour $\tilde{f}$ la première coordonnée. Alors $(\tilde{f},\tilde{T})$ a un développement caractéristique précis (avec $\lambda(t)=E(e^{\mathrm{i}tf})$ et $c(t)=1$), par indépendance. Ainsi, si $(S_nf-A_n)/B_n$ converge en distribution vers W non dégénérée, le théorème 4.1 montre qu'il en va de même de $(S_n\tilde{f}-A_n)/B_n$, distribué comme $(\sum_{k=0}^{n-1} Z_k - A_n)/B_n$. Cela montre une des implications du théorème 3.3. L'autre implication est identique en échangeant les rôles de T et $\tilde{T}$.

Démontrons maintenant le théorème 4.1. La preuve est assez amusante, car il s'agit de partir d'hypothèses très faibles (convergence en distribution de $(S_nf-A_n)/B_n$ et développement asymptotique de λ) pour les renforcer progressivement. On montrera successivement que B_n tend vers l'infini, puis que B_{n+1}/B_n tend vers 1, puis que B_n est à variation régulière.

Posons $\Phi(t) = E(1 - \cos(tf)) \geqslant 0$, c'est cette fonction qui va dicter tous les comportements asymptotiques. Tout d'abord, on vérifie aisément de la définition d'un développement asymptotique précis que

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O(|t|^{q+\varepsilon}) + o(\Phi(t)). \tag{4.1}$$

Le terme d'erreur délicat dans cette équation est $O(|t|^{q+\varepsilon})$. Si $|t|^q$ était comparable à $\Phi(t)$ pour tout t, on pourrait s'en débarrasser, dire que $\lambda(t)$ est vraiment comparable à $E(e^{\mathbf{i}tf})$ (et donc à $\tilde{\lambda}(t)$ par le même argument), et en déduire que $S_n f$ et $\tilde{S}_n \tilde{f}$ ont les mêmes asymptotiques. Malheureusement, ce n'est pas le cas en général, mais c'est vrai le long d'une sous-suite : comme f n'appartient pas à L^q, on peut trouver un ensemble infini $A \subset \mathbb{N}$ tel que, sur $\Lambda = \bigcup_{n\in A}[2^{-n-1}, 2^{-n}]$,

$$|t|^{q+\varepsilon/2} \leqslant \Phi(t).$$

Cela se démontre aisément par l'absurde.

Le long de Λ, le terme $|t|^{q+\varepsilon}$ est négligeable par rapport à $\Phi(t)$. On déduit donc de (4.1) que, le long de Λ,

$$|\lambda(t)|^2 = |E(e^{\mathbf{i}tf})|^2 + o(\Phi(t)) = 1 - (2 + o(1))\Phi(t). \tag{4.2}$$

Notons que cette asymptotique ne dépend pas de la transformation T, elle ne dépend que de la distribution de f. Même si elle n'est pas valable partout, elle permet de commencer à travailler. Illustrons par exemple son utilité pour montrer le lemme suivant :

LEMME 4.2. *La suite B_n tend vers l'infini.*

DÉMONSTRATION. Si B_n ne tendait pas vers l'infini, on pourrait extraire une sous-suite telle que la distribution de $S_{j(n)}f - A_{j(n)}$ soit tendue. La transformation T étant mélangeante, cela implique d'après [AARONSON et WEISS 2000, Theorem 2] que f s'écrit $u - u \circ T + c$ pour u mesurable et $c \in \mathbb{R}$. Ainsi, $S_n f - nc$ converge en distribution vers $Z = Z_1 - Z_2$, où les Z_i sont des copies indépendantes de u. En particulier, $e^{-\mathbf{i}tnc}E(e^{\mathbf{i}tS_n f})$ converge vers $E(e^{\mathbf{i}tZ})$, donc $|E(e^{\mathbf{i}tS_n f})| \to |E(e^{\mathbf{i}tZ})|$. Comme $E(e^{\mathbf{i}tS_n f}) = c(t)\lambda(t)^n + o(1)$, on obtient $E(e^{\mathbf{i}tZ}) = 0$ si $|\lambda(t)| < 1$.

Le long de Λ, la fonction Φ est positive et $|\lambda(t)| = 1-(2+o(1))\Phi(t)$. Ainsi, $|\lambda(t)| < 1$ pour $t \in \Lambda$ assez petit, puis $E(e^{\mathbf{i}tZ}) = 0$ pour ces mêmes valeurs de t. En particulier, $t \mapsto E(e^{\mathbf{i}tZ})$ n'est pas continue en 0, ce qui est absurde car une fonction caractéristique est toujours continue d'après le théorème de convergence dominée. □

Comme B_n tend vers l'infini, la convergence de $(S_n f - A_n)/B_n$ vers W se traduit en $e^{\mathbf{i}tA_n/B_n}\lambda(t/B_n)^n \to E(e^{\mathbf{i}tW})$. Par conséquent, $|\lambda(t/B_n)|^{2n} \to E(e^{\mathbf{i}tZ})$, où Z est la différence de deux copies indépendantes de W. En prenant le logarithme, on obtient

$$2n \log|\lambda(t/B_n)| \to \log E(e^{\mathbf{i}tZ}). \tag{4.3}$$

De plus, cette convergence est uniforme en t petit. On peut montrer également que B_{n+1}/B_n tend vers 1. Si on veut estimer $|\lambda(u)|$, on peut écrire $u = t/B_n$, et il y a beaucoup de telles décompositions qui doivent toutes donner des résultats compatibles.

En particulier, la fonction $\log E(e^{\mathbf{i}tZ})$ devrait être auto-similaire. Cette intuition est correcte et est formalisée par [Bingham, Goldie et Teugels 1987, Proposition 1.9.4] : la seule convergence (4.3), avec les hypothèses $B_n \to \infty$ et $B_{n+1}/B_n \to 1$, implique que $E(e^{\mathbf{i}tZ}) = e^{-ct^d}$ pour certains $c > 0$ et $d \in \mathbb{R}$ (le fait que ce soit une fonction caractéristique restreignant les possibilités à $d \in]0, 2]$). De plus, la suite normalisatrice B_n s'écrit $B_n = n^{1/d}L(n)$ pour une certaine fonction à variation lente L.

On a obtenu des estimées valables pour tout n, et on a réussi à exhiber un paramètre essentiel du problème, le nombre d (qui devrait être l'indice de la loi stable limite). En réutilisant l'estimée *a priori* (4.2), valable seulement sur l'ensemble de temps Λ, on montre ensuite que $d \leqslant q + \varepsilon/2$.

Par conséquent, si on regarde $\lambda(t/B_n)^n$, le terme d'erreur provenant de $O(|t|^{q+\varepsilon})$ dans (4.1) (qu'on ne savait contrôler que le long de Λ) est au plus

$$n \cdot (|t|/B_n)^{q+\varepsilon} \leqslant |t|^{q+\varepsilon} n \cdot n^{-(q+\varepsilon)/d} L(n)^{-(q+\varepsilon)},$$

qui tend vers 0 lorsque $n \to \infty$ car $d < q + \varepsilon$ Autrement dit, le terme d'erreur problématique dans (4.1) est maintenant contrôlé le long de toutes les valeurs de n. La fin de la preuve du théorème 4.1 est alors une suite de manipulations classiques de fonctions caractéristiques. □

Remarque 4.3. Un aspect de cette preuve mérite d'être souligné : on a vraiment besoin de partir de la convergence en loi de $(S_n f - A_n)/B_n$ pour réussir à dire quelque chose. En particulier, le résultat analogue au théorème 4.1 mais le long de sous-suites ne découle pas de cette preuve.

Plus précisément, on peut se poser la question suivante : soit f une fonction vérifiant $\sum \mu(a) Df(a)^\eta < \infty$, n'appartenant pas à L^2, telle que pour une certaine sous-suite $j(n)$ la suite $(S_{j(n)} f - A_{j(n)})/B_{j(n)}$ converge en distribution vers une loi limite non dégénérée W. Est-il vrai que $(Z_0 + \dots + Z_{j(n)-1} - A_{j(n)})/B_{j(n)}$ converge également en distribution vers W, où les Z_i sont des copies indépendantes de f ? Les arguments qui précèdent ne permettent pas de répondre à cette question (et je m'attendrais plutôt à une réponse négative en général).

5. Développement caractéristique pour les applications Gibbs-Markov

Dans cette section, on donne des éléments de preuve du théorème 3.2. Soit $T : X \to X$ une application Gibbs-Markov. Notons $\mathcal{B}$ l'espace des fonctions η-hölderiennes sur X, et considérons une fonction $f : X \to \mathbb{R}$ avec $\sum_{a \in \Pi} \mu(a) Df(a)^\eta < \infty$. Des calculs classiques permettent de vérifier que l'opérateur $\hat{T}_t$ donné par $\hat{T}_t(u) = \hat{T}(e^{\mathbf{i}tf} u)$ vérifie pour t assez petit

$$\left\| \hat{T}_t^n u \right\|_{\mathcal{B}} \leqslant C\rho^n \|u\|_{\mathcal{B}} + C \|u\|_{L^1}, \tag{5.1}$$

pour un certain $\rho < 1$, et

$$\left\| \hat{T}_t - \hat{T} \right\|_{\mathcal{B} \to \mathcal{B}} = O(|t|^\eta + E|e^{\mathbf{i}tf} - 1|). \tag{5.2}$$

Comme l'injection de $\mathcal{B}$ dans L^1 est compacte, le lemme de Hennion I.2.6 et l'équation (5.1) assurent que le rayon spectral essentiel de $\hat{T}_t$ sur $\mathcal{B}$ est < 1. De plus, comme

$\hat{T}$ a une unique valeur propre de module 1, en 1, l'opérateur $\hat{T}_t$ a pour t petit une unique valeur propre $\lambda(t)$ proche de 1 par continuité. Notons ξ_t la fonction propre correspondante, normalisée par $\int \xi_t = 1$. On déduit de (5.2) que

$$\lambda(t) - 1 = O(|t|^\eta + E|e^{\mathbf{i}tf} - 1|), \quad \|\xi_t - 1\|_{\mathcal{B}} = O(|t|^\eta + E|e^{\mathbf{i}tf} - 1|). \tag{5.3}$$

Comme $E(e^{\mathbf{i}tS_nf}) = \int \hat{T}_t^n 1$, l'existence d'un développement caractéristique pour (f, T) découle de la décomposition spectrale de $\hat{T}_t$. Pour démontrer le théorème 3.2, il s'agit de voir que ce développement est précis, i.e., d'obtenir une bonne estimée asymptotique sur $\lambda(t)$.

La première idée est de réécrire $\lambda(t)$ comme suit :

$$\lambda(t) = \int \lambda(t)\xi_t = \int \hat{T}_t\xi_t = \int \hat{T}_t 1 + \int (\hat{T}_t - \hat{T})(\xi_t - 1).$$

Ainsi,

$$\lambda(t) = E(e^{\mathbf{i}tf}) + \int (e^{\mathbf{i}tf} - 1)(\xi_t - 1). \tag{5.4}$$

Ceci suffit déjà pour conclure dans le cas de fonctions peu intégrables :

LEMME 5.1. *Si $f \notin L^{1+\eta/2}$, alors f admet un développement caractéristique précis.*

DÉMONSTRATION. L'estimée sur ξ_t donnée par l'équation (5.3), injectée dans l'expression (5.4) pour $\lambda(t)$, donne

$$\lambda(t) = E(e^{\mathbf{i}tf}) + O\left(|t|^\eta \int |e^{\mathbf{i}tf} - 1|\right) + O\left(\int |e^{\mathbf{i}tf} - 1|\right)^2.$$

Fixons $p \in [0, 1]$ tel que $f \in L^p$ et $f \notin L^{p+\eta/2}$ (par convention, toutes les fonctions mesurables appartiennent à L^0). Pour tout $x \in \mathbb{R}$, $|e^{\mathbf{i}x} - 1| \leqslant 2|x|^p$. Ainsi,

$$|t|^\eta \int |e^{\mathbf{i}tf} - 1| \leqslant 2|t|^\eta \int |t|^p |f|^p \leqslant C|t|^{p+\eta}.$$

Cela donne le développement caractéristique précis au sens de la définition 3.1, avec $q = p + \eta/2$ et $\varepsilon = \eta/2$. □

Dans le cas où f est plus intégrable, cette approche reposant uniquement sur l'inégalité $|e^{\mathbf{i}x} - 1| \leqslant 2|x|$ ne suffit pas, il faut plutôt exploiter l'inégalité $|e^{\mathbf{i}x} - 1 - \mathbf{i}x| \leqslant C|x|^p$ pour tout $p \in [1, 2]$, et utiliser un développement asymptotique de ξ_t avec plus de termes. Le résultat final qu'on obtient dans cette direction est le suivant :

THÉORÈME 5.2. *Soit $\eta \in]0, 1]$. Il existe une fonction continue $\varepsilon :]1, +\infty[\to \mathbb{R}_+^*$ avec la propriété suivante. Soient $T : X \to X$ une application Gibbs-Markov et $f : X \to \mathbb{R}$ une fonction dans L^p pour $p > 1$ avec $\sum_{a\in\Pi} \mu(a)Df(a)^\eta < \infty$. Il existe alors des nombres complexes c_i (pour $i \in \mathbb{N} \cap [2, p + \varepsilon(p)[$) tels que la valeur propre $\lambda(t)$ de l'opérateur correspondant $\hat{T}_t$ vérifie*

$$\lambda(t) = E(e^{\mathbf{i}tf}) + \sum_{2\leqslant i<p+\varepsilon(p)} c_i t^i + O(|t|^{p+\varepsilon(p)}).$$

L'intérêt de cet énoncé est que le terme d'erreur obtenu est strictement meilleur que l'intégrabilité de la fonction, avec un gain $\varepsilon(p)$ uniforme tant que p reste loin de 1 et $+\infty$ (puisque la fonction ε est continue). La preuve donne d'ailleurs une valeur explicite pour $\varepsilon(p)$, qui n'a pas vraiment d'intérêt en soi.

Ce théorème implique directement le théorème 3.2 dans le cas $f \in L^{1+\eta/2}$. En effet, si $f \in L^2$, le résultat est direct. Sinon, soit ε l'infimum de $\varepsilon(p)$ pour $p \in [1+\eta/2, 2]$. Soit $p \in [1+\eta/2, 2[$ tel que $f \in L^p$ et $f \notin L^{p+\varepsilon/2}$, alors le théorème précédent donne le développement $\lambda(t) = E(e^{\mathbf{i}tf}) + ct^2 + O(|t|^{p+\varepsilon})$, qui est précis pour $q = p + \varepsilon/2$. Comme le cas $f \notin L^{1+\eta/2}$ résulte du lemme 5.1, la démonstration du théorème 3.2 est terminée.

Comme le théorème 5.2 fonctionne pour des degrés d'intégrabilité arbitraires (sans modification de la condition de régularité !), il permet d'aller bien au-delà de la seule démonstration du théorème 3.2. Par exemple, il permet d'obtenir des caractérisations des fonctions f pour lesquelles le théorème central limite a lieu avec une certaine vitesse : ce raffinement du théorème de Berry-Esseen est prouvé pour les variables aléatoires indépendantes identiquement distribuées dans [Ibragimov 1966 ; Ibragimov et Linnik 1971], et leur preuve s'adapte aisément au cas dépendant dès qu'on dispose des estimées précises du théorème 5.2.

La preuve du théorème 5.2 est technique et délicate, je vais juste en dire quelques mots. Partant de l'équation (5.4), il s'agit d'obtenir un développement assez précis de $\xi_t - 1$ pour obtenir une bonne asymptotique. Autrement dit, on cherche une série de Taylor $\xi_t = \sum t^i u_i + O(|t|^{p+\varepsilon})$, où les u_i sont des fonctions avec une certaine intégrabilité. Ce développement asymptotique de la fonction propre de $\hat{T}_t$ serait automatique si $t \mapsto \hat{T}_t = \hat{T}(e^{\mathbf{i}tf}\cdot)$ lui-même était de classe C^p. Cependant, la dérivée première de cette quantité est l'opérateur $u \mapsto \hat{T}(\mathbf{i}fu)$, qui n'envoie en général pas $\mathcal{B}$ dans lui-même (la fonction $\hat{T}f$ n'est en général pas bornée, comme dans l'exemple 1). En revanche, on peut introduire un espace $\mathcal{B}_1$ tel que cet opérateur envoie $\mathcal{B}$ dans $\mathcal{B}_1$, et tel que $\mathcal{B}_1$ soit encore invariant par $\hat{T}$, avec de bonnes propriétés d'intégrabilité et de trou spectral (l'espace $\mathcal{B}_1$ est construit comme un espace d'interpolation entre L^r, pour r bien choisi, et $\mathcal{B}$). En continuant cette construction, on obtient une série de Taylor pour $\hat{T}_t$, mais où les dérivées successives envoient $\mathcal{B}$ dans des espaces de plus en plus faibles. En s'inspirant de [Keller et Liverani 1999 ; Liverani 2003], on peut déduire de ce type de structure une série de Taylor pour ξ_t, mais avec perte : on obtient un terme d'erreur $O(|t|^{p-\delta})$ pour tout $\delta > 0$, ce qui ne semble pas compatible avec l'énoncé du théorème 5.2 où l'on cherche à faire mieux que $O(|t|^p)$. Pour obtenir un gain dans l'exposant, il faut démontrer des propriétés régularisantes des opérateurs du type $u \mapsto \hat{T}(fu)$: lorsque $f \in L^p$, cet opérateur envoie automatiquement L^∞ dans L^p, mais il fait un peu mieux si f vérifie $\sum \mu(a) Df(a)^\eta < +\infty$: il envoie $\mathcal{B}$ dans L^r pour un certain $r > p$. C'est ce gain qui est à la source du gain $\varepsilon(p)$ dans l'énoncé du théorème.

CHAPITRE III

Principe d'invariance presque sûr par la méthode spectrale

Sommaire

1.	Résultat principal probabiliste	31
2.	Applications en dynamique	34
3.	Démonstration du théorème 1.3	37

1. Résultat principal probabiliste

Le principe d'invariance presque sûr est un renforcement très important du théorème central limite : il assure que les trajectoires d'un processus peuvent être mises en correspondance avec celles d'un mouvement brownien de telle sorte que, presque sûrement, l'écart entre les trajectoires soit négligeable par rapport à la taille typique de celles-ci (le résultat peut être plus ou moins précis en fonction des bornes obtenues). On en déduit que les trajectoires du processus étudié partagent de nombreuses propriétés du mouvement brownien.

Ce résultat est maintenant bien connu pour des sommes de variables aléatoires indépendantes identiquement distribuées, mais il est délicat : le terme d'erreur optimal (qui est de $O(n^{1/p})$ pour des variables dans L^p) n'a été démontré pour des processus à valeurs réelles qu'en 1975 dans [Komlós, Major et Tusnády 1975] pour $p \geqslant 4$ (le cas $p < 4$ est plus facile, et peut se traiter par des méthodes différentes, en particulier avec le plongement de Skorokhod). Pour des processus à valeurs dans $\mathbb{R}^d$, le résultat a été démontré encore plus récemment [Einmahl 1989 ; Zaĭtsev 1998], et est un véritable tour de force technique.

Pour les processus faiblement dépendants unidimensionnels, les techniques à base de plongement de Skorokhod fonctionnent bien [Denker et Philipp 1984 ; Hofbauer et Keller 1982], mais ce n'est plus le cas en dimension supérieure. Dans cette direction, un argument d'approximation introduit dans [Berkes et Philipp 1979] a été généralisé pour une large classe de processus faiblement dépendants dans [Melbourne et Nicol 2009] : leurs résultats donnent des termes d'erreur explicites dans le principe d'invariance presque sûr vectoriel, et sont utilisables lorsque les variables étudiées peuvent être suffisamment bien approchées par des variables aléatoires mesurables par rapport à une filtration donnée. Ces résultats s'appliquent en particulier à une large classe de systèmes dynamiques, dès que ceux-ci ont assez d'hyperbolicité et un caractère markovien. Ces techniques sont à rapprocher des techniques de martingales exposées dans la section I.1.

Malheureusement, pour de nombreux systèmes dynamiques, il n'existe pas de filtration naturelle qui se comporte bien. Dans ce cas, même si les techniques de martingales ne s'appliquent pas, on peut malgré tout souvent faire appel à la technique spectrale de Nagaev-Guivarc'h (voir la section I.4) pour obtenir des théorèmes limites classiques. En un mot, cette technique donne une approximation extrêmement précise de la fonction caractéristique de $S_n f$, ce qui permet essentiellement de faire fonctionner dans un cadre dépendant les preuves initialement inventées pour des variables indépendantes. Cette description des fonctions caractéristiques est tellement précise qu'on s'attend à ce que virtuellement n'importe quel théorème limite soit démontrable dans ce cadre, et en particulier le principe d'invariance presque sûr.

C'est ce problème que nous allons résoudre dans ce chapitre (correspondant à l'article [Gouëzel 2010a]). Les méthodes sont purement probabilistes : nous allons décrire comment des informations assez précises sur la fonction caractéristique d'un processus impliquent le principe d'invariance presque sûr (avec un terme d'erreur bien contrôlé, et indépendant de la dimension contrairement à [Melbourne et Nicol 2009]). Ces hypothèses sont automatiquement satisfaites quand on est dans le cadre de la méthode spectrale (et même sous des hypothèses plus faibles, voir la remarque 2.2), si bien qu'on obtient directement des applications à de nombreuses classes de systèmes dynamiques, comme on le verra dans la section 2.

Pour $d > 0$, considérons un processus $(A_0, A_1, \dots)$ à valeurs dans $\mathbb{R}^d$, borné dans L^p pour un certain $p > 2$. Sous certaines hypothèses, nous souhaitons voir qu'il vérifie un principe d'invariance presque sûr dans le sens suivant.

Définition 1.1. *Soient $\lambda \in]0, 1/2]$ et Σ^2 une matrice $d \times d$ symétrique positive (éventuellement dégénérée). Un processus $(A_0, A_1, \dots)$ à valeurs dans $\mathbb{R}^d$ satisfait un principe d'invariance presque sûr avec erreur $o(n^\lambda)$ et covariance limite Σ^2 s'il existe un espace de probabilité Ω et deux processus $(A_0^*, A_1^*, \dots)$ et $(B_0, B_1, \dots)$ sur Ω tels que :*

(1) *les processus $(A_0, A_1, \dots)$ et $(A_0^*, A_1^*, \dots)$ ont même loi ;*
(2) *les variables aléatoires $B_0, B_1, \dots$ sont indépendantes et de loi $\mathcal{N}(0, \Sigma^2)$;*
(3) *presque sûrement dans Ω,*

$$\left| \sum_{\ell=0}^{n-1} A_\ell^* - \sum_{\ell=0}^{n-1} B_\ell \right| = o(n^\lambda).$$

Un mouvement brownien aux temps entiers coïncidant avec une somme de variables gaussiennes i.i.d., cette définition peut aussi se formuler comme un résultat d'approximation presque sûre par des trajectoires browniennes avec un terme d'erreur $o(n^\lambda)$.

Ce résultat implique le théorème central limite, et le théorème central limite fonctionnel (aussi appelé principe d'invariance) : le processus renormalisé converge en distribution sur $[0, 1]$ vers la mesure de Wiener. Le principe d'invariance presque sûr est un résultat beaucoup plus fort, qui implique par exemple la loi du logarithme itéré.

Notre objectif est de donner des hypothèses portant seulement sur la fonction caractéristique du processus $(A_0, A_1, \dots)$ qui impliquent un principe d'invariance presque sûr. Pour simplifier les notations, nous écrirons $e^{\mathbf{i}tx}$ au lieu de $e^{\mathbf{i}\langle t,x\rangle}$ pour $t \in \mathbb{R}^d$ et $x \in \mathbb{R}^d$.

Formulons maintenant notre hypothèse principale (H), qui assure que le processus considéré est suffisamment proche d'un processus indépendant : il existe $\varepsilon_0 > 0$ et $C, c > 0$ tels que, pour tous $n, m > 0$, tous $b_1 < b_2 < \cdots < b_{n+m+1}$, tout $k > 0$ et tous $t_1, \dots, t_{n+m} \in \mathbb{R}^d$ avec $|t_j| \leqslant \varepsilon_0$, on ait

$$\tag{H} \begin{aligned} &\Bigg| E\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell\right)+\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell\right)}\right) \\ &\qquad - E\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell\right)}\right)\cdot E\left(e^{\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell\right)}\right)\Bigg| \\ &\leqslant C(1+\max|b_{j+1}-b_j|)^{C(n+m)}e^{-ck}. \end{aligned}$$

Cette hypothèse signifie que, si l'on regroupe les variables aléatoires en $n+m$ blocs, alors un intervalle de taille k entre deux blocs donne des fonctions caractéristiques qui sont exponentiellement proches (en k) de fonctions caractéristiques indépendantes, avec une erreur qui est, pour chaque bloc, polynomiale en la taille du bloc. Ce contrôle est uniquement requis pour des fréquences t_j proches de 0.

Cette condition est évidemment vraie pour des variables aléatoires indépendantes. Son intérêt est qu'elle est aussi automatiquement satisfaite pour des systèmes dynamiques dès qu'on est dans le cadre de la méthode spectrale, voir le théorème 2.1 ci-dessous.

Le résultat principal de ce chapitre est le théorème suivant.

THÉORÈME 1.2. *Soit* $(A_0, A_1, \dots)$ *un processus stationnaire centré à valeurs dans* $\mathbb{R}^d$*, dans* L^p *pour* $p > 2$*, satisfaisant* (H)*. Alors*

(1) *la matrice de covariance* $\operatorname{cov}(\sum_{\ell=0}^{n-1} A_\ell)/n$ *converge vers une matrice* Σ^2 *;*

(2) *la suite* $\sum_{\ell=0}^{n-1} A_\ell/\sqrt{n}$ *converge en distribution vers* $\mathcal{N}(0, \Sigma^2)$ *;*

(3) *le processus* $(A_0, A_1, \dots)$ *satisfait un principe d'invariance presque sûr avec covariance limite* Σ^2 *et erreur* $o(n^\lambda)$ *pour tout* λ *tel que*

$$\lambda > \frac{p}{4p-4} = \frac{1}{4} + \frac{1}{4p-4}.$$

Lorsque $p = \infty$, cette condition sur l'erreur devient $\lambda > 1/4$, ce qui est assez bon (notons en particulier que la borne obtenue ne dépend pas de la dimension) : cette condition n'avait auparavant été obtenue que pour des classes de systèmes dynamiques très spécifiques (invariantes en particulier par renversement du temps) pour des observables à valeurs réelles (voir par exemple [FIELD, MELBOURNE et TÖRÖK 2003; MELBOURNE et TÖRÖK 2004]).

Si le processus n'est pas stationnaire, on a besoin d'une hypothèse supplémentaire qui assure la convergence (assez rapide) des covariances :

THÉORÈME 1.3. *Soit* $(A_0, A_1, \dots)$ *un processus à valeurs dans* $\mathbb{R}^d$*, dans* L^p *pour* $p > 2$*, satisfaisant* (H)*. On suppose de plus que* $\sum |E(A_\ell)| < \infty$*, et qu'il existe une* *matrice* Σ^2 *telle que, pour tout* $\alpha > 0$,

$$\left| \operatorname{cov}\left(\sum_{\ell=m}^{m+n-1} A_\ell \right) - n\Sigma^2 \right| \leqslant C_\alpha n^\alpha, \tag{1.1}$$

uniformément en m *et* n. *Alors la suite* $\sum_{\ell=0}^{n-1} A_\ell/\sqrt{n}$ *converge en distribution vers* $\mathcal{N}(0,\Sigma^2)$, *et le processus* $(A_0, A_1, \dots)$ *satisfait un principe d'invariance presque sûr avec covariance limite* Σ^2 *et terme d'erreur* $o(n^\lambda)$ *pour tout* $\lambda > p/(4p-4)$.

Le théorème 1.2 est une conséquence du théorème 1.3, car un processus stationnaire vérifiant (H) vérifie toujours (1.1) (on peut même prendre $\alpha = 0$ dans cette inégalité). Ce résultat est une conséquence facile des techniques que nous allons exposer plus loin.

2. Applications en dynamique

Considérons un système dynamique $T : X \to X$ préservant une mesure μ, et notons $\hat{T}$ son opérateur de transfert. Soit $f : X \to \mathbb{R}^d$, on introduit comme il est habituel dans la méthode spectrale de NAGAEV-GUIVARC'H l'opérateur perturbé $\hat{T}_t(u) = \hat{T}(e^{\mathbf{i}tf}u)$, pour $t \in \mathbb{R}^d$. L'intérêt premier de cet opérateur est que $E(e^{\mathbf{i}tS_nf}) = \int \hat{T}_t^n(1)\,\mathrm{d}\mu$, mais on vérifie même que pour tous $t_0, \dots, t_{n-1} \in \mathbb{R}^d$,

$$E\left(e^{\mathbf{i}\sum_{\ell=0}^{n-1} t_\ell f\circ T^\ell}\right) = \int \hat{T}_{t_{n-1}} \cdots \hat{T}_{t_0} 1(x)\,\mathrm{d}\mu(x).$$

Si $\hat{T}_t$ a de bonnes propriétés spectrales sur un certain espace de Banach (par exemple les fonctions lipschitziennes si T est une application dilatante de l'intervalle à branches surjectives, voir la section I.4), on en déduit aisément que (H) est vérifiée.

Dans des situations plus complexes, il est souvent possible de coder de la même manière les fonctions caractéristiques du processus étudié par une famille d'opérateurs. Cependant, ces opérateurs peuvent agir sur des espaces de Banach plus compliqués (de distributions, de mesures, voire d'objets plus exotiques, comme dans le chapitre IV). Il est donc souhaitable d'introduire un cadre un peu plus abstrait pour formaliser les propriétés essentielles de ce type de codage, comme suit.

Considérons un processus $(A_0, A_1, \dots)$ à valeurs dans $\mathbb{R}^d$. Soient $\mathcal{B}$ un espace de Banach et $\mathcal{L}_t$ (pour $t \in \mathbb{R}^d$, $|t| \leqslant \varepsilon_0$) des opérateurs linéaires agissant continûment sur $\mathcal{B}$. On suppose qu'il existe $u_0 \in \mathcal{B}$ et $\xi_0 \in \mathcal{B}'$ (le dual de $\mathcal{B}$) tels que, pour tous $t_0, \dots, t_{n-1} \in \mathbb{R}^d$ avec $|t_j| \leqslant \varepsilon_0$,

$$E\left(e^{\mathbf{i}\sum_{\ell=0}^{n-1} t_\ell A_\ell}\right) = \langle \xi_0, \mathcal{L}_{t_{n-1}} \mathcal{L}_{t_{n-2}} \cdots \mathcal{L}_{t_1} \mathcal{L}_{t_0} u_0 \rangle.$$

On dit que la fonction caractéristique de $(A_0, A_1, \dots)$ est codée par $(\mathcal{B}, (\mathcal{L}_t)_{|t|\leqslant\varepsilon_0}, u_0, \xi_0)$.

Dans ces conditions, la propriété (H) est une conséquence de certaines hypothèses sur les opérateurs $\mathcal{L}_t$, que nous allons maintenant décrire.

(I1) On peut écrire $\mathcal{L}_0 = \Pi + Q$ où Π est une projection sur un espace de dimension 1 et Q est un opérateur sur $\mathcal{B}$, avec $Q\Pi = \Pi Q = 0$, et $\|Q^n\|_{\mathcal{B}\to\mathcal{B}} \leqslant C\rho^n$ pour un certain $\rho < 1$.

(I2) Il existe $C > 0$ tel que $\|\mathcal{L}_t^n\|_{\mathcal{B}\to\mathcal{B}} \leqslant C$ pour tout $n \in \mathbb{N}$ et tout $t \in \mathbb{R}^d$ avec $|t| \leqslant \varepsilon_0$.

On notera (I) cet ensemble de conditions.

THÉORÈME 2.1. *Soit (A_ℓ) un processus dont la fonction caractéristique est codée par une famille d'opérateurs $(\mathcal{L}_t)$, et qui est borné dans L^p, $p > 2$. Supposons (I) satisfaite. Il existe alors $a \in \mathbb{R}^d$ et une matrice Σ^2 tels que $\left(\sum_{\ell=0}^{n-1} A_\ell - na\right)/\sqrt{n}$ converge vers $\mathcal{N}(0, \Sigma^2)$. De plus, le processus $(A_\ell - a)_{\ell\in\mathbb{N}}$ vérifie le principe d'invariance presque sûr avec covariance limite Σ^2 et erreur $o(n^\lambda)$, pour tout $\lambda > p/(4p-4)$.*

La preuve consiste à construire a comme limite des $E(A_\ell)$, donner une formule explicite pour Σ^2, et déduire le résultat du théorème 1.3 en montrant que (H) et (1.1) sont des conséquences de (I). On a même sous les hypothèses du théorème 2.1

$$\left|\operatorname{cov}\left(\sum_{\ell=m}^{m+n-1} A_\ell\right) - n\Sigma^2\right| \leqslant C.$$

Donnons quelques détails sur la vérification de (H), qui est le point crucial. On écrit

$$\begin{aligned}
E&\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell\right)+\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell\right)}\right)\\
&= \left\langle \xi_0, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} \mathcal{L}_0^k \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} u_0 \right\rangle\\
&= \left\langle \xi_0, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} (\mathcal{L}_0^k - \Pi) \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} u_0 \right\rangle\\
&\quad + \left\langle \xi_0, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} \Pi \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} u_0 \right\rangle.
\end{aligned}$$

Tous les opérateurs $\mathcal{L}_{t_i}$ satisfont $\left\|\mathcal{L}_{t_i}^j\right\| \leqslant C$. Comme $\left\|\mathcal{L}_0^k - \Pi\right\| \leqslant C\rho^k$, on en déduit que le terme sur l'avant-dernière ligne est borné par $C^{n+m}\rho^k$, ce qui est un terme d'erreur compatible avec (H). Quant à la dernière ligne, la projection Π découple les deux parties, si bien que cette ligne est égale à

$$E\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} A_\ell\right)}\right) \cdot E\left(e^{\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} A_\ell\right)}\right) + O(C^m\rho^{b_{n+1}+k}).$$

Cela démontre (H).

REMARQUE 2.2. La méthode spectrale, dans son acception habituelle décrite dans la section I.4 (et exploitée également dans le chapitre II) consiste à partir de l'existence d'une valeur propre $\lambda(t)$ proche de 1 de l'opérateur perturbé $\hat{T}_t$, et à relier les asymptotiques de $S_n f$ (n grand) à celles de $\lambda(t)$ (t petit). Il est remarquable que le théorème précédent ne requière que des hypothèses significativement plus faibles : nous n'avons pas besoin de bonnes asymptotiques pour la valeur propre perturbée, ni même de l'existence de celle-ci. Cette différence est liée à la technique de preuve, qui est ici beaucoup plus probabiliste ; le prix à payer est l'hypothèse supplémentaire de moment L^p, $p > 2$, à comparer avec l'hypothèse L^2 qui suffit en général pour faire fonctionner la méthode spectrale. En particulier, sous les hypothèses du théorème 2.1, même le théorème central limite (qui est une conséquence triviale du principe d'invariance presque sûr) est nouveau.

Nous allons maintenant expliquer comment vérifier en pratique les hypothèses de ce théorème. Soient $T : X \to X$ un système dynamique, μ une mesure de probabilité (invariante ou non) et $f : X \to \mathbb{R}^d$. On désire étudier le processus $(f, f \circ T, f \circ T^2, \dots)$ sur l'espace probabilisé (X, μ). Supposons que sa fonction caractéristique est codée par une famille d'opérateurs $\mathcal{L}_t$ sur un espace de Banach $\mathcal{B}$, et que l'opérateur $\mathcal{L}_0$ satisfait (I1), i.e., il a une valeur propre simple en 1 et le reste de son spectre est contenu dans un disque de rayon strictement plus petit (ce type de comportement est souvent une conséquence du lemme de Hennion I.2.6).

Proposition 2.3. *Si la famille $\mathcal{L}_t : \mathcal{B} \to \mathcal{B}$ dépend continûment du paramètre t en $t = 0$ (au sens de la norme), alors* (I2) *est satisfaite.*

Démonstration. Par des arguments classiques de perturbation des opérateurs, la description spectrale de $\mathcal{L}_0$ survit pour t petit : on peut écrire $\mathcal{L}_t = \lambda(t)\Pi_t + Q_t$ où $\lambda(t) \in \mathbb{C}$, Π_t est une projection sur un espace de dimension 1 et $\|Q_t^n\| \leqslant C\rho^n$ avec $\rho < 1$, uniformément en t petit et $n \in \mathbb{N}$. Si $|\lambda(t)| \leqslant 1$ pour t petit, on en déduit (I2).

On a

$$
\begin{aligned}
E(e^{\mathbf{i}t \sum_{\ell=0}^{n-1} f\circ T^\ell}) &= \langle \xi_0, \mathcal{L}_t^n u_0 \rangle = \lambda(t)^n \langle \xi_0, \Pi_t u_0 \rangle + \langle \xi_0, Q_t^n u_0 \rangle \\
&= \lambda(t)^n \langle \xi_0, \Pi_t u_0 \rangle + O(\rho^n).
\end{aligned}
\tag{2.1}
$$

Lorsque $t \to 0$, par continuité, la quantité $\langle \xi_0, \Pi_t u_0 \rangle$ converge vers $\langle \xi_0, \Pi u_0 \rangle = 1$. En particulier, si t est assez petit, $\langle \xi_0, \Pi_t u_0 \rangle \neq 0$. Comme le membre de droite de (2.1) doit rester borné en module par 1, cela donne $|\lambda(t)| \leqslant 1$. □

Donnons maintenant des exemples spécifiques. Soient T un décalage de type fini irréductible apériodique, μ une mesure de Gibbs et $f : X \to \mathbb{R}^d$ une fonction hölderienne de moyenne nulle. Soient $\hat{T}$ l'opérateur de transfert associé et $\hat{T}_t = \hat{T}(e^{\mathbf{i}tf}\cdot)$ l'opérateur perturbé. Ces opérateurs codent la fonction caractéristique de $(f, f \circ T, \dots)$ et dépendent analytiquement de t (c'est une conséquence du développement en série $e^{\mathbf{i}x} = \sum(\mathbf{i}x)^n/n!$ et du fait que les fonctions hölderiennes forment une algèbre de Banach). La condition (I) est classique (voir par exemple [Guivarc'h et Hardy 1988 ; Parry et Pollicott 1990]). Par conséquent, le théorème 2.1 donne le principe d'invariance presque sûr avec terme d'erreur $o(n^\lambda)$ pour tout $\lambda > 1/4$. Ce résultat n'est pas nouveau, il est déjà démontré dans [Melbourne et Nicol 2009] (avec un terme d'erreur moins précis).

Si T est une transformation Anosov ou Axiom A et $f : X \to \mathbb{R}^d$ est hölderienne, on en déduit le même résultat en utilisant la dynamique symbolique. On peut aussi l'éviter et directement appliquer le théorème 2.1 à l'opérateur de transfert agissant sur un espace de Banach de formes différentielles généralisées, comme dans [Baladi et Tsujii 2007 ; Gouëzel et Liverani 2008] (voir le chapitre IV).

Soit maintenant $T : X \to X$ une application dilatante par morceaux, dans laquelle la dilatation domine la complexité (dans le sens de [Buzzi 1997] ou [Saussol 2000, Lemma 2.2]) ; ceci inclut en particulier toutes les applications dilatantes par morceaux de l'intervalle, puisque la complexité y est automatiquement sous-exponentielle. Soit $f : X \to \mathbb{R}^d$ une fonction β-hölderienne, $\beta \in]0,1]$. L'opérateur de transfert perturbé $\hat{T}_t$ agit continûment sur l'espace de Banach $\mathcal{B} = V_\beta$ de [Saussol 2000], et dépend analytiquement de t (puisque $\mathcal{B}$ est une algèbre de Banach). Le théorème 2.1 donne un principe d'invariance presque sûr avec terme d'erreur $o(n^\lambda)$ pour tout $\lambda > 1/4$. Ce résultat n'était connu que pour $\dim(X) = 1$ et $d = 1$ [Hofbauer et Keller 1982], car il n'y a pas en général de bonne filtration associée à ce type de dynamique.

Ce résultat s'applique aussi aux applications dilatantes de l'intervalle faiblement couplées sur réseau, puisque [Bardet, Gouëzel et Keller 2007] démontre la condition (I) dans ce cadre. L'espace de Banach $\mathcal{B}$ est ici un espace vraiment exotique (les éléments de $\mathcal{B}$ sont des limites projectives de classes compatibles de mesures), mais cela ne pose pas de problème vu que nous avons formulé nos résultats dans un cadre suffisamment général.

Supposons maintenant que T est le temps 1 d'un flot d'Anosov de contact. Tsujii construit dans [Tsujii 2010a,b] des espaces de Banach de distributions sur lesquels l'opérateur de transfert $\hat{T}$ agit avec un trou spectral. Si f est assez lisse, alors $\hat{T}_t$ dépend analytiquement de t, et on obtient encore un principe d'invariance presque sûr. Ce résultat était connu pour les observables réelles [Melbourne et Török 2002], mais est nouveau en dimension d. Notons cependant que nos méthodes ne s'appliquent pas à tous les flots hyperboliques qui mélangent superpolynomialement, contrairement aux arguments de martingales de [Melbourne et Török 2002].

Mentionnons aussi l'article [Luzzatto et Melbourne 2011], où le théorème 2.1 est appliqué à toute une classe d'applications unimodales.

Finalement, considérons le cas des applications Gibbs-Markov étudié dans le chapitre II (voir la définition II.2.1). Soient $T : X \to X$ une application Gibbs-Markov et $f : X \to \mathbb{R}^d$ une fonction de moyenne nulle vérifiant $\sum_{a\in\Pi} \mu(a) Df(a)^\eta < \infty$, pour un certain $\eta > 0$. Si $f \in L^2$, le théorème II.2.2 assure que les sommes de Birkhoff $S_n f$ vérifient un théorème central limite. Lorsque $f \in L^p$, $p > 2$, le théorème 2.1 montre qu'on a un principe d'invariance presque sûr avec terme d'erreur $o(n^\lambda)$ pour tout $\lambda > p/(4p-4)$.

3. Démonstration du théorème 1.3

La démonstration du théorème 1.3 utilise exclusivement des outils probabilistes, et repose de manière centrale sur la notion de couplage.

Si $Z_1 : \Omega_1 \to E_1$ et $Z_2 : \Omega_2 \to E_2$ sont deux variables aléatoires sur deux espaces probabilisés (éventuellement différents), un *couplage* entre Z_1 et Z_2 est une manière d'associer ces variables aléatoires, en général de manière à montrer que celles-ci sont proches en un certain sens. Formellement, un couplage entre Z_1 et Z_2 est donné par un espace probabilisé Ω' et deux variables aléatoires $Z'_1 : \Omega' \to E_1$ et $Z'_2 : \Omega' \to E_2$ telles que Z'_i est distribuée comme Z_i. En considérant la distribution de (Z'_1, Z'_2) dans $E_1 \times E_2$, on peut prendre sans perte de généralité $\Omega' = E_1 \times E_2$, et Z'_1 et Z'_2 les première et seconde projections. Si $E_1 = E_2$, on cherche souvent à assurer que Z'_1 et Z'_2 sont proches par exemple en minimisant $\|Z'_1 - Z'_2\|_{L^p}$ pour un certain p, ou $P(Z'_1 \neq Z'_2)$.

DÉFINITION 3.1. *Soient P, Q deux mesures de probabilité sur un espace métrique. On définit leur distance de Prokhorov $\pi(P,Q)$ comme étant le plus petit $\varepsilon > 0$ tel que $P(B) \leqslant \varepsilon + Q(B^\varepsilon)$ pour tout borélien B, où B^ε désigne le ε-voisinage de B.*

L'intérêt de cette distance (qui est bien symétrique, même si ce n'est pas complètement évident au vu de la définition) est qu'elle permet de construire de bons couplages, d'après le théorème suivant dû à STRASSEN et DUDLEY (voir [BILLINGSLEY 1999, Theorem 6.9]).

THÉORÈME 3.2. *Soient X et Y deux variables aléatoires à valeurs dans un espace métrique, de distributions respectives P_X et P_Y. Si $\pi(P_X, P_Y) < c$, il existe un couplage entre X et Y tel que $P(d(X,Y) > c) < c$.*

Pour construire de bons couplages, nous disposerons seulement d'informations sur les fonctions caractéristiques des processus. Il sera donc important de pouvoir estimer la distance de Prokhorov uniquement en termes de fonctions caractéristiques. C'est ce qui est fait dans le lemme suivant. Soient $d > 0$ et $N > 0$. On considère $\mathbb{R}^{dN}$ muni de la norme

$$|(x_1, \dots, x_N)|_N = \sup_{1 \leqslant i \leqslant N} |x_i|,$$

où $|x|$ est la norme euclidienne de $x \in \mathbb{R}^d$.

LEMME 3.3. *Il existe une constante $C(d)$ avec la propriété suivante. Soient F et G deux mesures de probabilité sur $\mathbb{R}^{dN}$, de fonctions caractéristiques respectives φ et γ. Pour tout $T > 0$,*

$$\pi(F,G) \leqslant \sum_{j=1}^{N} F(|x_j| > T) + \left(C(d)T^{d/2}\right)^N \left[\int_{\mathbb{R}^{dN}} |\varphi - \gamma|^2\right]^{1/2}.$$

DÉMONSTRATION. Grâce à un argument d'approximation, on peut supposer que F et G ont des densités f et g. Pour tout ensemble mesurable A,

$$\begin{aligned} F(A) - G(A) &\leqslant F(A \cap \max|x_j| \leqslant T) + F(\max|x_j| > T) - G(A \cap \max|x_j| \leqslant T) \\ &\leqslant \int_{|x_1|,\dots,|x_N| \leqslant T} |f - g| + \sum_{j=1}^{N} F(|x_j| > T). \end{aligned}$$

Par conséquent, $\pi(F,G)$ est borné par la dernière ligne de cette équation. Pour conclure, il suffit donc d'estimer $\int_{|x_1|,\dots,|x_N| \leqslant T} |f-g|$. On a

$$\int_{|x_1|,\dots,|x_N| \leqslant T} |f - g| \leqslant \|f - g\|_{L^2} \left\|1_{|x_1|,\dots,|x_N| \leqslant T}\right\|_{L^2} = \|\varphi - \gamma\|_{L^2} (CT)^{dN/2},$$

puisque la transformée de Fourier est une isométrie sur L^2 (à un facteur $(2\pi)^{dN/2}$ près). □

Illustrons l'utilité des outils précédents avec la proposition suivante.

Proposition 3.4. *Soit $(A_0, A_1, \dots)$ un processus centré, borné dans L^p, $p>2$, et vérifiant (H). Pour tout $\eta>0$, il existe $C>0$ tel que, pour tous $m,n\in\mathbb{N}$,*

$$\left\| \sum_{\ell=m}^{m+n-1} A_\ell \right\|_{L^{p-\eta}} \leqslant Cn^{1/2}.$$

Ce type d'estimées de moments est classique pour une large classe de processus faiblement dépendants. L'intérêt de la proposition est que, ici, ces bornes sont établies en partant uniquement de l'hypothèse (H) sur la fonction caractéristique du processus, qui ne donne à première vue aucune information sur les moments.

Démonstration. L'étape cruciale est d'obtenir une borne dans L^2 ; la borne dans $L^{p-\eta}$ s'en déduit en utilisant le même genre de techniques et l'inégalité de Rosenthal pour des sommes de variables aléatoires indépendantes [Rosenthal 1970].

Notons $u_n = \max_{m\in\mathbb{N}} \left\|\sum_{\ell=m}^{m+n-1} A_\ell\right\|_{L^2}^2$. On va montrer que u_{a+b} est essentiellement borné par u_a+u_b, ce qui permet de conclure par récurrence. Pour cela, on décompose la somme $\sum_{\ell=m}^{m+a+b-1} A_\ell$ comme $X_1+X_2+X_3$ où $X_1=\sum_m^{m+a-1} A_\ell$ et $X_3=\sum_{m+a+b^\alpha}^{m+a+b-1} A_\ell$, tandis que X_2 est la somme sur l'intervalle central. Si α est assez petit, X_2 est négligeable et peut être ignoré. Soient X_1' et X_3' deux copies indépendantes de X_1 et X_3. L'hypothèse (H) assure que la fonction caractéristique de (X_1, X_3) ressemble à celle de (X_1', X_3'). Ainsi, le lemme 3.3 permet de les coupler, avec une différence entre les deux très petite sur la plupart de l'espace. On obtient

$$\|X_1+X_3\|_{L^2} \leqslant \|X_1-X_1'\|_{L^2} + \|X_3-X_3'\|_{L^2} + \|X_1'+X_3'\|_{L^2}.$$

Les deux premiers termes sont petits, tandis que le troisième est égal à $(\|X_1'\|_{L^2}^2 + \|X_3'\|_{L^2}^2)^{1/2}$ par indépendance, et est donc borné par $(u_a+u_b)^{1/2}$. On a obtenu $u_{a+b}\leqslant u_a+u_b+r(a,b)$, où le reste $r(a,b)$ est suffisamment petit pour ne pas poser de problème. □

Je vais maintenant donner quelques idées de la preuve du théorème 1.3. Considérons donc un processus $(A_0, A_1, \dots)$ satisfaisant les hypothèses de ce théorème, pour une certaine matrice non dégénérée Σ^2 (le cas dégénéré doit être traité différemment, il est plus facile). On peut sans perte de généralité remplacer A_ℓ par $A_\ell - E(A_\ell)$, et supposer le processus centré.

La stratégie de preuve est très classique : c'est une stratégie de "gros blocs, petits blocs" où les gros blocs correspondent aux indices des variables aléatoires qu'on va vraiment coupler, et les petits blocs correspondent à des variables qu'on va négliger ; si les petits blocs sont effectivement petits, le contrôle de moment donné par la proposition 3.4 permet de majorer leur contribution presque sûrement, tandis que les gros blocs sont suffisamment éloignés les uns des autres pour que l'hypothèse (H) et le lemme 3.3 permettent de les considérer comme quasiment indépendants.

Cette méthode est généralement mise en œuvre en utilisant des blocs de taille polynomiale. Dans notre cas, on obtient de meilleures estimées en utilisant une approche triadique à la Cantor, comme suit. On commence par écrire $\mathbb{N}$ comme la réunion des intervalles $[2^n, 2^{n+1}[$. Dans chacun de ces intervalles, on met un petit bloc (pas trop petit) au centre, puis un petit bloc (un peu plus petit) au centre de chacun des intervalles ainsi créés, puis un petit bloc (encore plus petit) au milieu des quatre nouveaux intervalles, etc. L'intérêt de cette approche est que, pour obtenir n gros blocs bien séparés, l'argument classique a besoin de petits blocs dont la somme des tailles est de l'ordre de n^2, alors qu'ici on y arrive avec des petits blocs globalement de taille n.

La construction précise dépend de deux paramètres $\beta \in \left]0,1\right[$ et $\varepsilon < 1-\beta$. Soit $f = f(n) = \lfloor \beta n \rfloor$. On décompose $[2^n, 2^{n+1}[$ comme une réunion de $F = 2^f$ intervalles $(I_{n,j})_{0 \leqslant j < F}$ tous de même longueur (les gros blocs), et F petits blocs $(J_{n,j})_{0 \leqslant j < F}$ qui assurent suffisamment d'indépendance, construits comme expliqué ci-dessus. Le petit bloc central $J_{n,F/2}$ a une longueur $2^{\lfloor \varepsilon n \rfloor} 2^{f-1}$, les deux petits blocs suivants ont taille moitié, et ainsi de suite. Les gros blocs et les petits blocs sont disposés alternativement et par ordre croissant, comme suit :

$$\left[2^n, 2^{n+1}\right[= J_{n,0} \cup I_{n,0} \cup J_{n,1} \cup I_{n,1} \cdots \cup J_{n,F-1} \cup I_{n,F-1}.$$

On notera $(n', j') \prec (n, j)$ si l'intervalle $I_{n',j'}$ est à gauche de $I_{n,j}$ (cela correspond à l'ordre lexicographique), et $i_{n,j}$ le plus petit élément de $I_{n,j}$.

Soit $X_{n,j} = \sum_{\ell \in I_{n,j}} A_\ell$, pour $n \in \mathbb{N}$ et $0 \leqslant j < F(n)$. Notons $\mathcal{I} = \bigcup_{n,j} I_{n,j}$ la réunion des gros blocs (sur lesquels on va effectivement coupler) et $\mathcal{J} = \bigcup_{n,j} J_{n,j}$ la réunion des petits blocs (qu'on va négliger). Les différentes étapes de la preuve du théorème 1.3 sont les suivantes :

(1) Il existe un couplage entre $(X_{n,j})$ et une suite de variables aléatoires indépendants $(Y_{n,j})$, où $Y_{n,j}$ est distribuée comme $X_{n,j}$, tel que, presque sûrement, lorsque $(n, j) \to \infty$,

$$\left| \sum_{(n',j') \prec (n,j)} X_{n',j'} - Y_{n',j'} \right| = o(2^{(\beta+\varepsilon)n/2}).$$

(2) Il existe un couplage entre $(Y_{n,j})$ et une suite de variables aléatoires indépendantes gaussiennes $Z_{n,j}$, de covariance $\operatorname{cov}(Z_{n,j}) = |I_{n,j}|\Sigma^2$, tel que, presque sûrement, lorsque $(n, j) \to \infty$,

$$\left| \sum_{(n',j') \prec (n,j)} Y_{n',j'} - Z_{n',j'} \right| = o(2^{(\beta+\varepsilon)n/2} + 2^{((1-\beta)/2+\beta/p+\varepsilon)n}).$$

(3) En couplant $X_{n,j}$ avec $Z_{n,j}$ grâce aux deux premières étapes et en écrivant $Z_{n,j}$ comme la somme de $|I_{n,j}|$ variables gaussiennes $\mathcal{N}(0, \Sigma^2)$, on obtient un couplage entre $(A_\ell)_{\ell \in \mathcal{I}}$ et $(B_\ell)_{\ell \in \mathcal{I}}$ où les B_ℓ sont indépendants et distribués comme $\mathcal{N}(0, \Sigma^2)$. De plus, lorsque (n, j) tend vers l'infini,

$$\left| \sum_{\ell < i_{n,j},\ \ell \in \mathcal{I}} A_\ell - B_\ell \right| = o(2^{(\beta+\varepsilon)n/2} + 2^{((1-\beta)/2+\beta/p+\varepsilon)n}).$$

(4) On vérifie que, presque sûrement, lorsque $(n,j)\to\infty$,
$$\max_{m<|I_{n,j}|}\left|\sum_{\ell=i_{n,j}}^{i_{n,j}+m} A_\ell\right| = o(2^{((1-\beta)/2+\beta/p+\varepsilon)n}).$$
De plus, les B_ℓ vérifient une estimée similaire.

(5) En réunissant les deux dernières étapes, on obtient lorsque k tend vers l'infini
$$\left|\sum_{\ell<k,\ \ell\in\mathcal{I}} A_\ell - B_\ell\right| = o(k^{(\beta+\varepsilon)/2} + k^{(1-\beta)/2+\beta/p+\varepsilon}).$$

(6) Finalement, on montre que les petits blocs sont négligeables : presque sûrement,
$$\sum_{\ell<k,\ \ell\in\mathcal{J}} A_\ell = o(k^{\beta/2+\varepsilon}),$$
et les B_ℓ vérifient la même estimée.

On obtient ainsi un couplage tel que, presque sûrement
$$\left|\sum_{\ell<k} A_\ell - B_\ell\right| = o(k^{\beta/2+\varepsilon} + k^{(1-\beta)/2+\beta/p+\varepsilon}).$$
On choisit enfin β de telle sorte que les deux termes d'erreur coïncident, i.e., $\beta = p/(2p-2)$. On obtient le principe d'invariance presque sûr avec terme d'erreur $o(n^{p/(4p-4)+\varepsilon})$, pour tout $\varepsilon>0$. Comme le principe d'invariance presque sûr implique le théorème central limite, cela démontre le théorème 1.3.

Les étapes (3) et (5) ci-dessus étant triviales, il reste à justifier les autres. Elles ne sont pas très difficiles : (1) découle de l'hypothèse (H) et du lemme 3.3, (2) consiste à coupler des variables indépendantes et a donc déjà été fait dans la littérature (voir par exemple [Zaĭtsev 2007, Corollary 3]), et (4) et (6) reposent sur l'estimée de moments donnée dans la proposition 3.4. L'étape la plus délicate techniquement est l'étape (6), car la proposition 3.4 porte sur des blocs de variables consécutives, alors que l'ensemble $\mathcal{J}$ à contrôler est loin d'être connexe.

CHAPITRE IV

Étude statistique fine des applications Anosov lisses

Sommaire

1. Du cas dilatant au cas hyperbolique **43**
1.1. Applications dilatantes lisses 43
1.2. Mesures de Gibbs et codage pour les applications hyperboliques 45
2. Différents espaces candidats dans le cas hyperbolique **46**
2.1. Échec des fonctions lisses 47
2.2. Cas où les feuilletages sont lisses 47
2.3. Distributions pour les mesures SRB 49
2.4. Potentiels lisses et courants 51
3. Un espace de Banach convenable en général **53**

Le but de ce chapitre est de décrire comment on peut estimer très précisément les corrélations des applications Anosov de classe C^∞. Ce faisant, on introduira un formalisme qui permet de répondre très aisément à toute une classe de questions pour ces systèmes, tandis que l'approche classique reposant sur le codage donne des réponses partielles (et requiert des développements extrêmement techniques). On pourrait traiter le cas d'applications de régularité finie, hyperboliques sur des compacts invariants (appelés compacts hyperboliques localement maximaux dans la littérature). Par souci de simplicité, je ne considérerai ici que le cas où ce compact est la variété entière (i.e., le cas Anosov), et où l'application est C^∞. Les techniques développées dans ce chapitre, qui trouvent leur origine dans l'article [Blank, Keller et Liverani 2002], sont exposées dans [Gouëzel et Liverani 2006, 2008]. Mentionnons qu'une approche parallèle, plus analytique [Baladi et Tsujii 2007, 2008], permet d'obtenir des résultats plus fins (et optimaux) dans certaines situations.

1. Du cas dilatant au cas hyperbolique

1.1. Applications dilatantes lisses. On a décrit dans le chapitre I (en particulier dans les sections I.2 et I.3) les propriétés statistiques des applications uniformément dilatantes de l'intervalle. Les mêmes preuves fonctionnent pour des applications uniformément dilatantes en dimension plus grande, et donnent le résultat suivant (la définition d'une description matricielle a été donnée dans le paragraphe I.3) :

Théorème 1.1. *Soit $T : X \to X$ une application C^∞ sur une variété riemannienne compacte X, avec $\|DT(x)v\| \geqslant \lambda \|v\|$ uniformément en $x \in X$ et $v \in \mathcal{T}_x X$ (l'espace* *tangent à X en x), pour un certain $\lambda > 1$. Il existe alors une unique mesure de probabilité μ invariante absolument continue par rapport à la mesure de Lebesgue. Sa densité est C^∞. De plus, les corrélations de fonctions C^∞ admettent une description matricielle jusqu'à un terme d'erreur ε^n, pour tout $\varepsilon > 0$.*

La démonstration est la même que pour la proposition I.3.1. Alors que la partie sur la mesure invariante est classique, la description matricielle des corrélations est plus délicate et est due à RUELLE ; il l'a d'abord démontrée dans [RUELLE 1989] par une méthode de codage, avant d'en donner une autre démonstration sans partitions de Markov dans [RUELLE 1990]. L'idée de cette dernière preuve est de considérer l'opérateur de transfert (par rapport à la mesure de Lebesgue) donné par $\hat{T}u(x) = \sum_{T(y)=x} u(y)/\operatorname{Jac} DT(y)$, et de montrer que son rayon spectral essentiel sur l'espace C^k est borné par λ^{-k} (où λ est la dilatation de l'application).

Plus généralement, on peut s'intéresser aux mesures de Gibbs :

DÉFINITION 1.2. *Soient $T : X \to X$ une transformation sur un espace métrique et $\phi : X \to \mathbb{R}$ une fonction (appelée potentiel dans ce contexte). Une mesure de Gibbs pour le potentiel ϕ est une mesure de probabilité invariante μ vérifiant la propriété suivante. Il existe $P \in \mathbb{R}$ (appelé* pression *de ϕ) tel que, pour tout $\delta > 0$ assez petit, il existe C tel que, pour tout $x \in X$ et tout $n \in \mathbb{N}$,*

$$C^{-1} \leqslant \frac{\mu(B_n(x,\delta))}{e^{S_n\phi(x)-nP}} \leqslant C,$$

où $B_n(x,\delta)$ est la boule dynamique autour de x de paramètres n et δ, i.e.,

$$B_n(x,\delta) = \{y \in X \ : \ \forall 0 \leqslant i \leqslant n, d(T^i x, T^i y) \leqslant \delta\}.$$

Dans le cas dilatant, si δ est assez petit, la boule dynamique $B_n(x,\delta)$ est simplement la préimage de $B(T^n x, \delta)$ par la branche inverse de T^n qui envoie $T^n x$ sur x. Si le jacobien J_μ de T pour la mesure μ est assez régulier (par exemple hölderien), on en déduit que $\mu(B_n(x,\delta)) \asymp e^{-S_n(\log J_\mu)(x)}$. La mesure μ est alors une mesure de Gibbs pour le potentiel ϕ si et seulement si $S_n(\phi - P + \log J_\mu)$ est uniformément borné, ce qui équivaut au fait que $\phi - P + \log J_\mu$ soit un cobord. L'opérateur de transfert pour μ (qui s'écrit $\mathcal{L}u(x) = \sum u(y)/J_\mu(y)$) est donc conjugué à l'opérateur $\mathcal{L}_{\phi-P}u(x) = \sum_{T(y)=x} e^{\phi(y)-P}u(y)$.

Pour construire une mesure de Gibbs pour un potentiel donné ϕ, on procède dans l'autre sens : on part de l'opérateur $\mathcal{L}_\phi$ (qui ne dépend que de ϕ), on étudie ses propriétés spectrales (dans le cas dilatant, il a une unique valeur propre de module maximal, c'est un réel positif qu'on note e^P), et les fonctions propres correspondantes donnent la mesure de Gibbs (si h et ν sont respectivement la fonction propre de $\mathcal{L}_\phi$ et de son adjoint pour la valeur propre e^P, alors $\mu(u) = \nu(hu)$ définit une forme linéaire qui est en fait la mesure invariante recherchée). On vérifie comme dans le théorème 1.1 que le rayon spectral essentiel de $\mathcal{L}_\phi$ agissant sur C^k est borné par $e^P\lambda^{-k}$ si ϕ est C^k. Par conséquent, on obtient le résultat suivant également dû à RUELLE.

THÉORÈME 1.3. *Soient T une application uniformément dilatante de classe C^∞ sur une variété compacte X, et $\phi : X \to \mathbb{R}$ un potentiel C^∞. Il existe alors une unique mesure de Gibbs μ de potentiel ϕ, elle est ergodique et mélangeante. De plus, les* *corrélations de fonctions C^∞ admettent une description matricielle jusqu'à un terme d'erreur ε^n, pour tout $\varepsilon > 0$.*

Remarque 1.4. La mesure absolument continue est un cas particulier de mesure de Gibbs, pour le potentiel $\phi = -\log \operatorname{Jac}(T)$: cela résulte directement de la discussion qui précède. Ainsi, le théorème 1.1 est un cas particulier du théorème 1.3.

1.2. Mesures de Gibbs et codage pour les applications hyperboliques. Un difféomorphisme T d'une variété riemannienne compacte X est Anosov s'il existe une décomposition continue du fibré tangent en somme directe de deux sous-espaces $\mathcal{T}_x X = E^u(x) \oplus E^s(x)$, invariante par T, telle que $DT(x)$ dilate d'au moins λ dans la direction (instable) E^u et contracte d'au moins λ^{-1} dans la direction (stable) E^s (pour un certain $\lambda > 1$). Dans ce cas, il existe passant par chaque point x des variétés instables $W^u(x)$, tangentes à E^u (c'est l'ensemble des points dont l'orbite se rapproche de celle de x dans le passé), et des variétés stables $W^s(x)$ tangentes à E^s. Si T est C^∞ (ce qu'on supposera toujours dans le reste de ce chapitre), chaque variété $W^u(x)$ est C^∞, mais la dépendance transverse est seulement hölderienne (autrement dit, $y \mapsto E^u(y)$ est C^∞ le long de $W^u(x)$, mais seulement hölderienne globalement). On supposera aussi toujours que T est transitive (le fait que tous les Anosov soient transitifs est encore une question ouverte !)

Le cas dilatant correspond à $\dim E^s = 0$. Il est donc naturel de se poser dans ce cadre les mêmes questions que dans le cas dilatant. Le premier résultat dans cette direction est le suivant :

Théorème 1.5. *Si ϕ est un potentiel hölderien, il existe une unique mesure de Gibbs μ associée à ce potentiel. Elle est ergodique et mélangeante.*

Pour comprendre la notion de mesure de Gibbs dans le cas hyperbolique, il faut regarder la forme des boules dynamiques $B_n(x, \delta)$. Dans la direction stable, il s'agit simplement de supposer que $d(x, y) \leqslant \delta$, puisque la même condition sera alors automatique dans le futur. En revanche, dans la direction instable, il faut imposer une condition au temps n, comme dans le cas dilatant. Ainsi, si μ est une mesure invariante, $\mu(B_n(x, \delta))$ est essentiellement donné par $e^{-S_n(\log J^u_\mu)(x)}$, où J^u_μ est le jacobien de T pour la mesure μ dans la direction instable.

Remarque 1.6. En général, un Anosov n'a pas de mesure invariante absolument continue par rapport à la mesure de Lebesgue. Cette notion est remplacée dans ce cadre par celle de mesure SRB, i.e., une mesure qui est absolument continue par rapport à la mesure de Lebesgue le long des variétés instables. L'intérêt de cette définition est qu'une telle mesure est automatiquement une mesure physique, i.e., les moyennes ergodiques des fonctions continues convergent vers $\int f \, d\mu_{SRB}$ sur un ensemble de *mesure de Lebesgue* pleine. Intuitivement, la singularité de la mesure dans la direction stable ne joue pas de rôle dans l'asymptotique car cette direction est écrasée par la dynamique. Par définition, le jacobien instable de la mesure SRB est relié au jacobien $J^u(x)$ de la restriction de $DT(x)$ à $E^u(x)$ (où on munit $E^u(x)$ et $E^u(Tx)$ de la forme volume induite par la métrique riemannienne). Il suit que la mesure SRB est exactement la mesure de Gibbs pour le potentiel $-\log J^u(x)$. Notons que, comme $x \mapsto E^u(x)$ est seulement höldérien, ce potentiel n'est pas lisse en général.

La preuve classique du théorème 1.5, due à ANOSOV, SINAĬ, RUELLE et BOWEN, consiste à utiliser des partitions de Markov pour coder la dynamique : il existe un sous-décalage de type fini (X', T') et une projection surjective $\pi : X' \to X$ dont les fibres sont finies avec $T \circ \pi = \pi \circ T'$. On construit une mesure de Gibbs pour T' et le potentiel $\phi' = \phi \circ \pi$ en montrant que ϕ' est cohomologue à une fonction qui ne dépend que du futur, ce qui permet de se ramener à l'étude d'un système dilatant dans lequel les techniques d'opérateurs de transfert du paragraphe précédent s'appliquent bien. La projection de cette mesure de Gibbs par π donne la mesure de Gibbs recherchée pour T (voir par exemple [BOWEN 1975]).

Cette approche repose sur l'utilisation de partitions de Markov. Les bords d'une telle partition ne peuvent pas être réguliers en général (ce sont des réunions de variétés stables ou instables, ils sont höldériens mais pas mieux). Ainsi, on perd toutes les informations fines sur la régularité du système : on obtient les mêmes résultats qu'on parte d'une transformation C^2 (voire $C^{1+\varepsilon}$) ou C^∞. La description matricielle précise des corrélations est donc hors de portée de ces arguments de codage.

Certaines questions fines de régularité ou de différentiabilité ont reçu des réponses avec ce type de techniques. Mentionnons par exemple la question de la réponse linéaire : si $t \mapsto T_t$ est une famille lisse de difféomorphismes d'Anosov, de mesures SRB respectives μ_t, que peut-on dire de la régularité de $t \mapsto \int f \, \mathrm{d}\mu_t$ (où f est une fonction fixée), et peut-on donner des formules pour sa dérivée ? Il se trouve que cette fonction est lisse, et qu'on peut expliciter sa dérivée (voir par exemple [KATOK et al. 1989 ; RUELLE 1997, 2003 ; JIANG 2011]), mais au prix de difficultés techniques considérables (comme en témoigne le fait que les deux derniers articles cités sont des correctifs du précédent !). Les techniques que je vais exposer dans ce chapitre permettent, une fois en place, de répondre très simplement à ce genre de questions.

2. Différents espaces candidats dans le cas hyperbolique

La stratégie que je vais exposer consiste à trouver un bon espace de Banach sur lequel faire agir un opérateur de transfert. Notons que, pour les mesures de Gibbs générales, le choix de l'opérateur de transfert lui-même n'est pas évident : si l'on fait agir l'opérateur analogue du cas dilatant, donné par $\mathcal{L}_\phi u(x) = e^{\phi(y)} u(y)$ (où y est l'unique préimage de x par T), alors la mesure construite à partir de cet opérateur aurait un jacobien global J_μ tel que $\phi \sim -\log J_\mu$. Ce n'est pas ce qu'on cherche, on veut plutôt un jacobien *instable* comparable à ϕ.

En revanche, dans le cas SRB (et, disons, dans le cas où T préserve le volume pour simplifier), on sait déjà quel opérateur considérer : si on veut comprendre les corrélations $\int u \cdot v \circ T^n \, \mathrm{dLeb}$, il faut comprendre l'opérateur de composition par T ou, de manière équivalente, par T^{-1} (ce dernier choix est arbitraire, mais il est plus cohérent avec la définition habituelle des opérateurs de transfert). Il reste juste à trouver l'espace de Banach sur lequel le faire agir de telle sorte que son rayon spectral essentiel soit aussi petit que possible.

2.1. Échec des fonctions lisses. Comme dans le cas des applications dilatantes, le premier choix naturel à essayer est l'espace des fonctions lisses. Cependant, ce choix n'est guère judicieux, comme le montre la proposition suivante.

PROPOSITION 2.1. *Soit T un difféomorphisme C^∞ d'hyperbolicité $\lambda > 1$. Considérons l'opérateur $\mathcal{L} : u \mapsto u \circ T^{-1}$ agissant sur l'espace des fonctions C^k. Son rayon spectral essentiel est $\geqslant \lambda^k$.*

Comme $\lambda > 1$, le spectre devient donc de plus en plus mauvais lorsque l'on augmente k.

DÉMONSTRATION. Je vais donner une preuve complète de ce résultat car je n'en ai pas localisé dans la littérature. La différence avec le cas dilatant est la suivante : si v est un vecteur dans la direction stable en un point x, alors quand on développe la dérivée $k^{\text{ième}}$ $D^k(\mathcal{L}^j f)(x)(v, \dots, v)$ on obtient $D^k f(T^{-j}x)(DT^{-j}(x)v, \dots, DT^{-j}(x)v)$ et des termes d'ordre inférieur. Si v pointe dans la direction stable, la norme de $DT^{-j}(x)v$ croît au moins comme λ^j. Ainsi, la norme C^k de $\mathcal{L}^j f$ croît au moins comme λ^{kj} (à des termes d'ordre inférieur près), et ce à cause de tous les points x et tous les vecteurs v dans la direction stable, ce qui donne une infinité de raisons (linéairement indépendantes) à cette croissance. Cela ne pourrait se produire si le rayon spectral essentiel était $< \lambda^k$.

Pour formaliser ceci, considérons un point x et un vecteur $v \in E^s(x)$ de norme 1, et définissons une forme linéaire ν par $\nu(f) = D^k f(x)(v, \dots, v)$ (où la dérivée $k^{\text{ième}}$ est définie dans une carte fixée, sans importance). On a

$$\begin{aligned}
\left\| \mathcal{L}^j f \right\|_{C^k} &\geqslant \left| D^k(\mathcal{L}^j f)(T^j x) \left(\frac{DT^j(x)v}{\|DT^j(x)v\|}, \dots, \frac{DT^j(x)v}{\|DT^j(x)v\|} \right) \right| \\
&= \left\| DT^j(x)v \right\|^{-k} \left| D^k f(x)(v, \dots, v) \right| + O_j(\|f\|_{C^{k-1}}) \\
&\geqslant \lambda^{kj} |\nu(f)| + O_j(\|f\|_{C^{k-1}}).
\end{aligned}$$

Si le rayon spectral essentiel de $\mathcal{L}$ sur C^k était $< \rho < \lambda^k$, on pourrait décomposer C^k comme une somme directe $E \oplus F$ où E serait de dimension finie et $r(\mathcal{L}, F) < \rho$. Toutes les normes étant équivalentes sur E, la norme C^k y est équivalente à la norme C^{k-1}. On obtiendrait donc $\|\mathcal{L}^j f\|_{C^k} \leqslant C_j \|f\|_{C^{k-1}} + C\rho^j \|f\|_{C^k}$. Avec l'équation précédente, on trouve donc

$$\lambda^{kj} |\nu(f)| \leqslant C\rho^j \|f\|_{C^k} + C_j \|f\|_{C^{k-1}} .$$

Appliquons cette inégalité (avec j fixé) à une nouvelle fonction f_ε obtenue en zoomant sur f autour de x, de telle sorte que $\nu(f_\varepsilon)$ et $\|f_\varepsilon\|_{C^k}$ restent à peu près constants, tandis que $\|f_\varepsilon\|_{C^{k-1}}$ tend vers 0 (en coordonnées où $x = 0$ et f est supposée à support compact, on prend $f_\varepsilon(y) = \varepsilon^k f(y/\varepsilon)$). On obtient $\lambda^{kj}|\nu(f)| \leqslant C\rho^j \|f\|_{C^k}$. Comme $\rho < \lambda^k$, on peut faire tendre j vers l'infini pour trouver $\nu(f) = 0$, ce qui est absurde pour une fonction f générale. □

2.2. Cas où les feuilletages sont lisses. La première approche naturelle pour résoudre le problème évoqué dans le paragraphe précédent est de traiter séparément les directions stable et instable : comme dans le cas dilatant, on va utiliser une norme de type C^k dans la direction instable, mais on va plutôt utiliser une norme duale dans la direction stable, i.e., une norme distributionnelle. Cette idée naturelle est assez délicate à mettre en œuvre en général à cause du manque de régularité des feuilletages stable et instable (comme on le verra dans les paragraphes suivants). On va ici considérer le cas où les feuilletages stable et instable sont lisses, ce qui fait disparaître la plupart des difficultés. Ce cadre a été mis en avant dans [BALADI 2005], où elle a montré que des techniques d'opérateurs pseudo-différentiels pouvaient s'appliquer. Dans ce paragraphe, T sera donc une application hyperbolique C^∞ qui préserve le volume, et dont les directions stable et instable dépendent du point de manière C^∞.

Définissons une seminorme (correspondant à k dérivées dans la direction instable, et au dual de ℓ dérivées dans la direction stable), par

$$A_{k,\ell}(f) = \sup_{x\in X} \sup_{v_1,\dots,v_k} \sup_{\varphi} \int_{W^s_\varepsilon(x)} \varphi \cdot L_{v_1} \cdots L_{v_k} f \,\mathrm{dLeb} \tag{2.1}$$

où l'intégrale est prise sur la variété stable locale de x, la fonction φ est une fonction à support compact sur cet ensemble, de norme C^ℓ au plus 1, et les v_i sont des champs de vecteurs définis sur un voisinage de $W^s_\varepsilon(x)$, pointant partout dans la direction instable, et de norme C^k bornée par 1. La notation L_{v_i} désigne la dérivée dans la direction du vecteur v_i.

On a tout fait pour que cette seminorme se comporte bien sous l'itération de $\mathcal{L}$. C'est effectivement le cas :

LEMME 2.2. *On a*

$$A_{k,\ell}(\mathcal{L}^n f) \leqslant C\lambda^{-kn} A_{k,\ell}(f) + C(n) A_{k-1,\ell}(f). \tag{2.2}$$

DÉMONSTRATION. Il s'agit d'estimer des intégrales de la forme

$$\int_{W^s_\varepsilon(x)} \varphi \cdot L_{v_1} \cdots L_{v_k} (f\circ T^{-n}) \,\mathrm{dLeb} = \int_{T^{-n}(W^s_\varepsilon(x))} J^{(n)}_u \cdot \varphi\circ T^n L_{w_1}\cdots L_{w_k} f \,\mathrm{dLeb}, \tag{2.3}$$

où $J^{(n)}_u$ est le jacobien instable de T^n et $w_i = (T^n)^* v_i$. Les champs de vecteurs v_i pointant dans la direction instable, ils sont contractés, donc $|w_i| \leqslant \lambda^{-n}$. De plus, la dilatation le long de W^s assure que w_i est très régulier le long de $T^{-n}(W^s_\varepsilon(x))$. En revanche, dans la direction transverse, son comportement est très mauvais : sa norme C^k est seulement bornée par $C\Lambda^{kn}$ avec $\Lambda > 1$! Fixons un champ de vecteur w'_i qui coïncide avec w_i le long de $T^{-n}(W^s_\varepsilon(x))$ et de norme C^k bornée par $C\lambda^{-n}$. Comme une dérivée de Lie $L_v u$ ne dépend que de la valeur du vecteur v et pas de ses dérivées, on peut sans dommage remplacer w_1 par w'_1 dans l'équation précédente. Il n'en va pas de même de w_2 puisqu'il est soumis à la dérivée $L_{w'_1}$. On écrit alors $L_{w'_1} L_{w_2} = L_{w_2} L_{w'_1} + L_{[w'_1, w_2]}$. Le terme venant du crochet de Lie donne une intégrale impliquant $k-1$ dérivées de Lie de f contre des champs de vecteurs C^{k-1}, et est donc borné par $C(n)A_{k-1,\ell}(f)$ (avec une constante $C(n)$ exponentiellement grande), tandis que dans le terme $L_{w_2} L_{w'_1}$ on peut à nouveau sans dommage remplacer w_2 par w'_1. Recommençant le processus, on remplace tous les w_i par des w'_i. Comme $\prod \|w'_i\|_{C^k} \leqslant \lambda^{-kn}$, le résultat suit puisque la norme C^ℓ de $J^{(n)}_u$ est facilement contrôlée par un argument de distorsion. $\square$

Le lemme précédent permet d'obtenir de la contraction en ce qui concerne les termes $A_{k,\ell}$ avec k grand (modulo des termes d'ordre inférieur, dont on va aisément se débarrasser en utilisant des normes pondérées). En revanche, il ne donne rien pour les termes avec k petits, mais pour ceux-ci on peut utiliser la régularité de φ : en approchant φ par une fonction de classe $C^{\ell+1}$ par convolution et en utilisant la contraction le long de W^s, on obtient

$$A_{k,\ell}(\mathcal{L}^n f) \leqslant C\lambda^{-(k+\ell)n} A_{k,\ell}(f) + C(n)A_{k-1,\ell}(f) + C'(n)A_{k,\ell+1}(f), \tag{2.4}$$

où $C(n)$ et $C'(n)$ croissent encore une fois exponentiellement vite.

Définissons finalement une norme sur les fonctions C^∞ par

$$\|f\|_{K,\ell} = \sum_{k=0}^{K} A_{k,\ell}(f).$$

Pour $a > 0$, on définit une norme équivalente $\|f\|_{K,\ell,a} = \sum_{k=0}^{K} a^k A_{k,\ell}(f)$. Pour tout n, les équations (2.2) (pour $k = K$) et (2.4) (pour $k < K$) assurent que, si a est assez grand (pour éliminer le facteur $C(n)$ dans (2.2) et (2.4)),

$$\|\mathcal{L}^n f\|_{K,\ell,a} \leqslant C\lambda^{-\min(K,\ell)n} \|f\|_{K,\ell,a} + C(n,a) \|f\|_{K-1,\ell+1} . \tag{2.5}$$

Notons que le choix de a dépend de n. On vérifie ensuite que la boule unité pour $\|\cdot\|_{K,\ell}$ (ou pour $\|\cdot\|_{K,\ell,a}$ qui lui est équivalente) est compacte pour $\|\cdot\|_{K-1,\ell+1}$, ce qui n'est pas surprenant puisqu'on diminue la régularité dans toutes les directions : on passe moralement de C^K dans la direction instable et $C^{-\ell}$ dans la direction stable à C^{K-1} et $C^{-\ell-1}$. L'inégalité (2.5) est donc une inégalité DFLY (cette notion est définie page 14), ce qui permet d'appliquer le lemme I.2.6 de HENNION. On en déduit que le rayon spectral essentiel de $\mathcal{L}$ pour la norme $\|\cdot\|_{K,\ell,a}$ est au plus $C^{1/n}\lambda^{-\min(K,\ell)}$. Comme cette norme est équivalente à $\|\cdot\|_{K,\ell}$, la même estimée est valide pour cette dernière. En faisant tendre n vers l'infini, on obtient donc le théorème suivant :

THÉORÈME 2.3. *Le rayon spectral essentiel de $\mathcal{L}$ sur l'espace de Banach obtenu en complétant C^K pour la norme $\|\cdot\|_{K,\ell}$ est borné par $\lambda^{-\min(K,\ell)}$.*

En prenant K et ℓ assez grands, on peut rendre le rayon spectral essentiel arbitrairement petit. On en déduit immédiatement que les corrélations de fonctions C^∞ admettent une description matricielle avec une précision arbitraire.

2.3. Distributions pour les mesures SRB. Nous supposons dorénavant que T est un Anosov C^∞ qui préserve le volume, mais sans hypothèse de régularité supplémentaire sur les directions stable et instable : celles-ci seront uniquement hölderiennes en général. Les résultats du paragraphe précédent restent néanmoins valables : il existe un espace de Banach sur lequel le rayon spectral essentiel de $\mathcal{L}$ (l'opérateur de composition par T^{-1}) est arbitrairement petit. On en déduit à nouveau une description matricielle précise des corrélations.

Si l'on essaye d'utiliser l'espace décrit dans le paragraphe précédent, deux problèmes se posent :

(1) La direction instable n'étant pas lisse, il n'existe en général pas de champ de vecteurs non nul de classe C^k pointant dans la direction instable.

(2) Pour obtenir la compacité de $\|\cdot\|_{K,\ell}$ par rapport à $\|\cdot\|_{K-1,\ell+1}$, il faut être capable de passer d'une variété stable à une autre par l'intermédiaire d'une famille lisse de variétés stables. Ce n'est plus possible en général si les variétés stables ne dépendent pas du point de manière lisse.

Pour résoudre ces problèmes, l'idée la plus naturelle est de remplacer les objets pointant dans la direction stable ou la direction instable par des objets pointant approximativement dans ces directions. Plus précisément, ils devront pointer respectivement dans les cônes stable et instable.

Cette idée naïve fonctionne très bien pour résoudre le second problème. En effet, dans la définition (2.1) des coefficients $A_{k,\ell}(f)$, on peut remplacer l'intégrale le long d'une variété stable par l'intégrale le long d'un petit morceau de sous-variété pointant partout dans le cône stable (qu'on appellera variété quasi-stable). Comme le cône stable est invariant par T^{-1}, les préimages d'une telle sous-variété pointent encore partout dans celui-ci, ce qui permet encore d'utiliser le changement de variables (2.3) pour estimer $A_{k,\ell}(f)$. De plus, on peut aisément interpoler de manière lisse entre deux variétés quasi-stables (par exemple par interpolation linéaire dans une carte suffisamment petite). Le second problème évoqué ci-dessus est donc résolu.

Le premier problème est plus délicat. En effet, si v pointe dans une direction proche de la direction instable mais légèrement différente, alors $(T^n)^*v$ s'aplatit le long de la direction stable. Autrement dit, si l'on est prêt à accepter des champs de vecteurs dans des directions pas exactement instables, on est obligé d'accepter des champs de vecteurs pointant dans toutes les directions. Cette différence entre les deux problèmes provient du fait que la préimage du cône instable *contient* le cône instable, tandis que la préimage du cône stable est contenue dans celui-ci.

Définissons donc un nouveau coefficient $B_{k,\ell}(f)$ comme $A_{k,\ell}$, mais en autorisant des intégrales le long de variétés quasi-stables, et des champs de vecteurs pointant dans toutes les directions. On aura besoin d'utiliser des champs de vecteurs C^m, pas seulement C^k, notons donc $B_{k,\ell,m}(f)$ le coefficient correspondant. Cette définition semble incompatible avec l'utilisation de la dilatation dans la direction instable : on devrait perdre tout phénomène de contraction du champ de vecteurs, et donc tout espoir de quasi-compacité. Miraculeusement, ce n'est pas le cas, comme le montre l'estimée suivante, analogue dans ce cadre du lemme 2.2 :

Lemme 2.4. *Si $m \geqslant \ell$, on a*

$$B_{k,\ell,m}(\mathcal{L}^n f) \leqslant C\lambda^{-kn} B_{k,\ell,m}(f) + C(n) B_{k-1,\ell-1,m-1}(f).$$

La différence cruciale avec le lemme 2.2 est que, dans le terme d'erreur, on perd aussi en régularité sur la fonction test, qui passe de C^ℓ à $C^{\ell-1}$. Autrement dit, on voit apparaître une interaction entre les directions stable et instable qui était absente dans le paragraphe précédent.

Démonstration. Pour démontrer le lemme, dans l'expression $L_{v_1} \cdots L_{v_k}(f \circ T^{-n})$ considérons par exemple le champ de vecteurs v_1. On le décompose en $v_1^s + v_1^u$ où

v_1^s est tangent à la variété quasi-stable W sur laquelle on intègre, tandis que v_1^u est suffisamment proche de la direction instable pour être contracté de λ^{-n} par T^{-n}. On remarque ensuite que $\int_W \varphi \cdot L_{v_1^s}\psi = -\int_W L_{v_1^s}\varphi \cdot \psi$ (à un terme C^{m-1} de divergence près). Autrement dit, une intégration par parties permet de faire porter la dérivée de Lie sur la fonction test, ce qui laisse moins de dérivées sur f et fait perdre de la régularité à la fonction test (de C^ℓ, elle devient $C^{\min(\ell-1,m-1)}$). Pour un autre champ de vecteurs v_i, on procède de même en le faisant tout d'abord commuter avec les autres v_j $(j<i)$ pour le mettre en première position (ce qui fait apparaître comme dans le paragraphe précédent des crochets de Lie, contrôlés en termes de $B_{k-1,\ell,m-1}(f)$) puis en le décomposant en $v_i^u + v_i^s$ et en éliminant v_i^s par une intégration par parties.

Le reste de l'argument est analogue à la preuve du lemme 2.2. □

Dans ce cadre, on a également un analogue de l'estimée (2.4), donné lorsque $m \geqslant \ell$ par

$$B_{k,\ell,m}(\mathcal{L}^n f) \leqslant C\lambda^{-(k+\ell)} B_{k,\ell,m}(f) + C(n)(B_{k-1,\ell-1,m-1}(f) + B_{k,\ell+1,m}(f)). \tag{2.6}$$

Le terme d'erreur $B_{k-1,\ell-1,m-1}(f)$ provient, comme dans le lemme 2.4, de l'élimination des composantes des champs de vecteurs dans la direction stable, tandis que $B_{k,\ell+1,m}(f)$ correspond à la régularisation de la fonction test.

Finalement, la norme $\|f\|'_{K,\ell} = \sum_{k=0}^K B_{k,k+\ell,k+\ell}(f)$ se comporte bien sous l'action de l'opérateur de transfert. Notons qu'on a besoin d'augmenter la régularité de la fonction test et des champs de vecteurs en fonction de leur nombre, pour pouvoir utiliser le lemme 2.4 et l'inégalité (2.6). On a compacité de la boule unité de $\|\cdot\|'_{K,\ell}$ pour la norme $\|\cdot\|'_{K-1,\ell+1}$, et une inégalité DFLY pour une norme pondérée équivalente à $\|\cdot\|'_{K,\ell}$. On en déduit le résultat suivant (démontré dans [Gouëzel et Liverani 2006]) :

Théorème 2.5. *Le rayon spectral essentiel de $\mathcal{L}$ sur l'espace de Banach obtenu en complétant C^K pour la norme $\|\cdot\|'_{K,\ell}$ est borné par $\lambda^{-\min(K,\ell)}$.*

Remarque 2.6. Si l'on travaillait en différentiabilité finie, on aurait besoin que l'application soit $C^{K+\ell+1}$ pour pouvoir travailler avec la norme $\|\cdot\|'_{K,\ell}$ (car cette norme fait intervenir des champs de vecteurs $C^{K+\ell}$, qu'il faut pouvoir pousser par la dynamique). Autrement dit, si T est C^{r+1}, on ne peut prendre que $K=\ell=r/2$, i.e., le rayon spectral essentiel ne profite que de la moitié de la différentiabilité. Ce phénomène avait déjà été remarqué dans [Kitaev 1999]. Il faut noter que, dans le cas où le feuilletage instable est lisse, on peut prendre $K=\ell=r$ dans le théorème 2.3, et cette perte de moitié n'a donc pas lieu.

2.4. Potentiels lisses et courants. Il ressort du paragraphe précédent que des objets exploitant l'hyperbolicité doivent essentiellement pouvoir être intégrés sur des variétés stables (ou quasi-stables), et dérivés dans toutes les directions. Dans le paragraphe précédent, on a utilisé des fonctions lisses, et on a les a intégrées par rapport à la mesure riemannienne. Cependant, il est plus naturel de considérer directement des formes différentielles de degré d_s (la dimension des variétés stables), qui sont plus intrinsèques (le choix d'une métrique riemannienne n'importe plus). On va voir dans ce paragraphe comment cette idée permet de traiter les mesures de Gibbs pour certaines classes de potentiel. On supposera ici que T est une application Anosov C^∞, mais qui ne préserve plus nécessairement le volume.

Si ω est une forme différentielle lisse de degré d_s, définissons donc un coefficient

$$C_{k,\ell,m}(\omega) = \sup \int_W \varphi \cdot L_{v_1} \cdots L_{v_k} \omega, \tag{2.7}$$

où le suprémum porte sur toutes les variétés quasi-stables (de diamètre fixé) W, toutes les fonctions test φ sur W de norme C^ℓ au plus 1 et tous les champs de vecteurs $v_1, \dots, v_k$ définis sur un voisinage de W et de norme C^m au plus 1.

Posons $\|\omega\|_{K,\ell} = \sum_{k=0}^K C_{k,k+\ell,k+\ell}(\omega)$, comme dans le paragraphe précédent. Cette norme devrait bien se comporter sous l'itération de T. Notons $\mathcal{C}^{K,\ell}$ la complétion de l'ensemble des formes différentielles C^∞ pour cette norme. Si ω' est une forme de degré d_u, la forme linéaire $\omega \mapsto \int_X \omega \wedge \omega'$ est bien définie sur les formes différentielles lisses, et on vérifie aisément qu'elle est continue pour la norme $\|\cdot\|_{K,\ell}$. Ainsi, elle s'étend continûment à $\mathcal{C}^{K,\ell}$, ce qui permet de définir $\langle \omega, \omega' \rangle$ pour tout $\omega \in \mathcal{C}^{K,\ell}$ et tout ω' forme différentielle lisse de degré d_u. Autrement dit, les éléments de $\mathcal{C}^{K,\ell}$ vivent dans le dual des formes différentielles lisses, ce sont des *courants*.

Si ϕ est une fonction lisse, on peut définir un opérateur $\mathcal{L}_\phi$ agissant sur l'espace des d_s-formes par $\mathcal{L}_\phi \omega = T_*(e^\phi \omega)$. En reprenant les arguments du paragraphe précédent, on obtient que $\mathcal{L}_\phi$ a un trou spectral sur $\mathcal{C}^{K,\ell}$. Une étude soignée des fonctions propres de cet opérateur (nettement plus délicate que dans le cas dilatant ou que dans le cas précédent où on étudiait la mesure SRB) permet de démontrer le théorème suivant.

THÉORÈME 2.7. *Le rayon spectral de $\mathcal{L}_\phi$ sur $\mathcal{C}^{K,\ell}$ est e^P, où P est la pression du potentiel ϕ, tandis que son rayon spectral essentiel est au plus $\lambda^{-\min(K,\ell)} e^P$. De plus, $\mathcal{L}_\phi$ a une unique valeur propre de module maximal, en e^P, et celle-ci est simple.*

Notons $\omega_0 \in \mathcal{C}^{K,\ell}$ la fonction propre de $\mathcal{L}_\phi$ pour la valeur propre e^P (c'est un courant de degré d_s). L'adjoint de $\mathcal{L}_\phi$ agit sur $(\mathcal{C}^{K,\ell})'$ (qui est encore un espace de courants puisqu'on peut intégrer un élément de cet espace contre tout élément de $\mathcal{C}^{K,\ell}$, et en particulier contre toute forme différentielle lisse), avec le même spectre. En particulier, e^P est une valeur propre simple, et on peut considérer la fonction propre correspondante ν_0.

Les courants ω_0 et ν_0 sont de degrés complémentaires. Si on pouvait construire leur produit extérieur $\omega_0 \wedge \nu_0$, ce serait une mesure, nécessairement invariante. En général, le produit extérieur de courants ne peut pas toujours être défini, mais cette construction est ici triviale : définissons une forme linéaire sur les fonctions lisses par $\mu(f) = \langle \nu_0, f\omega_0 \rangle$ (où l'accouplement est bien défini puisque $\nu_0 \in (\mathcal{C}^{K,\ell})^*$ et $f\omega_0 \in \mathcal{C}^{K,\ell}$), on vérifie que c'est une mesure.

THÉORÈME 2.8. *Cette mesure μ est la mesure de Gibbs associée au potentiel ϕ.*

En un sens, ce résultat n'est pas surprenant : on a exhibé la mesure invariante en utilisant des objets dont la variation dans les directions stable et instable est mesurée par $e^{\pm\phi}$, ce qui implique que le jacobien instable de μ est comparable à $e^{-\phi}$. C'est exactement la propriété qui caractérise les mesures de Gibbs.

Des deux théorèmes précédents, on déduit que les corrélations des fonctions lisses pour la mesure de Gibbs admettent une description matricielle avec une précision arbitraire.

REMARQUE 2.9. En prenant $\phi = 1$, on obtient en particulier une construction de la mesure d'entropie maximale comme produit extérieur de deux courants de degrés respectifs d_s et d_u. Ce résultat est à rapprocher de la construction des mesures d'entropie maximale en dynamique complexe (par exemple pour les applications de HÉNON), qui se fait de même en construisant d'abord des courants (appelés courants de GREEN) et en prenant ensuite leur produit extérieur. Il est à noter que, dans ce cadre de dynamique complexe, il faut des arguments sophistiqués de théorie du potentiel pour montrer que ce produit extérieur est bien défini, alors que c'est essentiellement trivial dans notre contexte réel (mais uniformément hyperbolique).

3. Un espace de Banach convenable en général

Les constructions des paragraphes 2.3 et 2.4 ont des champs d'application différents : la première construction (basée sur des distributions) fonctionne pour la mesure de Lebesgue (correspondant au potentiel non lisse $-\log J^u$), tandis que la seconde (basée sur des courants) ne fonctionne que pour des potentiels lisses. Cependant, les arguments techniques sont très semblables, et les conclusions (trou spectral, description matricielle des corrélations) sont également similaires. Dans cette section, je vais décrire une théorie qui englobe ces deux cadres.

La remarque cruciale est que le potentiel $-\log J^u$ n'est pas lisse, mais qu'il provient d'une fonction lisse sur une grassmannienne, évaluée sur la direction instable. On utilisera plutôt la direction stable dans ce qui suit, mais cela ne change rien. Notons $\mathcal{G}^s$ la grassmannienne des sous-espace de $\mathcal{T}X$ de dimension d_s. Le théorème suivant englobe essentiellement les résultats des paragraphes 2.3 et 2.4.

THÉORÈME 3.1. *Soit $\overline{\phi}$ une fonction C^∞ de $\mathcal{G}^s$ dans $\mathbb{R}$, définissons une fonction höldérienne ϕ par $\phi(x) = \overline{\phi}(x, E^s(x))$. Alors les corrélations des fonctions C^∞ pour la mesure de Gibbs de potentiel ϕ admettent une description matricielle avec une précision arbitraire.*

La seule difficulté pour démontrer ce théorème, par rapport au paragraphe 2.4, est de trouver le bon cadre conceptuel dans lequel travailler, car toutes les idées techniques ont déjà été exposées. On a envie d'utiliser des espaces de courants, mais si ω est une forme différentielle (ou un courant), alors $e^{\overline{\phi}}\omega$ n'en est plus une car elle perd la structure alternée si $\overline{\phi}$ est arbitraire. Cependant, cette structure ne jouait pas de rôle réel dans les arguments du paragraphe 2.4, on va donc pouvoir s'en sortir en l'abandonnant.

Notons $\mathcal{D}$ l'ensemble des objets C^∞ qui, à un point $x \in X$ et à un sous-espace E de dimension d_s de $\mathcal{T}_xX$, associent une forme volume sur E (c'est l'ensemble des sections d'un fibré vectoriel de dimension 1 au-dessus de $\mathcal{G}^s$). Les formes différentielles de degré d_s définissent évidemment des éléments de $\mathcal{D}$. Un élément de $\mathcal{D}$ peut être multiplié par une fonction lisse sur $\mathcal{G}^s$ (en particulier par $e^{\overline{\phi}}$), et poussé en avant par T. Par conséquent, on peut définir un opérateur $\mathcal{L}_{\overline{\phi}}(\omega) = T_*(e^{\overline{\phi}}\omega)$ sur $\mathcal{D}$. Il joue le rôle des opérateurs de transfert des parties précédentes.

Pour poursuivre, il faut définir l'analogue des coefficients (2.7) pour $\omega \in \mathcal{D}$, i.e., il faut pouvoir définir $\int_W \varphi \cdot L_{v_1} \cdots L_{v_k}\omega$. Remarquons tout d'abord qu'un élément de $\mathcal{D}$, induisant une forme volume sur chaque espace tangent à W, peut être intégré le long de W. Autrement dit, $\int_W \varphi \cdot \omega$ est canoniquement défini. Il reste à traiter la dérivée de Lie $L_v\omega$. Si v est un champ de vecteurs sur X, il définit un flot ψ_t. Un élément de $\mathcal{D}$ peut être poussé par des difféomorphismes, et on peut donc définir $L_v\omega = \mathrm{d}\psi_t^*\omega / \mathrm{d}t_{|t=0}$. Cela correspond à dériver ω par rapport à un certain champ de vecteurs $\tilde{v}$ qui relève v à l'espace des grassmanniennes.

Pour pouvoir utiliser les arguments du paragraphe 2.4, il reste juste à vérifier que certains calculs formels restent vrais dans notre nouveau cadre. En particulier, il faut vérifier la formule d'intégration par parties $\int_W \varphi \cdot L_v\omega = -\int_W L_v\varphi \cdot \omega$ si v est tangent à W, et la formule de commutation $L_vL_w\omega - L_wL_v\omega = L_{[v,w]}\omega$. La première formule vient du fait que le flot de v conserve W, si bien que tout se passe le long de W où on peut considérer ω comme une forme volume normale. La seconde formule revient à vérifier que $[\tilde{v}, \tilde{w}] = \widetilde{[v, w]}$. Cette propriété, qui peut se montrer par un calcul élémentaire, est en fait vraie pour tout relèvement « naturel » de champs de vecteurs [Kolář, Michor et Slovák 1993, Lemma 6.19].

Une fois ceci fait, on peut définir des coefficients $D_{k,\ell,m}(\omega)$ pour $\omega \in \mathcal{D}$ comme en (2.7), puis une norme $\|\omega\|_{K,\ell}$, et définir $\mathcal{D}^{K,\ell}$ comme la complétion de $\mathcal{D}$ pour celle-ci. Les théorèmes 2.7 et 2.8 restent vrais dans ce cadre pour l'opérateur $\mathcal{L}_{\overline{\phi}}$, et impliquent le théorème 3.1.

Remarque 3.2. Un élément du dual de $\mathcal{D}^{K,\ell}$ peut s'intégrer contre n'importe quel élément de $\mathcal{D}^{K,\ell}$, et en particulier contre toute forme différentielle lisse, c'est donc un courant. La fonction propre ν_0 dominante de l'adjoint de $\mathcal{L}_{\overline{\phi}}$ est donc un courant. On peut l'expliciter comme suit : Margulis a construit une famille de mesures μ^u portées par les variétés stables et telles que $\mu^u = T_*(e^{P-\phi}\mu^u)$. On vérifie alors que, localement, si F est une petite variété instable,

$$\langle \nu_0, \omega \rangle = \int_{x\in F} \left(\int_{y \in W^s(x)} e^{\sum_{k=0}^{\infty} \phi(T^k y) - \phi(T^k x)} \omega \right) \mathrm{d}\mu_F^u(x).$$

En particulier, ν_0 a une structure laminaire, portée par les variétés stables et de mesure transverse μ^u. En revanche, la fonction propre dominante de $\mathcal{L}_{\overline{\phi}}$ ne peut pas s'interpréter comme un courant en général, contrairement à la situation du paragraphe 2.4.

Remarque 3.3. L'espace $\mathcal{D}^{K,\ell}$, qui fonctionne pour l'application T, fonctionne également pour des applications $\tilde{T}$ proches de T, puisque des variétés quasi-stables pour T le sont encore pour $\tilde{T}$ (c'est un autre intérêt d'utiliser des variétés quasi-stables dans la définition de l'espace de Banach, et pas uniquement les variétés stables). On peut donc changer la transformation, le potentiel, et obtenir une famille d'opérateurs de transfert $\mathcal{L}_t$ agissant tous sur $\mathcal{D}^{K,\ell}$. Si cette famille dépendait de manière lisse de t, on en déduirait immédiatement que toutes les données spectrales (et en particulier la mesure de Gibbs) dépendraient de manière lisse de t. Ce n'est pas le cas en général : les opérateurs de composition ne dépendent jamais de manière lisse de t, par exemple $U_t f(x) = f(x+t)$, agissant sur $C^1(\mathbb{R}/\mathbb{Z})$, ne définit pas une famille dépendant continûment de t. En revanche, dans cet exemple, on a une famille lisse d'opérateurs de C^1 dans C^0. Dans le cas des opérateurs de transfert $\mathcal{L}_t$, la famille $t \mapsto \mathcal{L}_t$ est continue comme opérateurs de $\mathcal{D}^{K,\ell}$ dans $\mathcal{D}^{K-1,\ell+1}$, et elle admet même des développements de Taylor d'ordre a si on la considère comme une famille d'opérateurs de $\mathcal{D}^{K,\ell}$ dans $\mathcal{D}^{K-a-1,\ell+1+a}$. Cela suffit pour appliquer des arguments de perturbation développés dans [Gouëzel et Liverani 2006], et pour en déduire que toutes les données spectrales sont lisses (si on fait un peu attention aux espaces de départ et d'arrivée). En particulier, si on note μ_t la mesure SRB de T_t (où $t \mapsto T_t$ est une famille lisse de difféomorphismes Anosov), alors $t \mapsto \int f \, \mathrm{d}\mu_t$ est C^∞ si f l'est (et ce résultat est essentiellement trivial une fois que le formalisme que je viens de décrire est mis en place).

CHAPITRE V

Spectre du laplacien sur l'espace des surfaces plates

Sommaire

1. Définitions et résultat principal **57**
1.1. Description de la preuve 59
2. Espaces de Banach distributionnels pour le flot de Teichmüller **60**
2.1. L'espace des modules des surfaces de translation 60
2.2. Hyperbolicité et récurrence 62
2.3. L'espace de Banach $\mathcal{B}$ 64
3. Relation entre informations géométriques et algébriques sur l'action de $\mathrm{SL}(2,\mathbb{R})$ **66**
3.1. Représentations unitaires de $\mathrm{SL}(2,\mathbb{R})$ 66
3.2. Démonstration du théorème principal 69

Dans ce chapitre, je vais présenter les résultats de l'article [Avila et Gouëzel 2010] sur le spectre de l'action de $\mathrm{SL}(2,\mathbb{R})$ sur l'espace des modules des surfaces plates. Nous y démontrons que le spectre du laplacien, dans l'intervalle $[0,1/4]$, est constitué de valeurs propres isolées de multiplicité finie, comme dans le cas des surfaces hyperboliques de volume fini. Ce résultat appelle un certain nombre de définitions et de commentaires, que je vais exposer avant de donner quelques idées centrales dans la preuve (qui est à la fois dynamique et algébrique).

1. Définitions et résultat principal

Il existe énormément d'actions de $\mathbb{Z}$ ou $\mathbb{R}$ préservant une mesure de probabilité, puisque toute action de ces groupes par homéomorphismes sur un espace compact admet des mesures invariantes. Si l'on cherche à déterminer si deux tels systèmes sont isomorphes ou non, la première idée est de regarder s'ils sont spectralement isomorphes, i.e., si les actions par composition sur L^2 sont isomorphes (ce qui est en général bien plus faible qu'un isomorphisme des dynamiques). Tous les systèmes présentant suffisamment d'hyperbolicité sont spectralement isomorphes (ils ont un spectre de Lebesgue avec multiplicité infinie), et il faut utiliser des invariants plus subtils comme l'entropie pour les distinguer.

La situation est radicalement différente si l'on considère des groupes plus gros, comme $\mathrm{SL}(2,\mathbb{R})$: ce groupe n'étant pas moyennable, une action typique (comme l'action de $\mathrm{SL}(2,\mathbb{R})$ par homographies sur S^1) n'a pas de mesure de probabilité invariante. Il est donc beaucoup plus difficile d'exhiber des exemples d'actions préservant une mesure de probabilité : en général, ces exemples proviennent de situations géométriques ou algébriques et sont à la fois riches et intéressants. Pour ces actions, la question de l'équivalence spectrale (qui revient à comprendre comment $L^2(X)$ se décompose en somme directe ou intégrale directe de représentations irréductibles de $\mathrm{SL}(2,\mathbb{R})$) est hautement non triviale. Elle équivaut plus ou moins à comprendre le spectre du laplacien (un opérateur non borné agissant sur $L^2(X)$ et provenant de l'action de $\mathrm{SL}(2,\mathbb{R})$), comme je l'expliquerai dans le paragraphe 3.1. Notons que la structure algébrique de $\mathrm{SL}(2,\mathbb{R})$ implique automatiquement certaines propriétés ergodiques de l'action. Par exemple, si l'action est ergodique, alors elle est automatiquement mélangeante (théorème de Howe-Moore).

Le premier exemple naturel est certainement un groupe G dont $\mathrm{SL}(2,\mathbb{R})$ est un sous-groupe, ainsi qu'un réseau Γ de G. En ce cas, $\mathrm{SL}(2,\mathbb{R})$ agit par multiplication à gauche sur $X = G/\Gamma$, en préservant la mesure de Haar. La structure homogène de X donne un certain nombre d'outils permettant l'étude de cette action. On peut en particulier prendre $G = \mathrm{SL}(2,\mathbb{R})$ lui-même : X est alors le fibré unitaire tangent à une surface hyperbolique M, et l'action du sous-groupe diagonal de $\mathrm{SL}(2,\mathbb{R})$ correspond au flot géodésique. En ce cas, le spectre du laplacien est bien compris (même si des conjectures importantes restent ouvertes) : dans le cas où M est compacte, le spectre du laplacien est discret avec multiplicités finies. En particulier, $L^2(X)$ se décompose comme une somme directe de représentations irréductibles (avec des paramètres spectraux correspondant aux valeurs propres du laplacien). En revanche, lorsque M est non compacte (mais encore de volume fini), le spectre du laplacien est discret dans $[0,1/4]$ mais il est continu dans $[1/4,+\infty[$. L'intervalle $[0,1/4]$ a ici une signification profonde puisqu'il correspond aux représentations de la série complémentaire, celles qui n'apparaissent pas dans la représentation régulière gauche de $\mathrm{SL}(2,\mathbb{R})$ et qu'on peut considérer comme étant les plus pathologiques (voir le paragraphe 3.1).

Dans ce chapitre, on va s'intéresser à une généralisation non homogène de cette situation. Plus précisément, on peut identifier $\mathrm{SL}(2,\mathbb{R})/\mathrm{SL}(2,\mathbb{Z})$ avec l'espace des modules des tores plats. On peut faire la même chose en augmentant le genre, i.e., travailler sur un espace de modules de surfaces plates de genre plus grand (avec singularités). Cet espace admet une action naturelle de $\mathrm{SL}(2,\mathbb{R})$, pour laquelle il existe une mesure invariante de support total, appelée mesure de Masur-Veech (et d'autres mesures invariantes, supportées par des sous-variétés affines). Nous définirons précisément ces objets dans la partie 2.1. Pour l'instant, contentons-nous de dire qu'un tel espace est noté $\mathcal{M}^{(1)}(S,\Sigma,\kappa)$ où S est une surface compacte de genre g, Σ est un ensemble fini de points sur S (les singularités) et κ donne l'ordre des singularités.

Le théorème principal de ce chapitre est le suivant :

Théorème 1.1. *Soit ν une mesure de probabilité invariante pour l'action de* $\mathrm{SL}(2,\mathbb{R})$ *sur un espace de modules de surfaces plates $\mathcal{M}^{(1)}(S,\Sigma,\kappa)$. Alors, pour tout $\delta > 0$, le spectre du laplacien dans $[0,1/4-\delta]$ est formé de valeurs propres isolées de multiplicité finie.*

En particulier, le spectre du laplacien ne rencontre pas l'intervalle $]0, \varepsilon[$, pour un certain $\varepsilon = \varepsilon(\nu) > 0$: on dit que ν a un trou spectral. L'existence d'un trou spectral avait déjà été démontrée pour la mesure de Masur-Veech dans [Avila, Gouëzel et Yoccoz 2006], sans aucune estimée sur $\varepsilon(\nu)$ (le théorème 1.1 n'en donne d'ailleurs pas plus). Ici, on obtient des informations précises jusqu'à la barrière de $1/4$ au-delà de laquelle le résultat est déjà faux en genre 1, et pour toutes les mesures invariantes.

D'après [Ratner 1987], le trou spectral implique que l'action du sous-groupe diagonal de $\mathrm{SL}(2, \mathbb{R})$ (aussi appelée *flot de Teichmüller*) mélange exponentiellement vite pour des observables lisses. Ainsi, la question du spectre du laplacien est reliée à la description des corrélations (que nous avons évoquée pour des difféomorphismes Anosov dans le chapitre IV). Cette relation est au cœur de la preuve du théorème 1.1 : nous allons étudier l'action du flot de Teichmüller (qui a certaines propriétés d'hyperbolicité) et en déduire la description du spectre.

1.1. Description de la preuve. La stratégie habituelle pour montrer dans une surface hyperbolique de volume fini M que le spectre du laplacien est fini dans $[0, 1/4]$ est la suivante : on décompose $L^2(M)$ comme $L^2_{\mathrm{cusp}}(M) \oplus L^2_{\mathrm{eis}}(M)$ où $L^2_{\mathrm{cusp}}(M)$ est formé des fonctions de moyenne nulle sur tous les horocycles fermés, et $L^2_{\mathrm{eis}}(M)$ est son supplémentaire orthogonal. On prouve ensuite que le spectre dans $L^2_{\mathrm{eis}}(M)$ est $[1/4, \infty[$ en construisant une base de fonctions propres avec des séries d'Eisenstein, et que le spectre dans $L^2_{\mathrm{cusp}}(M)$ est discret puisque la convolution avec les fonctions lisses est un opérateur compact.

Quand on essaye de reproduire cette stratégie dans des situations non homogènes, il y a deux difficultés. Tout d'abord, comme la géométrie à l'infini est plus compliquée, les analogues de $L^2_{\mathrm{eis}}(M)$ et des séries d'Eisenstein ne sont pas clairement identifiés. D'autre part, la convolution avec des fonctions lisses sur $\mathrm{SL}(2, \mathbb{R})$ a uniquement un effet régularisant dans la direction des orbites de $\mathrm{SL}(2, \mathbb{R})$, et pas dans la direction transverse (le même problème apparaîtrait si on essayait d'étudier directement le laplacien). Ainsi, cette convolution n'est probablement pas un opérateur compact.

Pour résoudre la première difficulté, nous éviterons complètement la décomposition en composantes Eisenstein et cuspidales, et travaillerons avec l'espace L^2 entier. En particulier, il ne faut pas s'attendre à avoir des opérateurs compacts (car on aurait alors uniquement du spectre discret), nous aurons plutôt affaire à des opérateurs quasi-compacts, i.e., des opérateurs avec un nombre fini de valeurs propres de grand module, le reste du spectre étant contenu dans un disque de petit rayon. Les premières correspondront au spectre du laplacien dans $[0, 1/4 - \delta]$, et la deuxième partie donnera le reste du spectre du laplacien.

En ce qui concerne la seconde difficulté, nous ne regarderons pas le laplacien ou des opérateurs de convolution, mais un autre élément de l'algèbre enveloppante : la différentiation dans la direction du flot diagonal T_t. Bien sûr, son comportement sur L^2 est très mauvais, mais nous construirons un espace de Banach de distributions $\mathcal{B}$ sur lequel il est quasi-compact. Pour relier les propriétés spectrales du flot diagonal sur $\mathcal{B}$ et du laplacien sur L^2, on utilisera ensuite des asymptotiques précises de fonctions sphériques dans les représentations unitaires de $\mathrm{SL}(2,\mathbb{R})$ (cette partie est complètement générale et n'utilise rien de spécifique à l'espace des modules des surfaces de translation).

La difficulté principale est la construction de $\mathcal{B}$. Le flot T_t a certaines propriétés d'hyperbolicité, si bien qu'on pourra faire appel au même genre d'idées que dans le chapitre IV (plus précisément dans le cas simple où les feuilletages sont lisses, voir le paragraphe IV.2.2). Cependant, la non-compacité de l'espace des modules (et le manque d'hyperbolicité uniforme du flot) compliquent beaucoup les choses, et on aura besoin d'utiliser toute la richesse de la structure géométrique des espaces de modules pour réussir à conclure. En particulier, on aura besoin d'estimées de récurrence près de l'infini dues à [Eskin et Masur 2001].

Décrivons rapidement un point clé de la preuve. On devra à un moment étudier les itérés $\mathcal{L}_{t_0}^n$ de l'opérateur $\mathcal{L}_{t_0} f = f \circ T_{t_0}$, pour un certain t_0 bien choisi. En utilisant une partition de l'unité, on décompose $\mathcal{L}_{t_0}$ comme $\tilde{\mathcal{L}}_1 + \tilde{\mathcal{L}}_2$ où $\tilde{\mathcal{L}}_1$ prend en compte ce qui se passe dans un très gros compact K, et $\tilde{\mathcal{L}}_2$ correspond à ce qui arrive hors de K. Développons $\mathcal{L}_{t_0}^n = \sum_{\gamma_i \in \{1,2\}} \tilde{\mathcal{L}}_{\gamma_1} \cdots \tilde{\mathcal{L}}_{\gamma_n}$. Dans un terme de cette somme, si la plupart des γ_i sont égaux à 2, on passe beaucoup de temps hors de K, et la fonction d'Eskin-Masur donne un gain du bon ordre de grandeur. Sinon, on passe une partie respectable du temps dans K, où le flot est hyperbolique, si bien qu'on obtient un gain lié à l'hyperbolicité λ du flot sur ce compact. Malheureusement, on sait seulement que λ est strictement plus grand que 1 (et comme le compact K est extrêmement grand, il est certainement très proche de 1). Cela suffit pour obtenir un trou spectral, mais ne permet pas d'atteindre $1/4$ dans le spectre du laplacien. Cependant, si on définit l'espace $\mathcal{B}$ en utilisant une régularité de type C^k et C^{-k}, le gain est meilleur, de l'ordre de λ^{-k}. En choisissant k assez grand tout à la fin de l'argument, on obtiendra des estimées aussi précises que l'on veut, ce qui permettra d'arriver arbitrairement près de $1/4$ dans le spectre du laplacien.

Au vu de cet argument, on peut faire deux remarques. Tout d'abord, comme on a besoin d'utiliser des régularités arbitrairement grandes, cet argument ne peut pas fonctionner dans un modèle symbolique avec des discontinuités (en particulier, l'utilisation d'échanges d'intervalles et de l'induction de Rauzy est proscrite). D'autre part, comme k est choisi tout à la fin de l'argument, on doit s'assurer que toutes les estimées, qui doivent déjà être uniformes sur un espace non compact, sont aussi uniformes en k.

2. Espaces de Banach distributionnels pour le flot de Teichmüller

2.1. L'espace des modules des surfaces de translation. Une surface de translation est un triplet (S, Σ, ω) où S est une surface compacte, connexe, sans bord, orientée, $\Sigma = \{P_1, \dots, P_s\}$ est un ensemble fini de points $P_i \in S$ et ω est un atlas de translation sur $S \setminus \Sigma$, i.e., un atlas pour lequel les changements de carte sont des translations. On demande de plus que, au voisinage de chaque singularité P_i, Σ soit isomorphe à un revêtement de $\mathbb{R}^2 - \{0\}$ avec un nombre fini κ_i de feuillets.

Une telle surface de translation peut être construite en partant d'une famille finie de polygones du plan dont les côtés sont regroupés par paires de même direction et de même longueur. En recollant cette famille de polygones le long des côtés correspondants, on obtient une surface de translation, dont les singularités correspondent aux sommets des polygones. De plus, toute surface de translation peut s'obtenir ainsi.

Fixons S, Σ et les ordres $\kappa_1, \dots, \kappa_s$ des singularités (avec la contrainte $\sum(\kappa_i - 1) = 2g-2$, où g est le genre de S). L'espace de Teichmüller $\mathcal{Q}(S,\Sigma,\kappa)$ est l'ensemble des structures de translation où l'on identifie deux telles structures si l'une s'obtient à partir de l'autre par un homéomorphisme de S isotope à l'identité et fixant Σ. C'est une variété de dimension finie, dont on peut décrire des cartes comme suit.

Soit ω une structure de translation sur (S,Σ). Si $\gamma \in C^0([0,t],S)$ est un chemin continu sur S, on peut le relever en un chemin dans $\mathbb{R}^2$ partant de 0 : le relèvement est clairement possible en dehors des singularités, et la forme locale des singularités assure qu'elles ne créent pas de problème. En regardant la valeur du relèvement en t, on obtient l'application développante $D_\omega : C^0([0,t],S) \to \mathbb{R}^2$. Elle induit une application linéaire $H_1(S,\Sigma;\mathbb{Z}) \to \mathbb{R}^2$, i.e., un élément de $H^1(S,\Sigma;\mathbb{R}^2)$. Comme celui-ci est invariant par isotopie, on obtient une application bien définie $\Theta : \mathcal{Q}(S,\Sigma,\kappa) \to H^1(S,\Sigma;\mathbb{R}^2)$, appelée application des périodes. Elle définit une carte locale autour de chaque point de $\mathcal{Q}(S,\Sigma,\kappa)$, qui est donc une variété linéaire de dimension réelle $2(2g+s-1)$.

Une structure de translation définit une forme volume sur $S - \Sigma$, pour laquelle S a une aire finie. Notons $\mathcal{Q}^{(1)}(S,\Sigma,\kappa)$ l'hypersurface des surfaces d'aire 1.

L'espace $H^1(S,\Sigma;\mathbb{R}^2)$ a une forme volume standard (la mesure de Lebesgue pour laquelle les points entiers ont covolume 1). En la tirant localement en arrière par Θ, on obtient une mesure lisse μ_* sur $\mathcal{Q}(S,\Sigma,\kappa)$. Elle induit à nouveau une mesure lisse $\mu_*^{(1)}$ sur l'hypersurface $\mathcal{Q}^{(1)}(M,\Sigma,\kappa)$.

Le groupe $\mathrm{SL}(2,\mathbb{R})$ agit sur $\mathcal{Q}(S,\Sigma,\kappa)$ par postcomposition dans les cartes de translation. Cette action préserve l'hypersurface $\mathcal{Q}^{(1)}(M,\Sigma,\kappa)$ ainsi que μ_* et $\mu_*^{(1)}$. En particulier, l'action de $T_t \coloneqq \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ est un flot qui préserve la mesure, appelé *flot de Teichmüller*.

Le groupe modulaire de (S,Σ) est l'ensemble des homéomorphismes de S qui fixent Σ, modulo isotopie fixant Σ. Il agit sur l'espace de Teichmüller $\mathcal{Q}(M,\Sigma,\kappa)$. Le quotient, noté $\mathcal{M}(S,\Sigma,\kappa)$, est appelé *espace des modules*. L'action du groupe modulaire sur $\mathcal{Q}(M,\Sigma,\kappa)$ est fidèle et propre, mais elle n'est pas libre. Ainsi, $\mathcal{M}(S,\Sigma,\kappa)$ a une structure d'orbifold complexe linéaire.

Comme l'action du groupe modulaire préserve la mesure μ_* et l'hypersurface $\mathcal{Q}^{(1)}$, on obtient également une mesure ν_* sur l'espace des modules, ainsi qu'une hypersurface $\mathcal{M}^{(1)}(S,\Sigma,\kappa)$ (correspondant aux surfaces de translation d'aire 1) munie d'une mesure $\nu_*^{(1)}$. Comme l'action de $\mathrm{SL}(2,\mathbb{R})$ commute avec celle du groupe modulaire, le groupe $\mathrm{SL}(2,\mathbb{R})$ agit encore sur $\mathcal{M}(S,\Sigma,\kappa)$ et $\mathcal{M}^{(1)}(S,\Sigma,\kappa)$, en préservant respectivement ν_* et $\nu_*^{(1)}$. En particulier, l'action de T_t définit à nouveau un flot sur ces espaces, qu'on appelle encore flot de Teichmüller. La mesure $\nu_*^{(1)}$ a une masse finie ; ce résultat, non trivial, est dû à Masur et Veech. Quitte à changer sa normalisation, on pourra supposer que c'est une mesure de probabilité. La restriction de $\nu_*^{(1)}$ à chaque composante connexe de $\mathcal{M}^{(1)}(M,\Sigma,\kappa)$ est ergodique (et donc mélangeante d'après le théorème de

Howe-Moore). Le flot T_t ayant des propriétés d'hyperbolicité, cette mesure est même Bernoulli.

Il est instructif de considérer le cas du genre 1 (avec un point marqué), où tous les objets que je viens de décrire correspondent à des objets plus classiques. En ce cas, $\mathcal{Q}(S, \Sigma, \kappa)$ est $\mathrm{GL}^+(2, \mathbb{R})$ puisqu'un tore plat orienté est donné par deux vecteurs de base, linéairement indépendants, positivement orientés. L'ensemble des surfaces d'aire 1 $\mathcal{Q}^{(1)}(S, \Sigma, \kappa)$ correspond à $\mathrm{SL}(2, \mathbb{R})$. Les mesures μ_* et $\mu_*^{(1)}$ sont les mesures de Haar respectives. L'action de $\mathrm{SL}(2, \mathbb{R})$ sur $\mathcal{Q}(S, \Sigma, \kappa)$ et $\mathcal{Q}^{(1)}(S, \Sigma, \kappa)$ est simplement la multiplication à gauche. Le groupe modulaire est $\mathrm{SL}(2, \mathbb{Z})$, les espaces $\mathcal{M}(S, \Sigma, \kappa)$ et $\mathcal{M}^{(1)}(S, \Sigma, \kappa)$ sont donc $\mathrm{GL}^+(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$ et $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$, et les mesures ν_* et $\nu_*^{(1)}$ sont les quotients des mesures de Haar. La finitude de $\nu_*^{(1)}$ correspond donc au fait que $\mathrm{SL}(2, \mathbb{Z})$ soit un réseau dans $\mathrm{SL}(2, \mathbb{R})$, et l'ergodicité du flot de Teichmüller traduit l'ergodicité du flot géodésique sur la surface modulaire.

2.2. Hyperbolicité et récurrence. Pour étudier les propriétés de mélange du flot de Teichmüller T_t sur un espace de modules $\mathcal{M}$ muni d'une mesure de probabilité ν invariante sous l'action de $\mathrm{SL}(2, \mathbb{R})$, on souhaite utiliser des techniques similaires à celles du chapitre IV. Pour cela, il faut que le flot soit hyperbolique, ce qui est le cas en un sens faible que nous allons maintenant décrire. Notons $\mathcal{Q}$ l'espace de Teichmüller qui revêt $\mathcal{M}$. Toutes les constructions seront effectuées dans $\mathcal{Q}$, on devra s'assurer qu'elles sont équivariantes si on veut obtenir des objets bien définis sur $\mathcal{M}$.

Localement, l'application des périodes Θ identifie $\mathcal{Q}$ avec $H^1(S, \Sigma; \mathbb{R}) \oplus \mathbf{i} H^1(S, \Sigma; \mathbb{R})$. Lorsqu'on fait agir la matrice $T_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$, on obtient donc une dilatation de e^t sur la partie réelle, et une contraction de e^{-t} sur la partie imaginaire. Le flot a donc l'air parfaitement hyperbolique. Cependant, quand on passe au quotient par le groupe modulaire pour travailler dans l'espace des modules $\mathcal{M}$, l'action du groupe modulaire peut contrebalancer cette dilatation et cette contraction. C'est inévitable, puisque certains champs de vecteurs sont invariants par le flot (ne serait-ce que la direction du flot !). Plus précisément, définissons deux champs de vecteurs $a(x) = \Re\Theta(x)$ et $b(x) = \Im\Theta(x)$ (où on a identifié l'espace tangent à $\mathcal{Q}$ avec $H^1(S, \Sigma; \mathbb{C})$ par l'application des périodes). Si on a $\Theta(x) = a + \mathbf{i}b$, alors $\Theta(T_t x) = e^t a + \mathbf{i} e^{-t} b$, ce qui montre que les champs de vecteurs a et b sont invariants par le flot. Ils passent au quotient et définissent donc encore sur $\mathcal{M}$ deux champs de vecteurs invariants. Géométriquement, ils s'interprètent de la manière suivante. Définissons un autre flot U_t qui dilate la surface d'un facteur e^t. Il commute avec T_t, si bien que le champ de vecteurs qu'il définit est invariant par T_t. On a $\mathrm{d}T_t(x)/\,\mathrm{d}t|_{t=0} = a - \mathbf{i}b$ et $\mathrm{d}U_t(x)/\,\mathrm{d}t|_{t=0} = a + \mathbf{i}b$.

Il se trouve que ces deux champs de vecteurs invariants sont la seule source de non-hyperbolicité du flot. Comme U_t disparaît quand on se restreint aux surfaces d'aire 1, la seule source de non-hyperbolicité du flot pour T_t sur $\mathcal{M}^{(1)}$ est donc la direction du flot, ce qui est inévitable. Autrement dit, T_t n'a pas d'exposants de Lyapunov nuls non triviaux sur $\mathcal{M}^{(1)}$ (et les directions dilatée et contractée correspondent bien aux parties réelle et imaginaire de la cohomologie, comme on s'y attend). En revanche, l'hyperbolicité dégénère près de l'infini, et on obtient des estimées quantitatives seulement si on reste assez longtemps dans un compact :

PROPOSITION 2.1. *Pour tout ensemble $K \subset \mathcal{Q}^{(1)}$ qui est compact modulo le groupe modulaire, il existe $A = A(K)$ tel que, pour tout $x \in K$ et tout temps t tels que $T_t x \in K$ et*

$$\mathrm{Leb}\{s \in [0,t] \ : \ T_s(x) \in K\} \geqslant A,$$

alors on a $\|DT_t(x)v\|_{T_t x} \leqslant \|v\|_x /2$ pour tout $v \in E^s(x) = \mathbf{i}H^1(S,\Sigma;\mathbb{R}) \cap \mathcal{T}_x\mathcal{Q}^{(1)}$, et $\|DT_t(x)v\|_{T_t x} \geqslant 2\|v\|_x$ pour tout $v \in E^u(x) = H^1(S,\Sigma;\mathbb{R}) \cap \mathcal{T}_x\mathcal{Q}^{(1)}$.

Cette proposition est due à [FORNI 2002, Lemma 2.1']. Dans sa formulation, on a utilisé une norme sur l'espace tangent à $\mathcal{Q}^{(1)}$ (provenant d'une norme sur $\mathcal{M}^{(1)}$) ; la preuve de FORNI utilise une norme spécifique avec de bonnes propriétés de contraction, mais le résultat est vrai pour n'importe quelle norme puisqu'elles sont toutes équivalentes sur K.

Notons une conséquence de cette proposition : les directions stables et instables, qui ont une description algébrique, dépendent de manière C^∞ du point considéré. En cela, le flot de Teichmüller se distingue des flots hyperboliques génériques.

Au vu de la proposition précédente, il est important de s'assurer qu'on passe la plupart du temps dans des compacts pour tirer profit de l'hyperbolicité. Les compacts de $\mathcal{M}^{(1)}$ peuvent être décrits géométriquement, comme suit. On appelle *connexion de selles* un segment tracé sur une surface plate dont les deux extrémités sont des singularités, et qui ne rencontre pas de singularité dans son intérieur. La *systole* d'une surface plate x, notée $\mathrm{sys}(x)$, est la longueur de la plus petite connexion de selles sur x. Une suite de surfaces tend vers l'infini dans $\mathcal{M}$ si et seulement si la systole tend vers 0.

La proposition suivante garantit un retour rapide de la plupart des points dans les compacts. Considérons une mesure μ sur $\mathcal{Q}^{(1)}$, invariante sous l'action de $\mathrm{SL}(2,\mathbb{R})$, qui induit une mesure de probabilité ν sur $\mathcal{M}^{(1)}$. Elle admet alors une structure de produit local : elle se désintègre localement comme le produit d'une mesure μ_u sur les variétés instables, d'une mesure μ_s sur les variétés stables, et de la mesure de Lebesgue dans la direction du flot T_t. Les propriétés de récurrence à un compact seront valides le long de chaque variété instable locale (disons de taille 1/100 autour de x), notée $W^u_{1/100}(x)$, et seront formulées en termes de la mesure μ_u.

PROPOSITION 2.2. *Soit $\delta \in (0,1/4)$. Il existe $C > 0$ tel que, pour tout $x \in \mathcal{Q}^{(1)}$ et pour tout $t \geqslant 0$,*

$$\frac{1}{\mu_u(W^u_{1/100}(x))} \int_{W^u_{1/100}(x)} V(T_t y) \,\mathrm{d}\mu_u(y) \leqslant Ce^{-(1-2\delta)t} V(x) + C,$$

où $V(x) = \max(1/\mathrm{sys}(x), 1)$.

Autrement dit, si on est assez proche de l'infini pour que $V(x)$ soit très grand, alors $V(T_t y)$ est significativement plus petit que $V(x)$ pour la plupart des points y autour de x, i.e., la plupart des points proches de x se rapprochent rapidement de la partie compacte de l'espace. En revanche, si $V(x)$ est borné, la proposition apporte moins d'informations, elle dit simplement qu'il n'y a pas trop de points qui partent rapidement vers l'infini.

Cette proposition a été essentiellement démontrée dans [Eskin et Masur 2001]. Plus précisément, ils ont prouvé une formule similaire pour des moyennes sur des cercles, avec une constante qui dépend de t, mais il s'agit ensuite simplement de manipulations formelles pour passer des cercles aux horocycles instables, puis de recouvrir un morceau de variété instable par une réunion de tels horocycles.

Dans la proposition précédente, on a intégré sur une variété instable locale de taille fixée autour de chaque point. Comme x varie dans l'espace non compact $\mathcal{Q}^{(1)}$ et qu'on veut des estimées uniformes, le choix précis de la distance est important. On a besoin d'une distance provenant d'une métrique sur chaque espace tangent (pour pouvoir mesurer les vecteurs tangents), mais pas nécessairement euclidienne. Elle doit principalement vérifier les propriétés suivantes :

(1) Elle doit être complète et invariante sous le groupe modulaire (pour définir une métrique sur $\mathcal{M}$).

(2) Son comportement à l'infini doit être contrôlé (plus précisément, elle doit varier lentement dans le sens de Hörmander).

(3) Sous l'action du flot de Teichmüller, la métrique doit être dilatée (au sens large) dans la direction instable, et contractée (au sens large) dans la direction stable.

Décrivons la construction d'une telle métrique. Comme l'espace tangent à $\mathcal{Q}$ est partout identifié à $H^1(S, \Sigma; \mathbb{C})$ par l'application des périodes, il suffit de définir une famille de normes sur $H^1(S, \Sigma; \mathbb{C})$, dépendant continûment d'un point $x \in \mathcal{Q}$, comme suit :

$$\|v\|_x = \sup \left| \frac{v(\gamma)}{\Theta(x)(\gamma)} \right|,$$

où γ parcourt les connexions de selles de la surface x. Dans [Avila, Gouëzel et Yoccoz 2006], nous avons démontré que cette formule définissait bien une norme, qui dépendait continûment du point x. Les autres propriétés exigées ci-dessus sont également assez faciles à vérifier. Le point le plus délicat, la variation lente, est crucial pour obtenir des partitions de l'unité avec des normes C^k bien contrôlées, grâce à [Hörmander 2003, Theorem 1.4.10]. L'équivariance sous l'action du groupe modulaire est triviale puisque la définition donnée est purement géométrique et ne dépend donc pas d'un marquage de l'homologie.

2.3. L'espace de Banach $\mathcal{B}$. On peut maintenant définir l'espace de Banach $\mathcal{B}$ de distributions sur lequel le flot de Teichmüller a de bonnes propriétés spectrales. On s'inspire du paragraphe IV.2.2, en ajoutant la fonction V de la proposition 2.2 dans la définition de l'espace (et en échangeant les rôles des directions stable et instable car on compose par T_t au lieu de son inverse). Fixons une fois pour toutes une mesure invariante μ sur $\mathcal{Q}^{(1)}$ qui induit une mesure de probabilité ν sur $\mathcal{M}^{(1)}$, et considérons les mesures μ_u qu'elle engendre sur les variétés instables locales. On définit des coefficients

$$A_{k,\ell}(f) = \sup_x \sup_{v_1,\dots,v_k} \sup_\varphi \frac{1}{V(x)} \frac{1}{\mu_u(W^u_{1/200}(x))} \left| \int_{W^u_{1/200}(x)} \varphi \cdot L_{\overline{v_1}} \cdots L_{\overline{v_\ell}} f \,\mathrm{d}\mu_u \right|,$$

où les champs de vecteurs v_i sont définis sur $W^u_{1/100}(x)$, pointent partout dans la direction stable ou la direction du flot, ont une norme C^k bornée par 1 le long de $W^u_{1/100}(x)$, et où la fonction test φ a une norme C^ℓ bornée par 1. La notation $\overline{v_i}$ désigne une extension canonique de v_i à un voisinage de $W^u_{1/200}(x)$ reposant sur la structure affine de $\mathcal{Q}$.

Pour ce type de coefficients, on a contraction pour les champs de vecteurs qui pointent dans la direction stable si on passe suffisamment de temps dans un compact, d'après la proposition 2.1. Notons que, contrairement aux arguments du chapitre IV, on n'a pas de perte de constante ou de besoin de norme pondérée car la norme C^k des v_i est calculée uniquement le long de la partie dilatée $W^u_{1/100}(x)$, et non pas dans la direction transverse : c'est la structure affine du flot qui rend ceci possible, le processus d'extension du champ de vecteurs à un voisinage étant canonique en fonction de la structure affine, et donc équivariant par le flot.

Si on est hors d'un compact, les propriétés de la fonction V énoncées dans la proposition 2.2 assurent également un gain. Il reste juste à traiter les v_i qui pointent dans la direction du flot : comme ils ne sont pas contractés, on n'obtient pas d'information sur l'opérateur de composition par T_t. En revanche, si on regarde un opérateur intégral du type $f \mapsto \int_0^\infty e^{-\delta t} f \circ T_t$, on peut se débarrasser de ces champs de vecteurs en faisant une intégration par parties dans la direction du temps. Par chance, ce sont ces opérateurs qu'il s'agit d'étudier naturellement, car ils correspondent à la résolvante de la dérivation dans la direction du flot.

Notons $\mathcal{B}^K$ la complétion de l'espace des fonctions C^∞ à support compact pour la norme donnée par $\sum_{k=0}^K A_{k,K}(f)$. En incorporant soigneusement tous les éléments évoqués rapidement dans le paragraphe précédent, on obtient le théorème suivant :

THÉORÈME 2.3. *Soit $\delta > 0$. L'opérateur $R(\delta) : f \mapsto \int_0^\infty e^{-\delta t} f \circ T_t \,\mathrm{d}t$ agit continûment sur $\mathcal{B}^K$. Si K est assez grand, son rayon spectral essentiel est $< 1 + \delta$.*

Notons que le rayon spectral de $R(\delta)$ est borné par $\int_0^\infty e^{-\delta t} = 1/\delta$. Ainsi, ce théorème affirme que le rayon spectral essentiel de $R(\delta)$ est beaucoup plus petit que son rayon spectral, c'est donc bien un analogue pour ce flot des résultats du chapitre IV.

REMARQUE 2.4. Ce type d'estimées n'est pas suffisant pour déduire le mélange exponentiel du flot de Teichmüller. En effet, il implique que le spectre du générateur du flot est discret dans une bande $\Re z \in [-1 + \delta, 0]$, mais ceci n'exclut pas l'accumulation de valeurs propres sur l'axe imaginaire (avec une partie réelle qui tend vers 0 mais une partie imaginaire qui diverge). Un argument supplémentaire à la DOLGOPYAT (voir par exemple [AVILA, GOUËZEL et YOCCOZ 2006 ; LIVERANI 2004]) serait nécessaire pour exclure ce phénomène. Ici, on va pouvoir s'en passer en tirant parti du fait que le flot de Teichmüller provient d'une action de $\mathrm{SL}(2, \mathbb{R})$.

3. Relation entre informations géométriques et algébriques sur l'action de $\mathrm{SL}(2,\mathbb{R})$

Dans cette section, on décrit comment déduire le théorème 1.1 (sur le spectre du laplacien) des propriétés spectrales de $R(\delta)$ sur l'espace $\mathcal{B} = \mathcal{B}^K$, décrites dans le théorème 2.3. Le problème est que les espaces L^2 et $\mathcal{B}$ n'ont *a priori* pas grand chose à voir, le second étant un espace de distributions. Pour contourner cette difficulté, on va tout reformuler en termes de corrélations (qui ont un sens tant sur L^2 que sur $\mathcal{B}$), et plus précisément en termes d'extensions méromorphes de leurs transformées de Laplace.

Notons tout d'abord une conséquence du théorème 2.3. Pour $\delta > 0$, soit $D_\delta \subset \mathbb{C}$ l'ensemble des points $z = x + \mathbf{i}y$ avec $x > 0$, ou $(x, y) \in \,]-1+2\delta, 0] \times]-\delta, \delta[$.

Proposition 3.1. *Considérons deux fonctions f_1 et f_2 C^∞ à support compact sur $\mathcal{M}^{(1)}$, et définissons une fonction $F(z) = F_{f_1,f_2}(z) = \int_{t=0}^{\infty} e^{-zt} \left(\int_{\mathcal{M}^{(1)}} f_1 \cdot f_2 \circ T_t \,\mathrm{d}\nu\right) \mathrm{d}t$ pour $\Re z > 0$. Alors, pour tout $\delta > 0$, la fonction F admet une extension méromorphe à D_δ.*

Les pôles de F dans D_δ sont situés dans l'ensemble $\{1/\delta - 1/\lambda_1, \dots, 1/\delta - 1/\lambda_I\}$, où les λ_i sont les valeurs propres de module au moins $1+\delta$ de l'opérateur $R(\delta)$ agissant sur l'espace construit dans le théorème 2.3.

Ce résultat est essentiellement une conséquence formelle du théorème 2.3, puisqu'on peut écrire la fonction F_{f_1,f_2} en fonction de l'opérateur $R(\delta)$. La description spectrale de celui-ci donne donc l'extension méromorphe recherchée.

L'intérêt de la proposition est que l'espace de distributions $\mathcal{B}$ a disparu de son énoncé : on a seulement affaire à des transformées de Laplace de corrélations. On va maintenant voir comment celles-ci peuvent s'exprimer abstraitement dans une représentation unitaire générale de $\mathrm{SL}(2,\mathbb{R})$, en termes de spectre du laplacien.

3.1. Représentations unitaires de $\mathrm{SL}(2,\mathbb{R})$.

Décrivons tout d'abord informellement la notion d'intégrale directe de représentations (les détails sont dans [Dixmier 1969]).

Soit H_ξ une famille de représentations unitaires de $\mathrm{SL}(2,\mathbb{R})$, dépendant mesurablement (en un certain sens) d'un paramètre ξ dans un espace Ξ. Si m est une mesure sur Ξ, on définit l'intégrale directe de représentations $\int_\Xi H_\xi \,\mathrm{d}m(\xi)$: un élément de cet espace est une fonction f sur Ξ telle que $f(\xi) \in H_\xi$ pour tout ξ, avec $\|f\|^2 \coloneqq \int \|f(\xi)\|_{H_\xi}^2 \,\mathrm{d}m(\xi) < \infty$. Le groupe $\mathrm{SL}(2,\mathbb{R})$ agit unitairement sur cette intégrale directe, par $(g \cdot f)(\xi) = g(f(\xi))$. Si m' est une mesure équivalente à m, les intégrales directes $\int H_\xi \,\mathrm{d}m(\xi)$ et $\int H_\xi \,\mathrm{d}m'(\xi)$ sont isomorphes comme représentations unitaires.

Prenons dorénavant pour Ξ l'espace de toutes les représentations unitaires irréductibles de $\mathrm{SL}(2,\mathbb{R})$, avec sa structure borélienne canonique (que nous décrirons ci-dessous). Toute représentation unitaire H de $\mathrm{SL}(2,\mathbb{R})$ est isomorphe à une intégrale directe $\int H_\xi \,\mathrm{d}m(\xi)$, où l'espace H_ξ est une somme directe (finie ou dénombrable) de copies de la représentation ξ (on dit que H_ξ est quasi-irréductible). De plus, la classe de la mesure m et la multiplicité de ξ dans H_ξ ne dépendent pas de la décomposition choisie, et caractérisent la représentation H [Dixmier 1969, Théorème 8.6.6].

Décrivons plus précisément Ξ. Les représentations unitaires irréductibles de $\mathrm{SL}(2,\mathbb{R})$ ont été classifiées par BARGMANN, comme suit. Une telle représentation appartient à l'une des familles suivantes :

- Les représentations $\mathcal{D}^+_{m+1}$ et $\mathcal{D}^-_{m+1}$, pour $m \in \mathbb{N}$. Elles constituent la série discrète (sauf pour $m = 0$).
- Les représentations $\mathcal{P}^{+,\mathbf{i}v}$ pour $v \in [0,+\infty)$ et $\mathcal{P}^{-,\mathbf{i}v}$ pour $v \in (0,\infty)$. Elles constituent la série principale (elles peuvent aussi être définies pour $v < 0$, mais elles sont alors isomorphes à la représentation de paramètre $-v > 0$).
- Les représentations $\mathcal{C}^u$ pour $0 < u < 1$. Elles forment la série complémentaire.
- La représentation triviale.

Ces représentations sont décrites plus en détails dans [KNAPP 2001, II.5]. Elles sont toutes irréductibles, deux à deux non isomorphes, et toutes les représentations irréductibles apparaissent dans cette liste. En particulier, on peut associer à toute représentation irréductible ξ de $\mathrm{SL}(2,\mathbb{R})$ un paramètre complexe $s(\xi)$ (égal à m dans le premier cas, $\mathbf{i}v$ dans le second, u dans le troisième et 1 dans le dernier).

L'opérateur de Casimir Ω est un générateur du centre de l'algèbre enveloppante de $\mathrm{SL}(2,\mathbb{R})$, i.e., c'est un opérateur différentiel sur $\mathrm{SL}(2,\mathbb{R})$, de degré minimal, commutant avec les translations. Il est unique à multiplication près. Il induit un opérateur non borné dans toutes les représentations unitaires de $\mathrm{SL}(2,\mathbb{R})$. Comme il commute avec les translations, il doit être scalaire dans chaque représentation irréductible. Avec les notations précédentes et une normalisation convenable, il vaut $(1 - s(\xi)^2)/4 \in \mathbb{R}$ dans une représentation irréductible ξ de paramètre $s(\xi)$.

On dit qu'une représentation unitaire irréductible ξ de $\mathrm{SL}(2,\mathbb{R})$ est *sphérique* si elle contient un vecteur non nul invariant sous l'action du groupe orthogonal $\mathrm{SO}(2,\mathbb{R})$. En ce cas, il est unique à multiplication scalaire près. Considérons un tel élément v de norme 1. La fonction sphérique φ_ξ est définie sur $\mathrm{SL}(2,\mathbb{R})$ par

$$\varphi_\xi(g) = \langle g \cdot v, v\rangle. \tag{3.1}$$

Elle ne dépend pas du choix de v. Si l'on prend $g = T_t$, la fonction sphérique décrit simplement les corrélations de v sous l'action du flot diagonal.

Les représentations sphériques sont les représentations $\mathcal{P}^{+,\mathbf{i}v}$ et $\mathcal{C}^u$ (ainsi que la représentation triviale, bien sûr).

Supposons maintenant que $\mathrm{SL}(2,\mathbb{R})$ agit sur un espace Y en préservant une mesure de probabilité μ. Alors $\mathrm{SL}(2,\mathbb{R})$ agit unitairement sur $L^2(Y,\mu)$ par $g \cdot f(x) = f(g^{-1}x)$. En particulier, l'opérateur de Casimir agit sur $L^2(Y,\mu)$ (en tant qu'opérateur non borné). Comme il commute avec les translations, il préserve l'espace $L^2(\mathrm{SO}(2,\mathbb{R})\backslash Y,\mu)$ (i.e., l'espace des fonctions sur Y qui sont invariantes sous $\mathrm{SO}(2,\mathbb{R})$ et de carré intégral). Sur cet espace, Ω peut aussi être décrit géométriquement comme un laplacien feuilleté, comme suit.

Pour $x \in Y$, son orbite modulo $\mathrm{SO}(2,\mathbb{R})$ s'identifie à $\mathbb{H} = \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2,\mathbb{R})$ par l'application $g\mathrm{SO}(2,\mathbb{R}) \mapsto \mathrm{SO}(2,\mathbb{R})g^{-1}x$ (changer le point base de l'orbite modifie la paramétrisation par un élément de $\mathrm{SL}(2,\mathbb{R})$). Ainsi, toute structure sur $\mathbb{H}$ qui est préservée par $\mathrm{SL}(2,\mathbb{R})$ peut être transférée sur $\mathrm{SO}(2,\mathbb{R})\backslash Y$. C'est en particulier le cas de la métrique hyperbolique, et du laplacien correspondant Δ.

Soit f_K une fonction sur $\mathrm{SO}(2,\mathbb{R})\backslash Y$ dans le domaine de Δ, et soit f son relèvement canonique à Y. Alors Ωf est encore invariante sous $\mathrm{SO}(2,\mathbb{R})$, et c'est le relèvement de Δf_K sur $\mathrm{SO}(2,\mathbb{R})\backslash Y$. C'est une conséquence directe des définitions (si on a choisi la bonne normalisation pour Ω).

Considérons la décomposition de $L^2(Y,\mu)$ en intégrale directe de représentations quasi-irréductibles, $L^2(Y,\mu) \simeq \int_\Xi H_\xi \,\mathrm{d}m(\xi)$. Notons $H_\xi^{\mathrm{SO}(2,\mathbb{R})}$ l'ensemble des éléments de H_ξ invariants par $\mathrm{SO}(2,\mathbb{R})$, de telle sorte que $L^2(\mathrm{SO}(2,\mathbb{R})\backslash Y,\mu) \simeq \int_\Xi H_\xi^{\mathrm{SO}(2,\mathbb{R})} \,\mathrm{d}m(\xi)$. Ainsi, le spectre de Δ sur $L^2(\mathrm{SO}(2,\mathbb{R})\backslash Y,\mu)$ coïncide avec l'ensemble $\{(1-s(\xi)^2)/4\}$, où ξ parcourt les représentations sphériques dans le support de m (de plus, la mesure spectrale de Δ est l'image de m par cette application). Comme le spectre de l'opérateur de Casimir dans l'intervalle $]0,1/4[$ ne vient que des représentations de la série complémentaire, qui sont toutes sphériques, on en déduit que $\sigma(\Delta)\cap]0,1/4[= \sigma(\Omega)\cap]0,1/4[$, et que les mesures spectrales sont les mêmes. Par conséquent, il est équivalent de comprendre $\sigma(\Delta)\cap]0,1/4[$ ou de comprendre les représentations de la série complémentaire qui interviennent dans $L^2(Y,\mu)$. Le premier point de vue est plus élémentaire, mais le second le place dans une perspective plus générale (et, peut-être, plus significative).

On a maintenant tous les outils pour décrire les transformées de Laplace des corrélations dans une représentation unitaire générale de $\mathrm{SL}(2,\mathbb{R})$. Comme plus haut, on part d'une telle représentation H, qu'on décompose comme intégrale directe $\int_\Xi H_\xi \,\mathrm{d}m(\xi)$. Pour $f \in H$, on notera f_ξ sa composante dans H_ξ.

Notons $\Xi^{\mathrm{SO}(2,\mathbb{R})}$ l'ensemble des représentations irréductibles sphériques. En utilisant le paramétrage complexe s des représentations irréductibles, on peut identifier $\Xi^{\mathrm{SO}(2,\mathbb{R})}$ avec $]0,1] \cup \mathbf{i}\,[0,+\infty[$. On notera ξ_s la représentation de paramètre s.

Proposition 3.2. *Soient $f_1, f_2 \in H$ des fonctions invariantes par* $\mathrm{SO}(2,\mathbb{R})$. *On définit la transformée de Laplace des corrélations par*

$$F(z) = F_{f_1,f_2}(z) = \int_{t=0}^{\infty} e^{-zt}\langle T_t \cdot f_1, f_2\rangle \,\mathrm{d}t,$$

initialement pour $\Re(z) > 0$.

La fonction F admet une extension holomorphe à $\{\Re(z) > -1, z \notin (-1,0]\}$. De plus, pour tout $\delta > 0$, on peut décomposer la fonction F sur le demi-espace $\{\Re(z) > -1+2\delta\}$ comme somme d'une fonction holomorphe bornée A_δ et de la fonction

$$B_\delta(z) = \frac{1}{\sqrt{\pi}} \int_{s\in[\delta,1]} \frac{\Gamma(s/2)}{\Gamma((s+1)/2)} \langle (f_1)_{\xi_s}, (f_2)_{\xi_s}\rangle \frac{\mathrm{d}m(\xi_s)}{z-s+1}.$$

La preuve consiste simplement à écrire, pour $\Re(z) > 0$,

$$F(z) = \int_{t=0}^{\infty} e^{-zt}\langle T_t\cdot f_1, f_2\rangle = \int_\Xi \left(\int_{t=0}^{\infty} e^{-zt}\langle T_t \cdot (f_1)_\xi, (f_2)_\xi\rangle \,\mathrm{d}t\right) \mathrm{d}m(\xi).$$

Supposons pour simplifier H_ξ irréductible. En ce cas, $(f_1)_\xi$ et $(f_2)_\xi$ sont tous deux multiples de l'unique vecteur de H_ξ invariant par $\mathrm{SO}(2,\mathbb{R})$, ce qui donne $\langle T_t \cdot (f_1)_\xi, (f_2)_\xi\rangle =$

$\langle (f_1)_\xi, (f_2)_\xi \rangle \varphi_\xi(T_t)$ (où φ_ξ est la fonction sphérique définie en (3.1)). On dispose d'asymptotiques extrêmement précises pour les fonctions sphériques φ_ξ (voir [Ratner 1987, Theorem 1] pour la série principale et [Helgason 2000, Theorem IV.5.5] pour la série complémentaire). En injectant ces asymptotiques dans la formule précédente pour $F(z)$, on en déduit l'extension méromorphe recherchée.

3.2. Démonstration du théorème principal. Pour démontrer le théorème 1.1, on va accoler les descriptions des extensions méromorphes des transformées de Laplace des corrélations, d'origine géométrique (proposition 3.1) et algébrique (proposition 3.2). Comme les deux fonctions $F(z)$ coïncident pour $\Re(z) > 0$, elles coïncident partout. En particulier, la fonction B_δ de la proposition 3.2 est méromorphe sur $]-1+2\delta, 0[$, avec un nombre fini de pôles. Ainsi, pour une certaine fonction $C(s)$ (qu'on peut choisir non nulle), la fonction

$$\int_{[2\delta,1]} C(s) \frac{\mathrm{d}m(\xi_s)}{z-s+1}$$

est méromorphe. On en déduit que m ne charge que les pôles de cette fonction (d'après la formule d'inversion de Stieltjes, voir [Wall 1948]), et que c'est donc une somme de masses de Dirac. Autrement dit, si on se restreint à $[2\delta, 1]$, il n'y a qu'un nombre fini de représentations de la série complémentaire dans cet intervalle qui apparaissent dans la décomposition de $L^2(\mathcal{M}^{(1)}, \nu)$ comme intégrale directe de représentations irréductibles. D'après la discussion du paragraphe 3.1, cela équivaut au théorème 1.1 sur le spectre du laplacien.

Bibliographie personnelle

Avila, Artur et Gouëzel, Sébastien (2010), « Small eigenvalues of the Laplacian for algebraic measures in moduli space, and mixing properties of the Teichmüller flow ». Annals of mathematics, to appear (cf. p. 57).

Avila, Artur, Gouëzel, Sébastien et Tsujii, Masato (2006), « Smoothness of solenoidal attractors ». *Discrete Contin. Dyn. Syst.* **15** p. 21–35. MR : 2191383.

Avila, Artur, Gouëzel, Sébastien et Yoccoz, Jean-Christophe (2006), « Exponential mixing for the Teichmüller flow ». *Publ. Math. Inst. Hautes Études Sci.* **104** p. 143–211. MR : 2264836 (cf. p. 59, 64, 65).

Baladi, Viviane et Gouëzel, Sébastien (2009), « Good Banach spaces for piecewise hyperbolic maps via interpolation ». *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** p. 1453–1481. MR : 2542733 (cf. p. 6).

— (2010), « Banach spaces for piecewise cone-hyperbolic maps ». *J. Mod. Dyn.* **4** p. 91–137. MR : 2643889 (cf. p. 6).

Bálint, Péter et Gouëzel, Sébastien (2006), « Limit theorems in the stadium billiard ». *Comm. Math. Phys.* **263** p. 461–512. MR : 2207652 (cf. p. 19).

Bardet, Jean-Baptiste, Gouëzel, Sébastien et Keller, Gerhard (2007), « Limit theorems for coupled interval maps ». *Stoch. Dyn.* **7** p. 17–36. MR : 2293068 (cf. p. 14, 37).

Chazottes, Jean-René et Gouëzel, Sébastien (2007), « On almost-sure versions of classical limit theorems for dynamical systems ». *Probab. Theory Related Fields* **138** p. 195–234. MR : 2288069 (cf. p. 6).

— (2012), « Optimal concentration inequalities for dynamical systems ». *Comm. Math. Phys.* **316** p. 843–889. MR : 2993935 (cf. p. 6).

Dedecker, Jérôme, Gouëzel, Sébastien et Merlevède, Florence (2010), « Some almost sure results for unbounded functions of intermittent maps and their associated Markov chains ». *Ann. Inst. Henri Poincaré Probab. Stat.* **46** p. 796–821. MR : 2682267 (cf. p. 6).

— (2012), « The almost sure invariance principle for unbounded functions of expanding maps ». *ALEA Lat. Am. J. Probab. Math. Stat.* **9** p. 141–163. MR : 2904480 (cf. p. 6).

Gouëzel, Sébastien (2001), « Spectre de l'opérateur de transfert en dimension 1 ». *Manuscripta Math.* **106** p. 365–403. MR : 1869227.

— (2004a), « Central limit theorem and stable laws for intermittent maps ». *Probab. Theory Related Fields* **128** p. 82–122. MR : 2027296 (cf. p. 6).

GOUËZEL, Sébastien (2004b), « Sharp polynomial estimates for the decay of correlations ». *Israel J. Math.* **139** p. 29–65. MR : 2041223 (cf. p. 6).

— (2004c), « Vitesse de décorrélation et théorèmes limites pour les applications non uniformément dilatantes ». Thèse de doct. Université Paris Sud.

— (2005), « Berry-Esseen theorem and local limit theorem for non uniformly expanding maps ». *Ann. Inst. H. Poincaré Probab. Statist.* **41** p. 997–1024. MR : 2172207 (cf. p. 6).

— (2006a), « Decay of correlations for nonuniformly expanding systems ». *Bull. Soc. Math. France* **134** p. 1–31. MR : 2233699.

— (2006b), « Regularity of coboundaries for nonuniformly expanding Markov maps ». *Proc. Amer. Math. Soc.* **134** p. 391–401. MR : 2176007.

— (2007a), « A Borel-Cantelli lemma for intermittent interval maps ». *Nonlinearity* **20** p. 1491–1497. MR : 2327135 (cf. p. 6).

— (2007b), « Statistical properties of a skew product with a curve of neutral points ». *Ergodic Theory Dynam. Systems* **27** p. 123–151. MR : 2297091 (cf. p. 6).

— (2009a), « An interval map with a spectral gap on Lipschitz functions, but not on bounded variation functions ». *Discrete Contin. Dyn. Syst.* **24** p. 1205–1208. MR : 2505699.

— (2009b), « Local limit theorem for nonuniformly partially hyperbolic skew-products and Farey sequences ». *Duke Math. J.* **147** p. 193–284. MR : 2495076 (cf. p. 6).

— (2010a), « Almost sure invariance principle for dynamical systems by spectral methods ». *Ann. Probab.* **38** p. 1639–1671. MR : 2663640 (cf. p. 32).

— (2010b), « Characterization of weak convergence of Birkhoff sums for Gibbs-Markov maps ». *Israel J. Math.* **180** p. 1–41. MR : 2735054 (cf. p. 22).

— (2011), « Correlation asymptotics from large deviations in dynamical systems with infinite measure ». *Colloq. Math.* **125** p. 193–212. MR : 2871313.

— (2012), « Local limit theorem for symmetric random walks in Gromov-hyperbolic groups ». Preprint (cf. p. 6).

GOUËZEL, Sébastien et LALLEY, Steven P. (2011), « Random walks on co-compact Fuchsian groups ». Annales scientifiques de l'ENS, to appear (cf. p. 6).

GOUËZEL, Sébastien et LANNEAU, Erwan (2010), « Un théorème de Kerckhoff, Masur et Smillie : unique ergodicité sur les surfaces plates ». *École de Théorie Ergodique*. T. 20. Sémin. Congr. Paris : Soc. Math. France p. 113–145. MR : 2856512.

GOUËZEL, Sébastien et LIVERANI, Carlangelo (2006), « Banach spaces adapted to Anosov systems ». *Ergodic Theory Dynam. Systems* **26** p. 189–217. MR : 2201945 (cf. p. 43, 51, 55).

— (2008), « Compact locally maximal hyperbolic sets for smooth maps : fine statistical properties ». *J. Differential Geom.* **79** p. 433–477. MR : 2433929 (cf. p. 36, 43).

GOUËZEL, Sébastien, MATHÉUS, Frédéric et MAUCOURANT, François (2012), « Sharp lower bounds for the asymptotic entropy of symmetric random walks ». Preprint.

Bibliographie générale

Aaronson, Jon et Denker, Manfred (2001a), « 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. Boston, MA : Birkhäuser Boston p. 215–223. MR : 1890328 (cf. p. 20, 24).

— (2001b), « Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps ». *Stoch. Dyn.* **1** p. 193–237. MR : 1840194 (cf. p. 24).

Aaronson, Jon et Weiss, Benjamin (2000), « Remarks on the tightness of cocycles ». *Colloq. Math.* **84/85**. Dedicated to the memory of Anzelm Iwanik p. 363–376. MR : 1784202 (cf. p. 27).

Baladi, Viviane (2005), « Anisotropic Sobolev spaces and dynamical transfer operators : C^∞ foliations ». *Algebraic and topological dynamics*. T. 385. Contemp. Math. Providence, RI : Amer. Math. Soc. p. 123–135. MR : 2180233 (cf. p. 48).

Baladi, Viviane et Tsujii, Masato (2007), « Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms ». *Ann. Inst. Fourier (Grenoble)* **57** p. 127–154. MR : 2313087 (cf. p. 36, 43).

— (2008), « Dynamical determinants and spectrum for hyperbolic diffeomorphisms ». *Geometric and probabilistic structures in dynamics*. T. 469. Contemp. Math. Providence, RI : Amer. Math. Soc. p. 29–68. MR : 2478465 (cf. p. 15, 43).

Bergh, Jöran et Löfström, Jörgen (1976), *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Berlin : Springer-Verlag p. x+207. MR : 0482275 (cf. p. 11).

Berkes, István et Philipp, Walter (1979), « Approximation theorems for independent and weakly dependent random vectors ». *Ann. Probab.* **7** p. 29–54. MR : 515811 (cf. p. 31).

Billingsley, Patrick (1999), *Convergence of probability measures*. Second. Wiley Series in Probability and Statistics : Probability and Statistics. A Wiley-Interscience Publication. New York : John Wiley & Sons Inc. p. x+277. MR : 1700749 (cf. p. 38).

Bingham, Nicholas H., Goldie, Charles M. et Teugels, Jozef L. (1987), *Regular Variation*. T. 27. Encyclopedia of Mathematics and its Applications. Cambridge : Cambridge University Press p. xx+491. MR : 898871 (cf. p. 22, 28).

Blank, Michael, Keller, Gerhard et Liverani, Carlangelo (2002), « Ruelle-Perron-Frobenius spectrum for Anosov maps ». *Nonlinearity* **15** p. 1905–1973. MR : 1938476 (cf. p. 43).

Bowen, Rufus (1975), *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Berlin : Springer-Verlag p. i+108. MR : 0442989 (cf. p. 46).

BUZZI, Jérôme (1997), « Intrinsic ergodicity of affine maps in $[0,1]^d$ ». *Monatsh. Math.* **124** p. 97–118. MR : 1462857 (cf. p. 36).

CHAZOTTES, Jean-René, COLLET, Pierre et SCHMITT, Bernard (2005), « Statistical consequences of Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems ». *Nonlinearity* **18** p. 2341–2364. MR : 2165706 (cf. p. 15).

CHERNOV, Nikolai et MARKARIAN, Roberto (2006), *Chaotic billiards*. T. 127. Mathematical Surveys and Monographs. Providence, RI : American Mathematical Society p. xii+316. MR : 2229799 (cf. p. 18).

CHERNOV, Nikolai et ZHANG, Hong-Kun (2008), « Improved estimates for correlations in billiards ». *Comm. Math. Phys.* **277** p. 305–321. MR : 2358286 (cf. p. 19, 20).

DENKER, Manfred et PHILIPP, Walter (1984), « Approximation by Brownian motion for Gibbs measures and flows under a function ». *Ergodic Theory Dynam. Systems* **4** p. 541–552. MR : 779712 (cf. p. 31).

DIXMIER, Jacques (1969), *Les C^*-algèbres et leurs représentations*. Deuxième édition. Cahiers Scientifiques, Fasc. XXIX. Gauthier-Villars Éditeur, Paris p. xv+390. MR : 0246136 (cf. p. 66).

EINMAHL, Uwe (1989), « Extensions of results of Komlós, Major, and Tusnády to the multivariate case ». *J. Multivariate Anal.* **28** p. 20–68. MR : 996984 (cf. p. 31).

ESKIN, Alex et MASUR, Howard (2001), « Asymptotic formulas on flat surfaces ». *Ergodic Theory Dynam. Systems* **21** p. 443–478. MR : 1827113 (cf. p. 60, 64).

FIELD, Michael, MELBOURNE, Ian et TÖRÖK, Andrew (2003), « Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions ». *Ergodic Theory Dynam. Systems* **23** p. 87–110. MR : 1971198 (cf. p. 33).

FORNI, Giovanni (2002), « Deviation of ergodic averages for area-preserving flows on surfaces of higher genus ». *Ann. of Math. (2)* **155** p. 1–103. MR : 1888794 (cf. p. 63).

GUIVARC'H, Yves et HARDY, Jean (1988), « Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov ». *Ann. Inst. H. Poincaré Probab. Statist.* **24** p. 73–98. MR : 937957 (cf. p. 36).

HELGASON, Sigurdur (2000), *Groups and geometric analysis*. T. 83. Mathematical Surveys and Monographs. Integral geometry, invariant differential operators, and spherical functions, Corrected reprint of the 1984 original. Providence, RI : American Mathematical Society p. xxii+667. MR : 1790156 (cf. p. 69).

HENNION, Hubert (1993), « Sur un théorème spectral et son application aux noyaux lipchitziens ». *Proc. Amer. Math. Soc.* **118** p. 627–634. MR : 1129880 (cf. p. 14).

HOFBAUER, Franz et KELLER, Gerhard (1982), « Ergodic properties of invariant measures for piecewise monotonic transformations ». *Math. Z.* **180** p. 119–140. MR : 656227 (cf. p. 31, 37).

HÖRMANDER, Lars (2003), *The analysis of linear partial differential operators. I*. Classics in Mathematics. Distribution theory and Fourier analysis, Reprint of the second (1990) edition. Berlin : Springer-Verlag p. x+440. MR : 1996773 (cf. p. 64).

Ibragimov, Ildar A. (1966), « On the accuracy of approximation by the normal distribution of distribution functions of sums of independent random variables ». *Teor. Verojatnost. i Primenen* **11** p. 632–655. MR : 0212853 (cf. p. 30).

Ibragimov, Ildar A. et Linnik, Yuri V. (1971), *Independent and stationary sequences of random variables*. With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman. Wolters-Noordhoff Publishing, Groningen p. 443. MR : 0322926 (cf. p. 30).

Jiang, Miaohua (2011), « Differentiating potential functions of SRB measures on hyperbolic attractors ». To appear in Ergod. Theory Dynam. Systems (cf. p. 46).

Kato, Tosio (1966), *Perturbation theory for linear operators*. Die Grundlehren der mathematischen Wissenschaften, Band 132. Springer-Verlag New York, Inc., New York p. xix+592. MR : 0203473 (cf. p. 17).

Katok, Anatole et al. (1989), « Differentiability and analyticity of topological entropy for Anosov and geodesic flows ». *Invent. Math.* **98** p. 581–597. MR : 1022308 (cf. p. 46).

Keller, Gerhard et Liverani, Carlangelo (1999), « Stability of the spectrum for transfer operators ». *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **28** p. 141–152. MR : 1679080 (cf. p. 30).

Kitaev, Alexei Yu. (1999), « Fredholm determinants for hyperbolic diffeomorphisms of finite smoothness ». *Nonlinearity* **12** p. 141–179. MR : 1668543 (cf. p. 51).

Knapp, Anthony W. (2001), *Representation theory of semisimple groups*. Princeton Landmarks in Mathematics. An overview based on examples, Reprint of the 1986 original. Princeton, NJ : Princeton University Press p. xx+773. MR : 1880691 (cf. p. 67).

Kolář, Ivan, Michor, Peter W. et Slovák, Jan (1993), *Natural operations in differential geometry*. Berlin : Springer-Verlag p. vi+434. MR : 1202431 (cf. p. 54).

Komlós, János, Major, Péter et Tusnády, Gábor (1975), « An approximation of partial sums of independent RV's and the sample DF. I ». *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **32** p. 111–131. MR : 0375412 (cf. p. 31).

Liverani, Carlangelo (2003), « Invariant measures and their properties. A functional analytic point of view ». *Dynamical systems. Part II*. Pubbl. Cent. Ric. Mat. Ennio Giorgi. Pisa : Scuola Norm. Sup. p. 185–237. MR : 2071241 (cf. p. 30).

— (2004), « On contact Anosov flows ». *Ann. of Math. (2)* **159** p. 1275–1312. MR : 2113022 (cf. p. 65).

Luzzatto, Stefano et Melbourne, Ian (2011), « Statistical properties and decay of correlations for interval maps with critical points and singularities ». Preprint (cf. p. 37).

Melbourne, Ian et Nicol, Matthew (2009), « A vector-valued almost sure invariance principle for hyperbolic dynamical systems ». *Ann. Probab.* **37** p. 478–505. MR : 2510014 (cf. p. 31, 32, 36).

Melbourne, Ian et Török, Andrew (2002), « Central limit theorems and invariance principles for time-one maps of hyperbolic flows ». *Comm. Math. Phys.* **229** p. 57–71. MR : 1917674 (cf. p. 37).

MELBOURNE, Ian et TÖRÖK, Andrew (2004), « Statistical limit theorems for suspension flows ». *Israel J. Math.* **144** p. 191–209. MR : 2121540 (cf. p. 20, 33).

PARRY, William et POLLICOTT, Mark (1990), « Zeta functions and the periodic orbit structure of hyperbolic dynamics ». *Astérisque* p. 268. MR : 1085356 (cf. p. 36).

RATNER, Marina (1987), « The rate of mixing for geodesic and horocycle flows ». *Ergodic Theory Dynam. Systems* **7** p. 267–288. MR : 896798 (cf. p. 59, 69).

ROSENTHAL, Haskell P. (1970), « On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables ». *Israel J. Math.* **8** p. 273–303. MR : 0271721 (cf. p. 39).

RUELLE, David (1989), « The thermodynamic formalism for expanding maps ». *Comm. Math. Phys.* **125** p. 239–262. MR : 1016871 (cf. p. 44).

— (1990), « An extension of the theory of Fredholm determinants ». *Inst. Hautes Études Sci. Publ. Math.* **72** p. 175–193. MR : 1087395 (cf. p. 44).

— (1997), « Differentiation of SRB states ». *Comm. Math. Phys.* **187** p. 227–241. MR : 1463827 (cf. p. 46).

— (2003), « Correction and complements : "Differentiation of SRB states" ». *Comm. Math. Phys.* **234** p. 185–190. MR : 1963142 (cf. p. 46).

SAUSSOL, Benoît (2000), « Absolutely continuous invariant measures for multidimensional expanding maps ». *Israel J. Math.* **116** p. 223–248. MR : 1759406 (cf. p. 36, 37).

STONE, Charles (1965), « A local limit theorem for nonlattice multi-dimensional distribution functions ». *Ann. Math. Statist.* **36** p. 546–551. MR : 0175166 (cf. p. 18).

TSUJII, Masato (2010a), « Contact Anosov flows and the FBI transform ». To appear in Ergod. Theory Dynam. Systems (cf. p. 37).

— (2010b), « Quasi-compactness of transfer operators for contact Anosov flows ». *Nonlinearity* **23** p. 1495–1545. MR : 2652469 (cf. p. 37).

VOLNÝ, Dalibor (1990), « On limit theorems and category for dynamical systems ». *Yokohama Math. J.* **38** p. 29–35. MR : 1093661 (cf. p. 7, 22).

WALL, Hubert S. (1948), *Analytic Theory of Continued Fractions*. D. Van Nostrand Company Inc., New York p. xiii+433. MR : 0025596 (cf. p. 69).

YOUNG, Lai-Sang (1998), « Statistical properties of dynamical systems with some hyperbolicity ». *Ann. of Math. (2)* **147** p. 585–650. MR : 1637655 (cf. p. 18).

ZAĬTSEV, Andrei Yu. (1998), « Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments ». *ESAIM Probab. Statist.* **2** 41–108 (electronic). MR : 1616527 (cf. p. 31).

— (2007), « Estimates for the rate of strong approximation in the multidimensional invariance principle ». *J. Math. Sci. (N. Y.)* **145** p. 4856–4865. MR : 2355400 (cf. p. 41).