

UNIVERSITE PIERRE ET MARIE CURIE - PARIS 6  
LABORATOIRE DE PROBABILITES ET MODELES  
ALEATOIRES

# EXPOSANTS DE LYAPUNOV DE SYSTÈMES DYNAMIQUES ALÉATOIRES SUR LE TORE

Thèse de doctorat

Discipline : MATHÉMATIQUES

Spécialité : SYSTÈMES DYNAMIQUES

Présentée et soutenue par

**Dominique MALICET**

le 15 juin 2012

<i>Directeur :</i>	Raphaël KRIKORIAN	-	Université Paris 6
<i>Rapporteurs :</i>	Konstantin KHANIN	-	Université de Toronto
	Jean-Christophe YOCOZ	-	Collège de France
<i>Examineurs :</i>	Artur AVILA	-	CNRS Paris 7-IMPA
	Sylvain CROVISIER	-	Université Paris-Sud
	Dmitry DOLGOPYAT	-	Université de Maryland
	Bassam FAYAD	-	CNRS Paris 7
	Yves LE JAN	-	Université Paris-Sud
	Jean-Paul THOUVENOT	-	CNRS Paris 6



## Remerciements

Je remercie évidemment en premier lieu mon directeur de thèse, Raphaël Krikorian. Je le remercie pour ses conseils, sa patience, sa sympathie, pour le sujet de thèse très intéressant qu'il m'a proposé, pour m'avoir laissé l'autonomie que je désirais tout en restant prêt à me donner un coup de main salvateur lorsque je me heurtais à de grosses difficultés. Et je le remercie particulièrement pour avoir pensé à regarder les commutateurs de cocycles D'Anzai.

Je remercie Konstantin Khanin et Jean-Christophe Yoccoz d'avoir accepté de rapporter ma thèse. Et je remercie aussi ce dernier pour toutes les remarques et suggestions pertinentes qu'il m'a données durant sa minutieuse lecture de ce mémoire.

Je remercie Artur Avila, Sylvain Crovisier, Dmitry Dolgopyat, Bassam Fayad, Yves Le Jan et Jean-Paul Thouvenot d'avoir accepté de faire partie de mon jury. Je remercie particulièrement Artur Avila pour certaines idées qu'il m'a suggérées (je lui dois mon résultat principal du chapitre 4), ainsi que Jean-Paul Thouvenot avec qui j'ai pu apprendre beaucoup en peu de temps (Son engouement à partager sa (grande) culture mathématique est très stimulant).

Je remercie les (ex)-Cachanais : Guillaume, Rafik, Martin, Antoine, Vincent, Ayman..

Je m'excuse auprès des (ex-)cachanais de ne pas leur offrir de remerciements personnalisés, mais mon esprit est ailleurs.

Je remercie de même les (ex-)bordelais, futur marseillais : Yavar, Stéphane, Rachid, Bro..

Je ne remercie pas Mickaël car il me trouverait trop conformiste sinon.

Je remercie Selda, Lynn et Manon pour tous les bons moments passés et à passer en leur compagnie.

Je remercie mes parents, et plus généralement ma famille : grand-parents, oncle, frères, soeurs (ah euh..non), nièce(s)..

Je remercie tous ceux qui lisent ces remerciements.

Ainsi que ceux qui ne les lisent pas.

Et tous ceux que j'ai oubliés.



# Table des matières

<b>1</b>	<b>Rappels</b>	<b>9</b>
1.1	Système dynamique aléatoire, mesure stationnaire . . . . .	9
1.2	Nombre de rotation, ensemble de rotation . . . . .	10
1.3	Exposants de Lyapunov . . . . .	12
<b>2</b>	<b>Théorèmes de rigidité sur les exposants de Lyapunov</b>	<b>15</b>
2.1	Quelques théorèmes de rigidité . . . . .	15
2.2	Démonstration d'un cas particulier du théorème de Avila/Viana	16
2.3	Conséquence de la négativité des exposants de Lyapunov . . . .	22
<b>3</b>	<b>Théorèmes de conjugaison locale, théorie KAM</b>	<b>29</b>
3.1	Le théorème de conjugaison locale d'Arnold . . . . .	29
3.2	Une généralisation d'un théorème de Moser . . . . .	32
<b>4</b>	<b>Théorèmes de conjugaison lisse simultanée, cocycles d'Anzai</b>	<b>41</b>
4.1	Mesures invariantes de perturbations de translations . . . . .	42
4.2	Cocycles d'Anzai, unique ergodicité . . . . .	44
4.3	Une caractérisation dynamique des cocycles d'Anzai généralisés	46
<b>5</b>	<b>Lyapunov exponents of dynamical systems on the circle</b>	<b>55</b>
5.1	Introduction . . . . .	55
5.1.1	Notations and main results . . . . .	56
5.2	Proof of the main Theorem . . . . .	58
5.2.1	Cohomological equation . . . . .	58
5.2.2	Estimation of $T$ -invariant measure . . . . .	59
5.2.3	Estimate of the Lyapunov exponent . . . . .	61
5.3	Simultaneous conjugation . . . . .	64
5.3.1	First conjugation . . . . .	66
5.3.2	KAM iteration . . . . .	68
<b>6</b>	<b>On simultaneous linearization of diffeomorphisms of <math>\mathbb{T}^2</math></b>	<b>71</b>
6.1	Notations and results . . . . .	72
6.1.1	Scheme of the proof . . . . .	74
6.2	Cohomological equation . . . . .	76

6.3	Estimates on the stationary measure and the sum of Lyapunov exponents, first conjugation . . . . .	85
6.4	Projective action . . . . .	88
6.5	KAM iteration . . . . .	105

# Introduction

Dans cette thèse, nous étudions des propriétés de certains systèmes dynamiques discrets aléatoires sur des variétés riemanniennes compactes. Nous nous intéresserons plus particulièrement aux compositions aléatoires de difféomorphismes du tore de dimension  $d$  ( $d \geq 1$ ).

La dynamique discrète classique consiste à étudier les itérées d'une fonction d'évolution  $f$  sur un espace topologique  $X$ , tandis que la dynamique aléatoire consiste à regarder par exemple deux fonctions  $f_1, f_2$  sur  $X$  et à étudier les compositions aléatoires de  $f_1$  et  $f_2$ . C'est à dire que l'on se donne une suite i.i.d.  $(i_n)$  à valeurs dans  $\{1, 2\}$ , et on regarde les itérations  $F_n = f_{i_{n-1}} \circ \dots \circ f_{i_0}$ , ou de manière équivalente, la chaîne de Markov

$$X_{n+1} = f_{i_n}(X_n), X_0 = x \in X.$$

La trajectoire ne dépend plus uniquement de la condition initiale  $x$  mais aussi des valeurs prises par la suite de Bernoulli  $(i_n)$ . On peut généraliser ce type de système à une infinité de fonctions en prenant d'autres lois que celle de Bernoulli.

Les systèmes dynamiques aléatoires englobent de nombreuses chaînes de Markov couramment étudiées en théorie des probabilités. Ils apparaissent notamment en discrétisant des équations différentielles aléatoires de la forme  $\frac{dX}{dt} = \phi(X_t, \xi_t)$  avec  $\xi$  un processus de Levy, lesquelles interviennent par exemple en physique statistique. Ces systèmes dynamiques apparaissent également pour étudier certains modèles de population ou pour simuler une intelligence artificielle en théorie des jeux.

Voyons un exemple tiré des travaux d'Erdos. Prenons  $0 < a < 1$  et  $\theta_k$  une suite de Bernoulli à valeurs dans  $\{-1, 1\}$ . On s'intéresse à la loi de  $\sum_k \theta_k a^k$ , et notamment aux valeurs de  $a$  pour lesquelles celle-ci est absolument continue par rapport à la mesure de Lebesgue. Il s'avère alors que cette loi est l'unique probabilité stationnaire du système dynamique  $X_{n+1} = aX_n + \theta_{n+1}$  (la définition de mesure stationnaire sera donnée dans le chapitre "Rappels"). Pour de plus amples détails sur les systèmes dynamiques aléatoires, on pourra consulter [10]

Les aspects "stochastiques" d'un système dynamique aléatoire (ou même

déterministes) sont en bonne partie régis par leurs exposants de Lyapunov :

Dans le cas d'un système dynamique déterministe de dimension 1, les exposants de Lyapunov sont :

$$\left\{ \int \ln |f'| d\mu \mid f_*\mu = \mu \right\}.$$

Par le théorème ergodique, si  $\mu$  est une mesure invariante ergodique, pour  $\mu$  presque tout  $x$  on a la formule :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln |(f^n)'(x)| = \gamma = \int \ln |f'| d\mu.$$

Heuristiquement, pour  $n \gg 1$  et  $x$  une condition initiale générique :

$$|f^n(x + \Delta) - f^n(x)| \approx |(f^n)'(x)|\Delta \approx e^{N\gamma}\Delta.$$

Il est à noter que l'exposant de Lyapunov dépend cruciallement de la mesure invariante choisie. En dimension  $d > 1$ , cette définition se généralise à une famille de  $d$  exposants de Lyapunov, représentant les sensibilités directionnelles du système. Ces définitions se transportent aisément au cas aléatoire.

La non-nullité des exposants de Lyapunov représente en un certain sens un certain "chaos" du système. Mais la théorie de Pesin [5] donne alors l'existence l'existence de variétés stables et instables qui peuvent donner une certaine compréhension de ce chaos.

Il s'avère fréquent qu'un système dynamique possède des exposants de Lyapunov non nuls, spécialement quand le système a une composante aléatoire, mais étonnamment ce fait est en général difficile à démontrer. Un résultat dû à H.Furstenberg [8] ayant eu une importance considérable dit que si  $(M_n)$  est une suite de variables aléatoires i.i.d. (indépendantes identiquement distribuées) à valeurs dans  $Sl_d(\mathbb{R})$ , alors le plus grand exposant de Lyapunov associé au produit  $P_n = M_{n-1} \cdots M_0$ , que l'on peut définir par

$$\gamma = \lim_{n \rightarrow +\infty} \frac{1}{n} \mathbb{E}[\ln \|P_n\|],$$

est strictement positif excepté dans certaines situations exceptionnelles, que nous qualifierons de rigide, où il existe une mesure laissée invariante par presque toute réalisation de  $M_0$ . Ce résultat fût par la suite généralisé dans différents contextes, par exemple dans des cas où les matrices  $M_n$  ne sont pas indépendantes (Virtser [18], Royer [17], Ledrappier [13]), ou en considérant des compositions de difféomorphismes plus nécessairement linéaires. (Avila-Viana [4]).



Chacune de ces généralisations illustre une certaine dichotomie :

-soit le système possède au moins un exposant de Lyapunov non nul, ce qui signifie que le système a une forte sensibilité aux conditions initiales et a donc une dynamique que l'on peut qualifier de "chaotique".

-soit le système a une certaine rigidité dynamique, qui s'exprime par certaines propriétés sur ses mesures invariantes, ce qui peut dans certains cas permettre une réduction du système à un plus simple. La dynamique du système sera alors plutôt régulière.

Les résultats énoncés plus haut sont d'une nature assez générale. En étudiant des systèmes dynamiques plus particuliers, on peut obtenir une dichotomie plus forte, où dans le premier cas de l'alternative on obtient des estimées plus précises sur les exposants de Lyapunov tandis que dans le second cas de l'alternative la rigidité du système s'exprime par une conjugaison, éventuellement lisse, à un système simple, auquel cas la dynamique du système sera très régulière. Citons par exemple l'article de D. Dolgopyat et R. Krikorian [6] sur l'étude de compositions indépendantes de difféomorphismes de la sphère proches de rotation, ou celui de A. Avila et R. Krikorian [3] sur les produits de matrices de Schrödinger à potentiels périodiques.

Une telle dichotomie forte peut s'avérer très utile pour l'étude d'un système dynamique : si ce système est conjugué à un système simple, alors sa dynamique est relativement aisée à étudier, tandis que des renseignements sur les exposants de Lyapunov permettent d'utiliser la théorie de Pesin [5]. C'est ainsi que dans [6] la dichotomie prouvée est utilisée pour montrer qu'un système dynamique aléatoire généré par des difféomorphismes de la sphère proches de rotations et préservant le volume est nécessairement ergodique, et dans [3] pour résoudre la conjecture d'Aubry-André sur la mesure de Lebesgue du spectre de certains opérateurs de Schrödinger.

Dans cette thèse nous désirons établir une dichotomie forte dans le cadre d'une composition aléatoire indépendante de difféomorphismes du tore  $d$ -dimensionnel ( $d \geq 1$ ), en montrant une alternative entre l'existence d'exposants de Lyapunov strictement négatif et une conjugaison simultanée des difféomorphismes composés à une famille de difféomorphismes simples. Nous nous intéresserons plus particulièrement aux cas  $d = 1$  et  $2$ .

Il existe beaucoup de résultats généraux sur les systèmes dynamiques sur  $\mathbb{T}$  (citons notamment le théorème de conjugaison lisse de M.R. Herman et J.C. Yoccoz [19]). En revanche pour  $d > 1$ , l'étude dynamique d'un difféomorphisme quelconque de  $\mathbb{T}^d$  est un problème très difficile, et nous nous restreindrons donc à un cas local (les difféomorphismes considérés sont choisis proche de difféomorphismes connus). De plus, à l'instar d'une majorité de

théorèmes de conjugaison lisse, nous imposerons des conditions arithmétiques sur les vecteurs de rotations des difféomorphismes.

Le premier chapitre présente les outils nécessaires par la suite, notamment la définition formelle de système dynamique aléatoire, la notion de mesure stationnaire et d'exposant de Lyapunov, et la définition de l'ensemble de rotation d'un difféomorphisme du tore.

Le second chapitre présente certains théorèmes de rigidité cités ci-dessus, et notamment une démonstration détaillée du théorème de A. Avila et M. Viana dans le contexte d'un système dynamique aléatoire sur le tore de dimension 1. Nous donnons ensuite quelques applications de ce résultat.

Dans le troisième chapitre, nous établissons quelques théorèmes de conjugaison locale sur les difféomorphismes du tore, dérivés du théorème de conjugaison d'Arnold [2], et présentons ainsi le procédé d'itération KAM, dont nous aurons besoin par la suite.

Dans le quatrième chapitre nous définissons une famille particulière de difféomorphismes du tore, que l'on appelle cocycles d'Anzai, et en donnerons quelques propriétés. Cette famille de difféomorphismes est intéressante en ce fait que toute composition aléatoire de cocycles d'Anzai n'a que des exposants de Lyapunov nuls. Un but de cette thèse est d'établir que dans certains contextes, la nullité des exposants de Lyapunov d'un système dynamique aléatoire sur  $\mathbb{T}^d$  est une caractérisation des difféomorphismes de cette famille, à changement de coordonnées près. Nous démontrerons dans ce chapitre un résultat allant en ce sens, basé sur le théorème de Avila-Viana énoncé au deuxième chapitre.

Dans les deux derniers chapitres nous nous intéressons à un résultat local de rigidité, à savoir que dans un contexte perturbatif, un système dynamique aléatoire lisse sur  $\mathbb{T}^d$  vérifiant certaines conditions arithmétiques dites diophantiennes admet des exposants de Lyapunov strictement négatifs excepté si l'on peut par un changement de coordonnées lisse ramener le système à un système de cocycles d'Anzai. Nous démontrons ce théorème dans les cas  $d = 1$  et  $2$ , utilisant des techniques présentées aux chapitres 2, 3 et 4.

# Chapitre 1

## Rappels

### 1.1 Système dynamique aléatoire, mesure stationnaire

Soit  $X$  un espace topologique. Nous appellerons transformation aléatoire de  $X$  toute application

$$f : \begin{array}{l} \Omega \times X \rightarrow X \\ (\omega, x) \mapsto f^\omega(x) \end{array}$$

où  $(\Omega, \mathcal{F}, \mathbb{P})$  est un espace probabilisé, telle que pour tout  $x$  dans  $X$ ,  $\omega \mapsto f^\omega(x)$  soit une application mesurable. Nous dirons que  $f$  vérifie une propriété fonctionnelle (continue,  $C^1$ , bijective..) si  $f^\omega$  vérifie cette propriété pour presque tout  $\omega$  dans  $\Omega$ . Un système dynamique aléatoire est la donnée d'une suite  $x_n$  définie par  $x_{n+1} = f^{\omega_n}(x_n)$ , où  $f$  est une transformation aléatoire et  $(\omega_n)$  une suite i.i.d. d'éléments de  $(\Omega, \mathcal{F}, \mathbb{P})$ .

La notion de mesure invariante devient dans ce contexte celle de *mesure stationnaire* (ou mesure d'équilibre). Les mesures stationnaires sont les mesures de probabilité  $\mu$  sur  $X$  telles que si  $x_0$  suit une loi  $\mu$  et est indépendante de  $(\omega_n)$ , alors  $x_n$  suit une loi  $\mu$  pour tout  $n$ . Cette propriété équivaut à vérifier la relation

$$\mu = \mathbb{E}[f_*\mu].$$

Autrement dit,  $\mu$  est stationnaire si elle est invariante par l'opérateur de transfert

$$T\varphi = \mathbb{E}[\varphi \circ f]$$

A l'instar des mesures invariantes, on dit qu'une mesure stationnaire est ergodique si elle est extrémale dans l'ensemble des mesures stationnaires. On peut vérifier qu'une mesure stationnaire  $\mu$  est ergodique si et seulement si pour toute fonction borelienne  $\varphi$  sur  $X$ ,  $T\varphi = \varphi \Rightarrow \varphi = \text{constante } \mu$ -pp.

Présentons maintenant un moyen de voir un système dynamique aléatoire comme un système dynamique classique. Soit  $\tilde{\Omega} = \Omega^{\mathbb{N}}$  muni de la probabilité  $\tilde{\mathbb{P}} = \mathbb{P}^{\otimes \mathbb{N}}$ , et soit  $\tau : \tilde{\Omega} \rightarrow \tilde{\Omega}$ , envoyant un élément  $\omega = (\omega_k)_{k \geq 0}$  sur  $\tau\omega = (\omega_{k+1})_{k \geq 0}$ . Considérons maintenant le système dynamique sur  $\tilde{\Omega} \times X$  défini par la transformation

$$F(\omega, x) = (\tau\omega, f^{\omega_0}(x)).$$

Les itérés de  $F$  sont donnés par

$$F^n(\omega, x) = (\tau^n\omega, f^{\omega_{n-1}} \circ \dots \circ f^{\omega_0}(x)).$$

Ainsi, la dynamique de  $F$  met en évidence la dynamique aléatoire de  $f$ . Une mesure  $\mu$  sera stationnaire pour  $f$  si et seulement si  $\mathbb{P} \times \mu$  est invariante pour  $F$ , et sera ergodique si et seulement si  $\mathbb{P} \times \mu$  l'est. Cela permet d'utiliser des théorèmes de dynamique classique. Par exemple, dans ce contexte le théorème de Birkhoff dit que si  $\mu$  est une mesure stationnaire ergodique et si  $\varphi$  est dans  $L^1(X, \mu)$ , alors pour  $\mu$ -presque tout choix de  $x_0$  on a presque sûrement

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(x_k) = \int_X \varphi(x) d\mu(x).$$

## 1.2 Nombre de rotation, ensemble de rotation

Soit  $f$  une application continue de  $\mathbb{T}$  dans  $\mathbb{T}$ . Cette application peut se relever en une application continue de  $\mathbb{R}$  dans  $\mathbb{R}$  sous la forme  $f(x) = kx + \phi(x)$ , où  $\phi$  est une fonction 1-periodique, et  $k$  un entier, ce relèvement étant qui plus est unique à une constante entière près. Si  $f$  est un homéomorphisme de  $\mathbb{T}$ , alors  $k = \pm 1$ . Si  $k = 1$ , on dit que  $f$  préserve l'orientation de  $\mathbb{T}$ , et on note  $\text{Hom}_+(\mathbb{T})$  l'ensemble des homéomorphismes de  $\mathbb{T}$  préservant l'orientation. Si  $f$  est un tel homéomorphisme, on peut vérifier que la quantité

$$\frac{f^n(x) - x}{n}$$

converge, uniformément en  $x$ , vers une constante que l'on notera  $\rho(f)$  et que l'on désignera comme le *nombre de rotation* de  $f$ . Les propriétés principales du nombre de rotation sont :

- 1) Si  $g$  appartient à  $\text{Hom}_+(\mathbb{T})$ ,  $\rho(gfg^{-1}) = \rho(f)$ .
- 2) Si  $r_\alpha(x) = x + \alpha$ ,  $\rho(r_\alpha) = \alpha$ .

Un objectif important en dynamique, ainsi que dans cette thèse, est l'étude de conjugaisons éventuelles de difféomorphismes du cercle à des rotation. Le nombre de rotation d'un homéomorphisme  $f$  donne la rotation à laquelle  $f$  doit être conjuguée si conjugaison il y a. De tels théorèmes de conjugaison nécessitent en général des hypothèses arithmétiques sur ces nombres de

rotations, entraînant certaines conséquences dynamiques sur  $f$ .

Présentons un moyen de calculer le nombre de rotation d'un homéomorphisme  $f$  à partir d'une mesure invariante  $\mu$ . En écrivant  $f = Id + \phi$  avec  $\phi$  1-périodique, on constate que

$$\frac{f^n(x) - x}{n} = \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x)).$$

Ainsi,  $\rho(f)$  apparaît comme la limite des sommes de Birkhoff de  $\phi$ . Si  $\mu$  est une mesure invariante par  $f$ , alors l'intégration de la formule précédente selon  $\mu$  donne

$$\rho(f) = \int_{\mathbb{T}} \phi(x) d\mu(x).$$

C'est cette formule que nous allons utiliser pour généraliser la notion de nombre de rotation en dimension supérieure.

Considérons donc maintenant une application continue de  $\mathbb{T}^d$  dans lui-même, où  $d$  est un entier strictement plus grand que 1. Une telle application se relève sur  $\mathbb{R}^d$ , de manière unique à un  $d$ -uplet d'entiers près, sous la forme  $f(x_1, \dots, x_d) = A(x_1, \dots, x_d) + \phi(x_1, \dots, x_d)$ , où  $A$  est une application linéaire sur  $\mathbb{R}^d$  stabilisant le réseau  $\mathbb{Z}^d$ , se représentant donc par une matrice  $d \times d$  à coefficients entiers, et  $\phi$  une application continue 1-périodique en chaque variable (i.e.  $\mathbb{Z}^d$ -périodique). Si de plus  $f$  est un homéomorphisme de  $\mathbb{T}^d$ , alors  $A$  doit être inversible, i.e.  $A$  appartient à  $Gl_d(\mathbb{Z})$ . On dit que  $A$  est la partie algébrique de  $f$ , et  $\phi$  sa partie périodique. Considérons maintenant le sous-ensemble d'homéomorphismes de  $\mathbb{T}^d$  dont la partie algébrique vaut  $Id$ , que l'on notera  $\text{Hom}_0(\mathbb{T}^d)$ . Si  $f$  est un élément de  $\text{Hom}_0(\mathbb{T}^d)$ , il s'écrit donc sous la forme  $f = Id + \phi$ , et à l'instar de la dimension 1, on a pour  $x$  dans  $\mathbb{R}^d$

$$\frac{f^n(x) - x}{n} = \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x)).$$

Cette quantité ne converge pas nécessairement quand  $n$  tend vers  $+\infty$ . Cependant, si  $\mu$  est une mesure invariante ergodique pour  $f$ , alors pour  $\mu$  presque toute valeur de  $x$  la quantité précédente converge vers  $\int_{\mathbb{T}^d} \phi(y) d\mu(y)$ . Et réciproquement, si la quantité précédente converge, alors la limite est de la forme  $\int_{\mathbb{T}^d} \phi(y) d\mu(y)$  pour une certaine mesure invariante  $\mu$ , choisie comme limite faible d'une sous-suite de  $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{f^k(x)}$ . Il est donc légitime d'appeler *vecteur de rotation* de  $f$  tout  $d$ -uplet de la forme  $\int_{\mathbb{T}^d} \phi(y) d\mu(y)$ , où  $\mu$  est une mesure invariante pour  $f$ . Et on appellera *ensemble de rotation* de  $f$ , que l'on notera  $\mathcal{R}(f)$ , l'ensemble des vecteurs de rotations de  $f$ . C'est un compact convexe de  $\mathbb{T}^d$ , et plus précisément l'enveloppe convexe des limites des quantités  $\frac{f^n(x) - x}{n}$  quand  $x$  parcourt  $\mathbb{T}^d$ . Comme en dimension 1, on a les

propriétés :

- 1) Si  $f, g$  appartiennent à  $\text{Hom}_0(\mathbb{T}^d)$ ,  $\mathcal{R}(gf g^{-1}) = \mathcal{R}(f)$ .
- 2) Pour  $\alpha$  dans  $\mathbb{T}^d$ , si  $r_\alpha(x) = x + \alpha$ ,  $\mathcal{R}(r_\alpha) = \{\alpha\}$ .

Regardons cette définition sur un exemple. Considérons un homéomorphisme de  $\mathbb{T}^2$  de la forme

$$f(x, y) = (x + \alpha, y + \beta(x))$$

où  $\alpha$  est un réel et  $\beta$  une fonction continue sur  $\mathbb{T}$ . C'est une classe d'homéomorphismes de  $\mathbb{T}^2$  qui s'avèrera particulièrement importante dans cette thèse et que nous étudierons plus profondément par la suite. Un tel homéomorphisme préserve la mesure de Lebesgue sur  $\mathbb{T}^2$ , donc, la partie periodique de  $f$  étant  $\phi(x, y) = (\alpha, \beta(x))$ , on en déduit que  $(\alpha, \int_{\mathbb{T}} \beta(x) dx)$  est un vecteur de rotation de  $f$ . Si  $\mu$  est une mesure  $f$ -invariante, alors

$$\int_{\mathbb{T}^2} \phi(x, y) d\mu(x, y) = (\alpha, \int_{\mathbb{T}^2} \beta(x) d\mu(x, y)),$$

donc  $\mathcal{R}(f) \subset \{\alpha\} \times \mathbb{T}$ . De plus, dans le cas où  $\alpha$  est irrationnel, la projection de  $\mu$  sur la première coordonnée étant invariante par  $x \mapsto x + \alpha$  est égale à la mesure de Lebesgue, donc  $\int_{\mathbb{T}^2} \beta(x) d\mu(x, y) = \int_{\mathbb{T}} \beta(x) dx$ , et  $\mathcal{R}(f)$  est donc réduit au seul point  $\{\alpha, \int_{\mathbb{T}} \beta(x) dx\}$ .

Cela n'est en revanche plus vrai si  $\alpha$  est rationnel. Par exemple, si  $\alpha = 0$ , et donc  $f(x, y) = (x, y + \beta(x))$ , alors pour tout  $x$  dans  $\mathbb{T}$  la mesure  $\delta_x \otimes dy$  est invariante par  $f$ , d'où on déduit que  $(\alpha, \beta(x))$  appartient à  $\mathcal{R}(f)$ . Donc dans le cas  $\alpha = 0$ ,  $\mathcal{R}(f) = \{0\} \times [\min \beta, \max \beta]$ .

On aurait pu aussi obtenir ces résultats en étudiant directement les limites éventuelles de

$$\frac{f^n(x, y) - (x, y)}{n} = (\alpha, \frac{1}{n} \sum_{k=0}^{n-1} \beta(x + k\alpha)).$$

### 1.3 Exposants de Lyapunov

Les exposants de Lyapunov sont un moyen de calculer les contractions et les expansions d'un système dynamique. Ils sont définis comme suit : Soient  $X$  un espace compact,  $f : X \rightarrow X$  une application continue, et  $\mu$  une mesure invariante ergodique. Soit  $(A(x))_{x \in X}$  une famille de matrices réelles  $d \times d$  indexée par  $X$ , de manière mesurable (i.e.  $x \mapsto A(x)$  est une application mesurable) tel que  $\ln \|A\| \in L^1(\mu)$ . Posons

$$A_n(x) = A(f^{n-1}(x)) \cdots A(x)$$

Les matrices  $A_n(x)$  se récupèrent naturellement par la dynamique d'un cocycle : définissant sur  $X \times \mathbb{R}^d$  la transformation  $T(x, h) = (f(x), A(x)h)$ , on a  $T^n(x, h) = (f^n(x), A_n(x)h)$ . On s'intéresse à la quantité

$$\lambda(x, h) = \limsup_{n \rightarrow +\infty} \frac{1}{n} \ln |A_n(x)h| \quad (1.1)$$

pour  $h$  dans  $\mathbb{R}^d$ . On vérifie aisément que pour tout réel  $t$  non nul,  $\lambda(x, th) = \lambda(x, h)$  et pour tout  $h_1, h_2$  de  $\mathbb{R}^d$   $\lambda(x, h_1 + h_2) \leq \max(\lambda(x, h_1), \lambda(x, h_2))$  avec égalité si  $\lambda(x, h_1) \neq \lambda(x, h_2)$ , et on en déduit qu'à  $x$  fixée,  $\lambda(x, h)$  ne prend qu'un nombre fini de valeurs  $\lambda_1 > \dots > \lambda_p$ ,  $p \leq d$ , quand  $h$  parcourt  $\mathbb{R}^d$ , et que les espaces  $V_k = V_k(x) = \{h \in \mathbb{R}^d \mid \lambda(x, h) \leq \lambda_k\}$  ( $k = 1, \dots, p$ ) forment une suite décroissante de sous espaces vectoriels de  $\mathbb{R}^d$ , vérifiant  $V_k(f(x)) = A(x)V_k(x)$ .

Notant  $m_k = \dim V_k - \dim V_{k+1}$  (convenant  $V_{d+1} = \{0\}$ ), on peut montrer (voir [15],[12]) que si  $\mu$  est une mesure invariante ergodique pour  $f$ , alors il existe un sous ensemble d'éléments  $x$  de  $X$  de  $\mu$ -mesure totale sur lequel les quantités  $p$ ,  $\lambda_k$  et  $m_k$  ne dépendent pas de  $x$  et tel que pour tout  $h$  de  $\mathbb{R}^d$ , la limite supérieure 1.1 est en réalité une vraie limite. Ces réels  $\lambda_1, \dots, \lambda_p$  sont appelés exposants de Lyapunov du cocycle  $T$  associés à la mesure invariante  $\mu$ , et  $m_k$  est la multiplicité de  $\lambda_k$ . Ainsi, pour  $\mu$ -presque tout  $x$  dans  $X$  et pour tout  $h$  appartenant à  $V_k - V_{k+1}$ ,

$$\lambda_k = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln |A_n(x)h|.$$

Comme  $\sum_k m_k = d$ , en comptant les exposants de Lyapunov avec leur multiplicité on définit ainsi  $d$  exposants de Lyapunov  $\gamma_1 \geq \gamma_2 \geq \dots \geq \gamma_d$ .

Le plus grand exposant  $\gamma_1 = \lambda_1$  et le plus petit exposant  $\gamma_d = \lambda_p$  sont donnés par les formules

$$\gamma_1 = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|A_n(x)\|, \gamma_d = \lim_{n \rightarrow +\infty} \frac{1}{n} \ln \|A_n(x)^{-1}\|^{-1},$$

valides pour  $\mu$ -presque tout point  $x$  de  $X$ .

Lorsque  $f$  est inversible, on a pour  $\mu$  presque tout  $x$  la décomposition de l'espace

$$\mathbb{R}^d = E_1(x) \oplus \dots \oplus E_p(x),$$

où

$$E_k(x) = \left\{ h \in \mathbb{R}^d \mid \lim_{n \rightarrow \pm\infty} \frac{1}{n} \ln |A_n(x)h| = \lambda_k \right\} \cup \{0\}$$

vérifie  $E_k(f(x)) = A(x)E_k(x)$ . On a alors

$$V_k(x) = E_k(x) \oplus \dots \oplus E_d(x).$$

Quand  $X$  est une variété riemannienne et  $f$  un difféomorphisme de  $X$ , une famille  $(A(x))$  naturelle associée est donnée par  $A(x) = Df(x)$ . Si  $f : \Omega \times X \rightarrow X$  est un difféomorphisme aléatoire de mesure stationnaire ergodique  $\mu$ , on lui

associe la transformation  $(\omega, x) \mapsto (\tau\omega, f^{\omega_0}(x))$  sur  $\Omega^{\mathbb{N}} \times X$  muni de la mesure invariante  $\mathbb{P}^{\mathbb{N}} \otimes \mu$  et on choisit  $A(\omega, x) = Df^{\omega}(x)$ . Les exposants de Lyapunov sont invariants par conjugaison : si  $g$  est un difféomorphisme de  $X$ , les exposants de Lyapunov de  $f$  associés à la mesure  $\mu$  sont identiques aux exposants de Lyapunov de  $gfg^{-1}$  à la mesure  $f_*\mu$ .



## Chapitre 2

# Théorèmes de rigidité sur les exposants de Lyapunov

### 2.1 Quelques théorèmes de rigidité

Les exposants de Lyapunov mesurent les contractions et expansions exponentielles d'un système. Ceux ci ne donnent *a priori* que peu d'information lorsqu'ils sont nuls. Mais il s'avère que la situation où tous les exposants de Lyapunov d'un système sont nuls ne se produit que rarement, divers théorèmes montrant que cette propriété entraîne souvent une importante rigidité dynamique. Nous allons explorer certains de ces théorèmes dans ce chapitre, ainsi que certaines de leurs conséquences.

Le premier résultat de rigidité de ce type est dû à H. Furstenberg [8], et concerne l'exposant de Lyapunov caractéristique d'un produit de matrices aléatoires i.i.d. Si  $(M_n)$  est une suite de variable aléatoires à valeurs dans  $Sl_d(\mathbb{R})$ , Furstenberg montre que l'exposant de Lyapunov caractéristique

$$\gamma = \lim_{n \rightarrow +\infty} \frac{1}{n} \mathbb{E}[\ln \|M_n\|]$$

est strictement positif, excepté s'il existe une mesure  $\mu$  sur l'espace projectif des droites de  $\mathbb{R}^d$  presque sûrement invariante par  $M_0$ .

F. Ledrappier [13] étudie le cas général d'un produit de matrices au dessus d'un système dynamique mesuré quelconque. Etant donné un espace probabilisé  $(X, \mu)$ , une transformation inversible  $f$  de  $X$  préservant une probabilité  $\mu$ , et une application mesurable  $A : X \rightarrow Sl_d(\mathbb{R})$ , il obtient un critère général assurant la positivité stricte de l'exposant de Lyapunov caractéristique en fonction des mesures invariantes  $\hat{\mu}$  du cocycle  $(x, h) \mapsto (f(x), A(x)h)$  se projetant sur  $\mu$ . Ledrappier récupère ainsi notamment le résultat de Furstenberg ainsi que des résultats généralisant celui de Furstenberg d'un produit de matrices sur un

système dynamique Markovien (voir [17] et [18]).

En utilisant la méthode de Ledrappier, A. Avila et F. Viana [4] ont démontré un résultat analogue sur les exposants de Lyapunov associés à un cocycle de la forme  $(x, y) \rightarrow (f(x), g_x(y))$ ,  $f$  étant une transformation inversible d'un espace probabilisé et  $(g_x)$  une famille de difféomorphismes d'une variété riemannienne compacte. Ce résultat a pour conséquence un résultat semblable à celui de Furstenberg sur les exposants de Lyapunov d'une composition aléatoire d'une suite i.i.d. de difféomorphismes, que l'on peut énoncer ainsi :

**Théorème 1.** *Soit  $f$  un difféomorphisme aléatoire  $C^1$  d'une variété riemannienne  $V$ ,  $\mu$  une mesure stationnaire de  $f$ . On suppose que  $\ln \|Df(\cdot)\|$  et  $\ln \|Df(\cdot)\|^{-1}$  sont  $\mathbb{P} \times \mu$  intégrables, et on note  $\gamma_1 \geq \dots \geq \gamma_d$  les exposants de Lyapunov associés à  $(f, \mu)$ . Alors  $\gamma_d < 0$ , excepté si  $f_*\mu = \mu$  presque sûrement.*

## 2.2 Démonstration d'un cas particulier du théorème de Avila/Viana

Nous allons maintenant démontrer le théorème de Avila/Viana en se restreignant au cas d'un système dynamique aléatoire sur le cercle afin d'alléger les notations et les calculs (le cas général est un peu plus technique mais suit la même trame de démonstration).

**Théorème 2.** *Soit  $f : \mathbb{T} \rightarrow \mathbb{T}$  un difféomorphisme aléatoire  $C^1$ , et  $\mu$  une mesure stationnaire ergodique pour  $f$ . On suppose  $\ln |f'|$  est  $\mathbb{P} \times \mu$ -intégrable, et on pose  $\gamma = \mathbb{E} \int_{\mathbb{T}} \ln |f'(x)| d\mu(x)$  l'exposant de Lyapunov associé à  $(f, \mu)$ . Alors si  $\gamma \leq 0$ , on a  $f_*\mu = \mu$  presque sûrement.*

L'idée est d'étudier la quantité

$$-h = \mathbb{E} \int_{\mathbb{T}} \ln \left( \frac{d(f^{-1})_*\mu}{d\mu}(x) \right) d\mu(x),$$

où  $\frac{d(f^{-1})_*\mu}{d\mu}$  représente la dérivée de Radon-Nikodym de  $f_*^{-1}\mu$  par rapport à  $\mu$  (plus précisément : c'est la fonction positive  $g$  définie  $\mu$  presque partout telle que  $(f^{-1})_*\mu$  s'écrit  $d(f^{-1})_*\mu = g d\mu + dv$  avec  $v$  mesure singulière par rapport à  $\mu$ ).

La quantité  $\ln \left( \frac{d(f^{-1})_*\mu}{d\mu} \right)$  n'étant *a priori* pas nécessairement intégrable, on prend la convention  $-h = -\infty$  si ce n'est pas le cas. Par l'inégalité de Jensen,

$$-h \leq \ln \left( \mathbb{E} \int_{\mathbb{T}} \frac{d(f^{-1})_*\mu}{d\mu}(x) d\mu(x) \right) \leq 0,$$

et  $h = 0$  si et seulement si  $\frac{d(f^{-1})_*\mu}{d\mu} = 1$   $\mu$ -presque partout et presque sûrement, c'est à dire si et seulement si  $f_*\mu = \mu$  presque sûrement.

## 2.2. DÉMONSTRATION D'UN CAS PARTICULIER DU THÉORÈME DE AVILA/VIANA17

On va montrer l'inégalité  $\min(\gamma, 0) \leq -h$ , qui implique clairement le théorème 2 par ce qui a été dit ci dessus. Pour ce faire nous utiliserons un lemme de géométrie de la mesure disant d'une part que l'on peut voir une dérivée de Radon Nikodym du type  $\frac{d\mu}{dv}(x)$  comme limite de rapport du type  $\frac{\mu(I)}{v(I)}$  où  $I$  est un intervalle centré en  $x$  de longueur tendant vers 0, et d'autre part qu'une probabilité  $\mu$  sur  $\mathbb{T}$  a une dimension inférieure à 1, dans le sens où pour  $r > 1$  il n'existe que peu d'intervalles de longueurs petites satisfaisant des inégalités du type  $\mu(I) \leq l(I)^r$ . On désignera dans la suite par  $I_\varepsilon(x)$  l'intervalle  $[x - \varepsilon, x + \varepsilon]$ .

**Proposition 1.** Soient  $\nu$  une probabilité sur  $\mathbb{T}$ , et  $\mu$  une mesure signée sur  $\mathbb{T}$ . Alors :

i) Pour  $\nu$  presque tout  $x$ ,

$$\frac{d\mu}{d\nu}(x) = \lim_{\varepsilon \rightarrow 0} \frac{\mu(I_\varepsilon(x))}{\nu(I_\varepsilon(x))}.$$

ii) Notant  $q^*(x) = \sup_{r>0} \frac{|\mu(I_r(x))|}{\nu(I_r(x))}$  la fonction maximale associée à  $\mu$ , on a

$$\int_{\mathbb{T}} \ln q^*(x) d\nu(x) \leq 2.$$

iii) Pour  $\nu$ -presque tout  $x$ ,

$$\limsup_{\varepsilon \rightarrow 0} \frac{\ln \nu(I_\varepsilon(x))}{\ln \varepsilon} \leq 1$$

Cette proposition étant un énoncé classique de géométrie de la mesure et n'étant que peu liée aux systèmes dynamiques, nous ne donnerons qu'une brève démonstration. Celle ci est basée sur le lemme de recouvrement de Besicovitch, dont une version simplifiée aisée à montrer sur  $\mathbb{T}$  dit que pour tout recouvrement d'un compact de  $\mathbb{T}$  par des intervalles, on peut en extraire un recouvrement fini tel que tout point de  $\mathbb{T}$  appartient à au plus 2 intervalles du sous-recouvrement.

On commence par établir l'inégalité maximale

$$\nu(\{q^* > \alpha\}) \leq \frac{2}{\alpha},$$

qui se montre en prenant un compact arbitraire  $K \subset \{q^* > \alpha\}$ , en choisissant pour chaque point  $x$  de  $K$  un intervalle  $I_x$  centré en  $x$  tel que  $|\mu(I_x)| > \alpha \nu(I_x)$ , en extrayant de  $(I_x)$  un sous recouvrement  $(I_k)$  de  $K$  par le lemme de Besicovitch puis en écrivant

$$|\mu(K)| \leq \sum_k |\mu(I_k)| \leq \alpha \sum_k \nu(I_k) \leq 2\alpha.$$

Cette inégalité maximale permet de déduire directement le point ii) par la formule

$$\int_{\mathbb{T}} \ln q^*(x) d\nu(x) = \int_1^{+\infty} \nu(\{q^* > t\}) \frac{dt}{t},$$

et de montrer qu'à  $\nu$  fixée, l'ensemble des mesures  $\mu$  satisfaisant i) est stable par limite en variation totale. Il suffit de vérifier alors que i) est vrai dans les cas simples où la mesure  $\mu$  a une densité continue par rapport à  $\nu$  et où le support topologique de  $\mu$  est de  $\nu$ -mesure nulle, puis de vérifier que ces mesures engendrent un sous-espace dense des mesures signées pour la convergence en variation totale.

Pour démontrer iii), on fixe  $\delta > 0$ , et  $K$  un compact inclus dans l'ensemble des points  $x$  vérifiant

$$\limsup_{\varepsilon \rightarrow 0} \frac{\ln \nu(I_\varepsilon(x))}{\ln \varepsilon} \geq 1 + 2\delta.$$

Pour chaque  $x$  de  $K$ , on choisit un intervalle  $I_x$  centré en  $x$  tel que  $\nu(I_x) \leq m(I_x)^{1+\delta}$  et  $m(I_x) \leq \varepsilon$ , on extrait un recouvrement  $(I_k)$  par le lemme de Besicovitch puis on écrit

$$\nu(K) \leq \sum_k \nu(I_k) \leq \varepsilon^\delta \sum_k m(I_k) \leq 2\varepsilon^\delta,$$

ce qui permet de conclure.

Retournons maintenant au contexte et aux notations du théorème 2.

**Lemme 1.** *Soit  $x$  générique pour  $\mu$  (i.e.  $x$  appartient à un certain ensemble de  $\mu$ -mesure 1), et  $(f_n)_{n \geq 0}$  une suite i.i.d. générique de copies de  $f$ , et soit  $(x_n)_{n \geq 0}$  la suite définie par  $x_0 = x$ ,  $x_{n+1} = f_n(x_n)$ . Si  $(\varepsilon_n)_{n \geq 0}$  est une suite de réels strictement positifs tendant vers 0 et vérifiant  $I_{\varepsilon_{n+1}}(x_{n+1}) \subset f_n(I_{\varepsilon_n}(x_n))$  pour tout  $n$ , alors on a les inégalités*

$$\liminf_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} \leq \limsup_{n \rightarrow +\infty} \frac{\mu(I_{\varepsilon_n}(x_n))}{n} \leq -h.$$

**Remarque :** Seule l'inégalité extrême  $\liminf_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} \leq -h$  nous intéresse par la suite, le terme central est uniquement un intermédiaire de calcul.

*Démonstration.* Commençons par prouver l'inégalité de droite. On commence par poser

$$q(y, \varepsilon) = \sup_{r < \varepsilon} \frac{(f^{-1})_* \mu(I_r(y))}{\mu(I_r(y))},$$

dont on sait par la proposition 1 que pour  $\mu$ -presque tout  $y$  de  $\mathbb{T}$ ,  $\lim_{\varepsilon \rightarrow 0} q(y, \varepsilon) = \frac{d(f^{-1})_* \mu}{d\mu}(y)$ , et que  $q(y, \varepsilon) \leq q^*(y) = \sup_{r > 0} \frac{(f^{-1})_* \mu(I_r(y))}{\mu(I_r(y))}$  avec  $\int_{\mathbb{T}} \ln q^*(y) d\mu(y) \leq 2$ . Maintenant, écrivons

$$\begin{aligned} \mu(I_{\varepsilon_n}(x_n)) &= \mu(I_{\varepsilon_0}(x)) \prod_{k=0}^{n-1} \frac{\mu(I_{\varepsilon_{k+1}}(x_{k+1}))}{\mu(I_{\varepsilon_k}(x_k))} \\ &\leq \mu(I_{\varepsilon_0}(x)) \prod_{k=0}^{n-1} \frac{(f_k^{-1})_* \mu(I_{\varepsilon_k}(x_k))}{\mu(I_{\varepsilon_k}(x_k))} \\ &\leq \mu(I_{\varepsilon_0}(x)) \prod_{k=0}^{n-1} q(x_k, \varepsilon_k) \end{aligned}$$

## 2.2. DÉMONSTRATION D'UN CAS PARTICULIER DU THÉORÈME DE AVILA/VIANA19

donc

$$\frac{\ln \mu(I_{\varepsilon_n}(x_n))}{n} \leq \frac{\ln(\mu(I_{\varepsilon_0}(x)))}{n} + \frac{1}{n} \sum_{k=0}^{n-1} \ln q(x_k, \varepsilon_k).$$

Le terme  $\frac{\ln(\mu(I_{\varepsilon_0}(x)))}{n}$  tend vers 0 quand  $n$  tend vers  $+\infty$ . Pour majorer asymptotiquement le second terme, fixons  $\varepsilon > 0$  et remarquons que pour  $k$  assez grand,  $\ln q(x_k, \varepsilon_k) \leq \ln q(x_k, \varepsilon)$ , de sorte que

$$\limsup_{n \rightarrow +\infty} \frac{\ln \mu(I_{\varepsilon_n}(x_n))}{n} \leq \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln q(x_k, \varepsilon).$$

L'idée est maintenant que  $x$  étant générique, on peut appliquer le théorème de Birkhoff à  $\ln q(\cdot, \varepsilon)$ . La fonction  $\ln q(\cdot, \varepsilon)$  n'étant pas nécessairement  $\mathbb{P} \times \mu$  intégrable, on majore celle-ci par  $\widetilde{\ln} q(\cdot, \varepsilon)$ , où on a posé  $\widetilde{\ln}(t) = \max(-A, \ln t)$ ,  $A$  étant un réel arbitrairement grand. Les inégalités  $-A \leq \widetilde{\ln} q(y, \varepsilon) \leq A + \ln q^*(y)$  entraînent l'intégrabilité de  $\widetilde{\ln} q(\cdot, \varepsilon)$  (uniforme en  $\varepsilon$ ), et on obtient donc par le théorème de Birkhoff

$$\limsup_{n \rightarrow +\infty} \frac{\ln \mu(I_{\varepsilon_n}(x_n))}{n} \leq \mathbb{E} \int_{\mathbb{T}} \widetilde{\ln} q(y, \varepsilon) d\mu(y).$$

En faisant tendre  $\varepsilon$  vers 0 on obtient par convergence dominée

$$\limsup_{n \rightarrow +\infty} \frac{\ln \mu(I_{\varepsilon_n}(x_n))}{n} \leq \mathbb{E} \int_{\mathbb{T}} \widetilde{\ln} \left( \frac{d(f^{-1})_* \mu}{d\mu}(y) \right) d\mu(y),$$

puis en faisant tendre  $A$  vers  $+\infty$ , une convergence monotone donne l'inégalité souhaitée.

Démontrons maintenant l'inégalité de gauche. Fixons  $c > 1$ . Grâce au point iii) de la proposition 1, on peut trouver une fonction mesurable  $\varepsilon(\cdot)$  définie sur  $\mathbb{T}$  à valeurs dans  $]0, 1]$  telle que pour  $\mu$ -presque tout  $y$  de  $\mathbb{T}$ , si  $\varepsilon < \varepsilon(y)$  alors  $\ln \mu(I_\varepsilon(y)) \geq c \ln \varepsilon$ . Il s'agit maintenant de montrer que l'on a  $\varepsilon_n \leq \varepsilon(x_n)$  pour une infinité de  $n$ . Supposons que ce ne soit pas le cas. On aurait donc  $\varepsilon(x_n) < \varepsilon_n$  pour  $n$  assez grand, ce qui implique en particulier que  $\varepsilon(x_n)$  tend vers 0 quand  $n$  tend vers  $+\infty$ , donc par le théorème de Birkhoff,

$$\int_{\mathbb{T}} \varepsilon(y) d\mu(y) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varepsilon(x_k) = 0,$$

et donc  $\varepsilon(y) = 0$  pour  $\mu$ -presque tout  $y$ , ce qui est absurde. Ainsi,  $\ln \mu(I_{\varepsilon_n}(x_n)) \geq c \ln \varepsilon_n$  pour une infinité de  $n$ , donc

$$\liminf_{n \rightarrow +\infty} c \frac{\ln \varepsilon_n}{n} \leq \limsup_{n \rightarrow +\infty} \frac{\ln \mu(I_{\varepsilon_n}(x_n))}{n}$$

et on obtient l'inégalité voulue en faisant tendre  $c$  vers 1.  $\square$

Prouvons maintenant le théorème 2. Fixons  $x$  et  $(f_n)$  génériques et définissons  $(x_n)$  comme dans le lemme 1. Il s'agit maintenant de construire une suite de réels  $\varepsilon_n$  vérifiant les hypothèses du lemme. Le théorème des accroissements finis montre que l'inclusion  $I_{\varepsilon_{n+1}}(x_{n+1}) \subset f_n(I_{\varepsilon_n}(x_n))$  est entraînée par l'inégalité  $\varepsilon_{n+1} \leq m_n(x_n, \varepsilon_n)\varepsilon_n$  où on a posé  $m_n(x, \varepsilon) = \inf_{I_\varepsilon(x)} |f'_n|$ . Fixons  $r > \max(0, \gamma)$  et définissons  $\varepsilon_n$  par  $\varepsilon_0 = 1$  et  $\varepsilon_{n+1} = m_n(x_n, \varepsilon_n)\varepsilon_n e^{-r}$ , ce qui implique en particulier  $\varepsilon_{n+1} \leq m_n(x_n, \varepsilon_n)\varepsilon_n$ . D'autre part,

$$\varepsilon_n = e^{-nr} \prod_{k=0}^{n-1} m_k(x_k, \varepsilon_k) \leq e^{-nr} \prod_{k=0}^{n-1} |f'_k(x_k)|,$$

donc par le théorème de Birkhoff,

$$\limsup_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} \leq \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln |f'_k(x_k)| - r = \gamma - r < 0,$$

donc en particulier  $\varepsilon_n$  tend (géométriquement) vers 0 quand  $n$  tend vers  $+\infty$ . Ainsi  $(\varepsilon_n)$  satisfait les hypothèses du lemme précédent, donc

$$\liminf_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} \leq -h,$$

et nous pouvons minorer le terme de gauche par la même méthode que pour montrer la première inégalité du lemme précédent : pour tout  $\varepsilon > 0$ , on a  $\varepsilon_k \leq \varepsilon$  si  $k$  assez grand, donc, en posant  $\tilde{\ln} t = \min(\ln t, A)$  où  $A$  est arbitrairement grand, de sorte que  $A \geq \tilde{\ln} m_0(y, \varepsilon) \geq -A - \ln |f'_0(y)|$ , on a

$$\begin{aligned} \liminf_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} &= \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \ln m_k(x_k, \varepsilon_k) - r \\ &\geq \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \tilde{\ln} m_k(x_k, \varepsilon) - r = \mathbb{E} \int_{\mathbb{T}} \tilde{\ln} m_0(y, \varepsilon) d\mu(y) - r, \end{aligned}$$

et en faisant tendre  $\varepsilon$  vers 0 puis  $A$  vers  $+\infty$ , le terme de droite tend vers  $\gamma - r$ , donc

$$\gamma - r \leq \liminf_{n \rightarrow +\infty} \frac{\ln \varepsilon_n}{n} \leq -h$$

L'inégalité  $\gamma - r \leq h$  étant vérifié pour tout  $r > \max(\gamma, 0)$ , on a donc  $\min(\gamma, 0) \leq -h$ , ce qui prouve le théorème 2.

**Remarque** : La différence principale dans la démonstration lorsqu'on remplace le tore  $\mathbb{T}$  par une variété compacte de dimension supérieure est la suivante : une analyse analogue à celle présentée ci dessus permet d'obtenir l'inégalité

$$\min\left(0, \mathbb{E} \int_{\mathbb{T}} \ln \|Df(x)^{-1}\|^{-1} d\mu(x)\right) \leq -dh,$$

## 2.2. DÉMONSTRATION D'UN CAS PARTICULIER DU THÉORÈME DE AVILA/VIANA21

$d$  étant la dimension de la variété (ce facteur  $d$  vient du fait que en dimension  $d$ , l'inégalité iii) de la proposition 1 devient  $\limsup_{\varepsilon \rightarrow 0} \frac{\ln \mu(B(x, \varepsilon))}{\ln \varepsilon} \leq d$ ). Cependant, lorsque  $d \geq 2$ , la quantité  $\mathbb{E} \int_{\mathbb{T}} \ln \|Df(x)^{-1}\|^{-1} d\mu(x)$  ne correspond en général pas à un exposant de Lyapunov du système. Pour remédier à cela, on applique cette inégalité en remplaçant  $f$  par  $F_p = f_{p-1} \circ \dots \circ f_0$ , ( $f_k$ ) étant comme d'habitude une suite de copies i.i.d. de  $f$ , ce qui permet d'obtenir l'inégalité

$$\min\left(0, \mathbb{E} \int_{\mathbb{T}} \ln \|DF_p(x)^{-1}\|^{-1} d\mu(x)\right) \leq -pdh,$$

et on peut montrer que lorsque  $p$  tends vers  $+\infty$ , la quantité  $\mathbb{E} \int_{\mathbb{T}} \frac{\ln \|DF_p(x)^{-1}\|^{-1}}{p} d\mu(x)$  converge vers le plus petit exposant de Lyapunov  $\gamma_d$  du système, et on obtient ainsi l'inégalité  $\min(\gamma_d, 0) \leq -dh$ , qui permet de conclure.

Etant donné un homéomorphisme aléatoire du cercle préservant l'orientation et minimal (un homéomorphisme aléatoire d'un espace topologique  $X$  est dit minimal s'il n'existe pas de fermés autre que  $X$  et  $\emptyset$  presque sûrement stables par l'homéomorphisme), l'existence d'une mesure presque sûrement invariante entraîne une conjugaison du système à un système de rotations. En effet, soit  $\mu$  est une mesure sur  $\mathbb{T}$  presque sûrement invariante par un difféomorphisme aléatoire  $f : \mathbb{T} \rightarrow \mathbb{T}$ . Commençons par montrer que la mesure  $\mu$  est à support total et sans atomes.

Notant  $U = (\text{supp } \mu)^c$ , on a  $0 = \mu(U) = \mathbb{E}[\mu(f^{-1}(U))]$ , donc  $f^{-1}(U)$  est presque sûrement un ouvert de  $\mu$ -mesure nulle donc inclus dans  $U$ , et on a donc presque sûrement  $f(\text{supp } \mu) \subset \text{supp } \mu$ , ce qui entraîne  $\text{supp } \mu = \mathbb{T}$  par minimalité.

D'autre part, si on suppose que  $M = \sup_{x \in \mathbb{T}} \mu(\{x\})$  est strictement positif, alors l'ensemble  $\{x \in \mathbb{T} | \mu(\{x\}) > \frac{M}{2}\}$  est fini, donc la borne supérieure est atteinte. L'ensemble  $F = \{x \in \mathbb{T} | \mu(\{x\}) = M\}$  est non vide, fini donc fermé, et si  $x$  est dans  $F$ ,  $M = \mu(\{x\}) = \mathbb{E}[\mu(\{f^{-1}(x)\})] \leq M$ , donc  $\mu(\{f^{-1}(x)\}) = M$  presque sûrement. Comme  $F$  est fini, cela entraîne  $f(F) = F$  presque sûrement. Mais  $F$  n'étant ni vide ni égal à  $\mathbb{T}$ , cela est absurde par l'hypothèse de minimalité, donc  $\mu$  n'a aucun atome.

Maintenant si on pose  $g(x) = \mu([0, x])$ , on a  $g \circ f = g + \alpha$ , où  $\alpha = \mu([0, f(0)])$ . Comme  $\mu$  est à support total sans masse de Dirac,  $g$  est un homéomorphisme de  $\mathbb{T}$ , et  $gf g^{-1} = Id + \alpha$ . Ceci permet d'obtenir le corollaire suivant au théorème 2.

**Corollaire 1.** *Soit  $f : \mathbb{T} \rightarrow \mathbb{T}$  un difféomorphisme aléatoire préservant l'orientation vérifiant les hypothèses du théorème 2, et dont l'action sur  $\mathbb{T}$  est minimale, et soit  $\mu$  une mesure stationnaire ergodique pour  $f$ . On a alors l'alternative :*

- i) *Soit l'exposant de Lyapunov associé à  $\mu$  est strictement négatif.*
- ii) *Soit il existe un homéomorphisme  $g : \mathbb{T} \rightarrow \mathbb{T}$  tel que  $gf g^{-1}$  soit presque sûrement une rotation.*

### 2.3 Conséquence de la négativité des exposants de Lyapunov

Nous allons parler maintenant des conséquences dynamiques qu'entraînent l'existence d'exposants de Lyapunov négatifs. Plus précisément nous allons nous intéresser aux systèmes dynamiques aléatoires dont tous les exposants de Lyapunov associés à une mesure stationnaire sont strictement négatifs. C'est une hypothèse forte qui n'est pas entraînée par le théorème d'Avila-Viana, celui-ci impliquant seulement qu'en général, au moins un exposant de Lyapunov est négatif. Cependant, dans le cas d'un système dynamique aléatoire sur le tore  $\mathbb{T}$ , il n'y a qu'un seul exposant de Lyapunov associé à une mesure stationnaire donnée. Ainsi, le théorème d'Avila-Viana montrant qu'en général, cet exposant est strictement négatif, les théorèmes que nous allons présenter s'appliqueront, ce qui permet d'obtenir des phénomènes dynamiques généraux intéressants sur  $\mathbb{T}$ .

Nous allons présenter deux théorèmes. Le premier est dû à Y. Le Jan : le fait que tous les exposants de Lyapunov d'un système dynamique sont strictement négatifs entraîne l'existence de contractions locales, et Le Jan montre qu'il y a en fait un phénomène de contraction globale. Le deuxième théorème (n'existant pas dans la littérature à ma connaissance) étudie des propriétés de convergence des itérés de l'opérateur de transfert du système.

Énonçons le contexte. Soit  $V$  une variété riemannienne compacte munie d'une métrique  $d$ ,  $f : V \rightarrow V$  un difféomorphisme aléatoire  $C^1$  vérifiant  $\mathbb{E}[|\ln(\|Df\|_\infty)| + |\ln\|Df^{-1}\|_\infty|] < +\infty$ , et  $\mu$  une mesure stationnaire ergodique, et  $\gamma_1 > \dots > \gamma_d$  les exposants de Lyapunov associés. Soit  $(f_n)$  une suite i.i.d. de copies de  $f$ , et on note  $F_n = f_{n-1} \circ \dots \circ f_0$ .

Avant d'énoncer les théorèmes en question, énonçons la propriété de contractions locales sur laquelle sont basés les preuves :

**Proposition 2.** *Si  $\gamma_1 < 0$ , alors il existe  $q$  dans  $]0, 1[$  tel que pour  $\mu$ -presque tout  $x$ , il existe presque sûrement un voisinage  $U$  de  $x$  et une constante  $C > 0$  tels que pour tout  $y$  dans  $U$  et pour tout entier  $n$ ,  $d(F_n(y), F_n(x)) \leq Cq^n d(x, y)$ .*

*Démonstration.* Afin d'éviter le formalisme des variétés riemanniennes, nous traiterons seulement le cas  $V = \mathbb{T}^d$ . Supposons dans un premier temps que

$$\mathbb{E} \int_{\mathbb{T}^d} \ln \|Df(x)\| d\mu(x) < 0.$$

Pour  $x$  dans  $V$  et  $\varepsilon > 0$ , on pose  $M(x, \varepsilon) = \sup_{y \in B(x, \varepsilon)} \|Df(y)\|$ . Comme  $-\ln\|Df^{-1}\| \leq \ln M(\cdot, \varepsilon) \leq \ln\|Df\|$  et  $\lim_{\varepsilon \rightarrow 0} M(x, \varepsilon) = \|Df(x)\|$ , par convergence dominée on peut trouver  $\varepsilon > 0$  tel que

$$\mathbb{E} \int_{\mathbb{T}^d} \ln M(x, \varepsilon) d\mu(x) < 0.$$



### 2.3. CONSÉQUENCE DE LA NÉGATIVITÉ DES EXPOSANTS DE LYAPUNOV 23

Soit  $x$  générique pour  $\mu$ , et  $x_n = F_n(x)$ . Le théorème de Birkhoff appliqué à  $\ln M(\cdot, \varepsilon)$  montre que, choisissant  $q$  tel que  $0 > \ln q > \mathbb{E} \int_{\mathbb{T}^d} \ln M(z, \varepsilon) d\mu(z)$ , il existe presque sûrement une constante  $C$  telle que pour tout  $n$ ,

$$\prod_{k=0}^{n-1} M(x_k, \varepsilon) \leq Cq^n$$

Soit maintenant  $y$  proche de  $x$ ,  $y_n = F_n(y)$  et  $\varepsilon_n = d(x_n, y_n)$ . Par l'inégalité des accroissements finis,  $\varepsilon_{n+1} \leq M(x_n, \varepsilon_n)\varepsilon_n$ , donc par récurrence,

$$\varepsilon_n \leq \varepsilon_0 \prod_{k=0}^{n-1} M(x_k, \varepsilon_k).$$

Si on suppose que  $\varepsilon_0, \dots, \varepsilon_{n-1}$  sont inférieurs à  $\varepsilon$ , alors

$$\varepsilon_n \leq \varepsilon_0 \prod_{k=0}^{n-1} M(x_k, \varepsilon) \leq C\varepsilon_0 q^n,$$

donc en particulier  $\varepsilon_n \leq \varepsilon$  si  $C\varepsilon_0 \leq \varepsilon$ . Ainsi, si  $C\varepsilon_0 \leq \varepsilon$ , on obtient par récurrence que  $\varepsilon_n \leq \varepsilon$  pour tout  $n$ , puis que  $\varepsilon_n \leq C\varepsilon_0 q^n$ .

Traitons maintenant le cas général. La quantité  $\mathbb{E} \int_{\mathbb{T}^d} \frac{\ln \|DF_p(x)\|}{p} d\mu(x)$  tend vers l'exposant de Lyapunov caractéristique du système qui est strictement négatif, donc il existe  $p$  tel que

$$\mathbb{E} \int_{\mathbb{T}^d} \ln \|DF_p(x)\| d\mu(x) < 0.$$

Soit  $q$  tel que  $0 > \ln q > \mathbb{E} \int_{\mathbb{T}^d} \ln \|DF_p(x)\| d\mu(x) < 0$ , soit  $x$  générique pour  $\mu$ ,  $y$  proche de  $x$  et posons  $x_n = F_n(x)$ ,  $y_n = F_n(y)$ . En appliquant ce qui précède à  $F_p$  appliqués au points  $x_0, \dots, x_{p-1}$ , on en déduit l'existence d'une constante  $C$  telle que si  $d(x_r, y_r)$  est assez petit pour  $r = 1, \dots, p-1$ , alors tout entier  $k$ ,  $d(x_{kp+r}, y_{kp+r}) \leq Cq^n d(x, y)$ . Par suite, écrivant un entier  $n$  sous la forme  $n = kp + r$  avec  $r < p$ , on a  $d(x_n, y_n) \leq Cq^{\frac{n}{p}} \sup_{0 \leq r \leq p-1} d(x_k, y_k) \leq CMq^{\frac{n}{p}} d(x, y)$  où  $M = \prod_{k=0}^{p-1} \|Df_k\|_{\infty}$ , ce qui conclut la preuve.  $\square$

Enonçons maintenant le théorème de Le Jan.

**Théorème 3.** (*Y. Le Jan [11]*)

Si  $\gamma_1 < 0$ , il existe un entier  $p$  et un réel  $q$  dans  $]0, 1[$  tel que presque sûrement, on peut trouver des ouverts connexes disjoints  $U_1, \dots, U_p$  tels que pour tout  $x, y$  dans  $U_k$ ,  $\text{dist}(F_n(x), F_n(y)) \leq Cq^n$  pour un certain  $C > 0$ , et vérifiant de plus  $\mu(U_k) = \frac{1}{p}$  (en particulier,  $U_1 \cup \dots \cup U_p$  est de  $\mu$ -mesure totale).

Nous ne donnerons pas la preuve intégrale, pour laquelle nous renvoyons à [11], seulement les idées.

On a vu qu'un système dynamique aléatoire peut être vu comme un système dynamique usuel en regardant le cocycle sur  $\Omega^{\mathbb{N}} \times V$  défini par  $T(\omega, x) = (\tau\omega, f^{\omega_0}(x))$ , où  $f$  est vue comme une variable aléatoire  $\omega_0 \mapsto f^{\omega_0}$  définie sur un certain espace probabilisé  $(\Omega, \mathbb{P})$ . Ce cocycle est  $\mathbb{P}^{\mathbb{N}} \times \mu$ -invariant. Mais si on regarde maintenant le même cocycle, mais défini sur  $\Omega^{\mathbb{Z}} \times V$ , il n'est plus vrai en général que  $\mathbb{P}^{\mathbb{Z}} \times \mu$  est une mesure invariante (c'est en fait le cas si et seulement si  $\mu$  est presque sûrement invariante par  $f$ ).

Il existe cependant une méthode pour exhiber une mesure invariante sur  $\Omega^{\mathbb{Z}} \times V$  se projetant sur  $\Omega^{\mathbb{N}} \times V$ , dont l'étude permet d'obtenir des conséquences dynamiques sur le système, et est le point clé de la démonstration du théorème 3. Pour ce faire, considérons pour  $\omega$  dans  $\Omega^{\mathbb{Z}}$  la suite de mesures

$$\mu_{\omega}^n = (f^{\omega_{-1}} \circ \dots \circ f^{\omega_{-n}})_* \mu.$$

Pour  $\varphi : V \rightarrow \mathbb{R}$  continue, on vérifie que la suite de variables aléatoires  $M_n(\omega) = \int_V \varphi d\mu_{\omega}^n$  est une martingale bornée pour la filtration  $\mathcal{F}_n = \sigma(\{\omega_k, k \geq -n\})$ , donc converge presque sûrement, et la limite définit presque sûrement une forme linéaire sur  $\mathcal{C}(V, \mathbb{R})$  positive et valant 1 en 1, donc une probabilité notée  $\mu_{\omega}$ . On vérifie que les  $\mu_{\omega}$  satisfont les relations  $\mu_{\tau\omega} = f^{\omega_0} * \mu_{\omega}$  et  $\mathbb{E}[\mu_{\omega}] = \mu$ , puis on en déduit que  $d\pi(\omega, x) = d\mu_{\omega}(x) d\mathbb{P}^{\mathbb{Z}}(\omega)$  est une probabilité invariante pour le cocycle  $T$ . Sa projection sur  $\Omega^{\mathbb{N}} \times V$  est  $\mathbb{P}^{\mathbb{N}} \times \mu$ , et on peut montrer que l'ergodicité de  $\mu$  entraîne celle de  $\pi$ .

Il s'avère que l'hypothèse de négativité des exposants de Lyapunov entraîne que  $\mu_{\omega}$  est presque sûrement constituée uniquement de masses de Dirac. Pour montrer cela, il faut étudier la probabilité  $d\mu_{\omega}(x) d\mu_{\omega}(y) d\mathbb{P}^{\mathbb{Z}}(\omega)$  sur  $\Omega^{\mathbb{Z}} \times V \times V$ . Pour  $\varphi : V \times V \rightarrow \mathbb{R}$ , on a

$$\int \varphi(x, y) d\tilde{\pi}(\omega, x, y) = \int \varphi(F_n^{\omega}(x), F_n^{\omega}(y)) d\tilde{\pi}(\omega, x, y).$$

L'idée est de choisir  $\varphi$  à support dans un voisinage de la diagonale  $D = \{(x, x), x \in V\}$ , et d'utiliser que presque sûrement et pour  $\mu$  presque tout  $x$ ,  $d(F_n(x), F_n(y))$  tend vers 0 si  $y$  est suffisamment proche de  $x$ , afin de montrer que  $\int \varphi d\tilde{\pi}$  ne dépend essentiellement que des valeurs de  $\varphi$  sur  $D$ . Cette méthode permet de montrer que pour presque tout  $\omega$  et pour presque tout  $\mu_{\omega}$  tout  $x$ , on a  $\mu_{\omega}(B(x, \delta) - \{x\}) = 0$  si  $\delta$  est assez petit, et cela implique que  $\mu_{\omega}$  est atomique.

L'ergodicité de  $\pi$  permet ensuite de montrer que le nombre d'atomes de  $\mu_{\omega}$  est fini et ne dépend pas de  $\omega$ , et que tous ces atomes ont même masse. Finalement pour un événement  $\omega$  générique fixé, notant  $x_1, \dots, x_p$  les atomes de  $\mu_{\omega}$ , et notant  $U_p$  la composante connexe de  $\{x \in V \mid d(F_n^{\omega}(x), F_n^{\omega}(x_p)) \rightarrow 0\}$

### 2.3. CONSÉQUENCE DE LA NÉGATIVITÉ DES EXPOSANTS DE LYAPUNOV 25

contenant  $x_p$ , les  $U_p$  sont des ouverts vérifiant les propriétés souhaitées.

Dans le cas où  $V = \mathbb{T}$  et  $f$  est minimale, les ouverts  $U_p$  sont des intervalles, et  $U_1 \cup \dots \cup U_p$  est de  $\mu$ -mesure pleine donc dense puisqu'on a vu que quand  $f$  est minimale  $\mu$  est à support total. Ainsi,  $\mathbb{T} - (U_1 \cup \dots \cup U_p)$  est nécessairement fini de cardinal  $p$ , et en utilisant le corollaire 1, on obtient l'alternative :

**Corollaire 2.** Soit  $f : \mathbb{T} \rightarrow \mathbb{T}$  un difféomorphisme aléatoire  $C^1$  préservant l'orientation tel que  $\mathbb{E}[\ln \|f'\|_\infty] < +\infty$ . Alors :

i) soit il existe un homéomorphisme  $g$  de  $\mathbb{T}$  tel que  $gf g^{-1}$  soit presque sûrement une rotation.

ii) soit il existe presque sûrement un ensemble fini  $F$  de cardinal indépendant de l'aléa tel que pour tout  $x, y$  appartenant à la même composante connexe de  $\mathbb{T} - F$ ,  $|F_n(x) - F_n(y)|$  converge vers 0 exponentiellement vite.

Retournant au contexte général d'un système dynamique aléatoire sur une variété compacte dont tous les exposants de Lyapunov sont strictement négatifs, nous allons maintenant étudier l'opérateur de transfert  $T$  associé à  $f$ , défini par  $T\varphi = \mathbb{E}[\varphi \circ f]$ .

**Théorème 4.** Si  $\gamma_1 < 0$  et si de plus l'action de  $f$  est minimale, alors pour toute fonction  $\varphi : V \rightarrow \mathbb{R}$  continue (resp dans  $L^p(\mu)$ ),  $T^n\varphi$  converge vers  $\int_V \varphi d\mu$  uniformément (resp dans  $L^p(\mu)$ ). En particulier  $\mu$  est uniquement ergodique.

Notons qu'il suffit de traiter le cas où  $\varphi$  est continue, le cas où  $\varphi$  est dans  $L^p(\mu)$  s'en déduisant par densité des fonctions continues et le fait que  $\|T^n\psi\|_{L^p(\mu)} \leq \|\psi\|_{L^p(\mu)}$ . La démonstration s'effectue en plusieurs étapes. Premièrement, nous utilisons la proposition 2 pour obtenir que pour  $\mu$ -presque tout  $x$  de  $V$ , la suite  $(F_n)$  est presque sûrement contractante au voisinage de  $x$ . Ensuite, en utilisant la minimalité de  $f$ , nous montrerons que cette propriété est en fait vérifiée pour tout  $x$ . Cela nous permettra alors de montrer que la famille  $(T^n\varphi)$  est équicontinue, puis nous montrerons finalement que toute valeur d'adhérence de cette famille est constante égale à  $\int_V \varphi d\mu$ .

Posons  $U$  l'ouvert (aléatoire) constitué des  $x$  dans  $V$  tels que  $d(F_n, F_n(x))$  converge uniformément vers 0 sur un voisinage de  $x$ . Par la proposition 2, on a que pour  $\mu$ -presque tout  $x$ ,  $x \in U$  presque sûrement, et on va montrer qu'en fait cette propriété est vérifiée pour tout  $x$  de  $V$ . Pour cela, nous utiliserons le lemme suivant :

**Lemme 2.** Si  $\phi$  une fonction semicontinue supérieurement vérifiant  $T\phi = \phi$ , alors  $\phi$  est constante.

*Démonstration.* Soit  $M = \sup \phi$  et  $F = \phi^{-1}(\{M\})$ . Comme  $\phi$  est semicontinue,  $F$  est fermé et non vide, et si  $x$  est dans  $F$ , on a

$$M = \phi(x) = \mathbb{E}[\phi(f(x))] \leq M$$

donc  $\phi(f(x)) = M$  presque sûrement, donc  $f(x)$  appartient à  $F$  presque sûrement. Par minimalité,  $F = V$  et  $\phi$  est constante.  $\square$

Montrons maintenant la propriété souhaitée :

**Lemme 3.** *Pour tout  $x$  de  $U$ ,  $\mathbb{P}(x \notin U) = 0$*

*Démonstration.* Posons  $\phi(x) = \mathbb{P}(x \notin U)$ , et vérifions que  $\phi$  satisfait les hypothèses du lemme précédent. L'ouvert  $U$  ne dépend que de la suite  $\omega = (f_0, f_1, f_2, \dots)$ , notons le  $U(\omega)$ . Pour tout  $x$  de  $V$ ,  $x \in U(\omega) \Leftrightarrow f_0(x) \in U(\tau\omega)$ , donc

$$\phi(x) = \mathbb{P}(\{\omega | f_0(x) \notin U(\tau\omega)\}) = \mathbb{E}[\mathbb{P}(\omega' | f_0(x) \notin U(\omega'))] = \mathbb{E}[\phi(f_0(x))] = T\phi(x).$$

Montrons maintenant que  $\phi$  est semicontinue supérieurement. Soit donc  $c$  un réel positif et  $(x_n)_{n \in \mathbb{N}}$  une suite d'éléments de  $V$  convergeant vers  $x$  telle que  $\phi(x_n) \geq c$  et vérifions que  $\phi(x) \geq c$ . On a  $\mathbb{P}(x_n \notin U) \geq c$  pour tout  $n$ , donc  $\mathbb{P}(\limsup\{x_n \notin U\}) \geq c$ . Ainsi avec probabilité supérieure à  $c$ , une infinité de  $x_n$  appartient à  $V - U$ . Puisque  $(x_n)$  converge vers  $x$  et  $V - U$  est fermé, on en déduit que  $x$  appartient à  $V - U$  avec probabilité supérieure à  $c$ , et donc  $\phi(x) \geq c$ .

Ainsi,  $\phi$  est  $T$ -invariante et semicontinue supérieurement, donc par le lemme précédent,  $\phi$  est constante. Comme de plus  $\phi$  est nulle  $\mu$  presque partout par la proposition 2, cette constante est 0 et  $\phi$  est identiquement nulle.  $\square$

Le lemme précédent nous permettra de montrer que si  $\varphi$  est continue sur  $V$  alors  $(T^n\varphi)$  est équicontinue. Le lemme suivant nous sera utile pour montrer que les seules valeurs d'adhérence sont constantes.

**Lemme 4.** *Si  $\varphi$  est une fonction continue sur  $V$ , alors pour tout réel  $c$  la suite  $(\|T^n\varphi\|_{L^1(\mu)})$  est décroissante et si on a l'égalité  $\|T^n\varphi\|_{L^1(\mu)} = \|\varphi\|_{L^1(\mu)}$  pour un entier  $n$  alors presque sûrement  $\varphi \circ F_n$  a même signe que  $T^n\varphi$  sur  $V$*

*Démonstration.* Pour tout entier  $p$ ,

$$\|T^p\varphi\|_{L^1(\mu)} = \int_V |T^p\varphi| d\mu \leq \int_V T^p|\varphi| d\mu = \int_V \varphi d\mu = \|\varphi\|_{L^1(\mu)}$$

En particulier  $p = 1$  donne  $\|T\varphi\|_{L^1(\mu)} \leq \|\varphi\|_{L^1(\mu)}$  donc  $(\|T^n\varphi\|_{L^1(\mu)})$  est décroissante. De plus si dans la chaîne d'inégalités précédentes les termes extrémaux sont égaux, alors on doit avoir  $|T^p\varphi| = T^p|\varphi|$   $\mu$ -presque partout. Comme  $\mu$  est à support total (car le support de  $\mu$  est presque sûrement invariant par  $f$ ), on a en fait par continuité que  $|T^p\varphi| = T^p|\varphi|$  sur  $V$ . Donc pour tout  $x$  de  $V$ ,  $|\mathbb{E}[\varphi \circ F_p(x)]| = \mathbb{E}[|\varphi \circ F_p(x)|]$ , ce qui n'est possible que si  $\varphi \circ F_p(x)$  a presque sûrement même signe que  $\mathbb{E}[\varphi \circ F_p(x)] = T^p\varphi(x)$ . Ainsi, en fixant un sous ensemble dénombrable dense  $D$  de  $V$ , la fonction  $T^n\varphi \cdot \varphi \circ F_n$  est presque sûrement positive sur  $D$ , donc sur  $V$ .  $\square$

Démontrons maintenant le théorème 4. Soit donc  $\varphi$  une fonction continue sur  $V$ . Fixons  $x$  dans  $V$  et  $\varepsilon > 0$ . Il existe  $\delta > 0$  tel que

$$\forall y, z \in V, d(y, z) < \delta \Rightarrow |\varphi(y) - \varphi(z)| < \varepsilon.$$

### 2.3. CONSÉQUENCE DE LA NÉGATIVITÉ DES EXPOSANTS DE LYAPUNOV 27

Par le lemme 3, il existe presque sûrement un voisinage de  $x$  tel que  $d(F_n, F_n(x))$  converge uniformément vers 0 sur ce voisinage. On en déduit que l'on peut presque sûrement trouver  $r > 0$  tel que

$$\forall y \in V, d(x, y) < r \Rightarrow d(F_n(x), F_n(y)) < \delta$$

On peut donc trouver  $r > 0$  ne dépendant pas de l'aléa et un ensemble  $A$  de probabilité supérieure à  $1 - \varepsilon$  tel que pour tout événement  $\omega$  de  $A$ ,

$$d(x, y) < r \Rightarrow d(F_n^\omega(x), F_n^\omega(y)) < \delta \Rightarrow |\varphi(F_n^\omega(x)) - \varphi(F_n^\omega(y))| < \varepsilon$$

Donc si  $d(x, y) < r$ ,

$$\begin{aligned} |T^n \varphi(x) - T^n \varphi(y)| &= |\mathbb{E}[\varphi(F_n(x)) - \varphi(F_n(y))]| \\ &\leq \varepsilon P(A) + 2\|\varphi\|_0 P(A^c) \\ &\leq (1 + 2\|\varphi\|_0)\varepsilon \end{aligned}$$

Ainsi,  $(T^n \varphi)$  est équicontinue en  $x$  pour tout  $x$  de  $V$ . Par compacité de  $V$ ,  $(T^n \varphi)$  est uniformément équicontinue sur  $V$ .

Soit maintenant  $\psi$  une valeur d'adhérence de  $(T^n \varphi)$ , disons  $\psi = \lim_{k \rightarrow +\infty} T^{m_k} \varphi$ . Si  $p$  est un entier positif, par le lemme 4, on a pour tout  $k$  l'inégalité

$$\|T^{m_k} \varphi\|_{L^1(\mu)} \leq \|T^{m_k+p} \varphi\|_{L^1(\mu)} \leq \|T^{m_k+p} \varphi\|_{L^1(\mu)}$$

Quand  $k$  tend vers  $+\infty$ , on obtient  $\|T^p \psi\|_{L^1(\mu)} = \|\psi\|_{L^1(\mu)}$  donc toujours par le lemme 4, on a presque sûrement l'équivalence

$$\forall x \in V, \psi \circ F_p(x) \geq 0 \Leftrightarrow T^p \psi(x) \geq 0$$

Si  $c$  est un réel quelconque, on a aussi  $\psi - c = \lim_{k \rightarrow \infty} T^{m_k}(\varphi - c)$ , donc en appliquant l'argument précédent à  $\psi - c$ , on obtient que presque sûrement, pour tout rationnel  $c$

$$\forall x \in V, \psi \circ F_p(x) \geq c \Leftrightarrow T^p \psi(x) \geq c,$$

et on en déduit donc que presque sûrement,  $\psi \circ F_p = T^p \psi$  pour tout  $p$ .

D'autre part, si on pose  $m_k = n_{k+1} - n_k$ , en écrivant  $T^{m_k} \psi = T^{m_{k+1}} \varphi + T^{m_k}(\psi - T^{m_k} \varphi)$ , et en utilisant que  $\|T^{m_k}\|_{C(V, \mathbb{R})} \leq 1$ , on obtient que  $T^{m_k} \psi$  tend uniformément vers  $\psi$  quand  $k \rightarrow +\infty$ .

Fixons maintenant  $x$  dans  $V$ . Presque sûrement, si  $y$  est suffisamment proche de  $x$ ,  $d(F_n(x), F_n(y))$  tend vers 0 quand  $n$  tend vers  $+\infty$ , et  $\psi = \lim_{k \rightarrow +\infty} T^{m_k} \psi = \lim_{k \rightarrow +\infty} \psi \circ F_{m_k}$ , donc  $\psi(x) = \psi(y)$ . On a montré que  $\psi$  est constante au voisinage de tout point de  $V$ , donc constante sur  $V$ . Et comme  $T$  est  $\mu$ -invariant, cette constante vaut nécessairement  $\int_V \varphi d\mu$ . Ainsi,  $\int_V \varphi d\mu$  est la seule valeur d'adhérence de  $(T^n \varphi)$ , ce qui prouve le théorème.

Comme pour le théorème de Le Jan, on peut obtenir un corollaire intéressant dans le cas  $V = \mathbb{T}$  grâce au corollaire 1 du théorème de Avila-Viana :

**Corollaire 3.** Soit  $f : \mathbb{T} \rightarrow \mathbb{T}$  un difféomorphisme aléatoire minimal préservant l'orientation tel que  $\mathbb{E}[|\ln(\|f'\|_\infty) + |\ln(\|f'^{-1}\|_\infty)|] < +\infty$ , et soit  $T$  l'opérateur de transfert associé. Alors, sauf si il existe  $g$  un homéomorphisme de  $\mathbb{T}$  tel que  $g^{-1} \circ f$  appartient presque sûrement à un sous-groupe fini de  $\text{Hom}(\mathbb{T})$ ,  $f$  admet une seule mesure stationnaire  $\mu$ , et pour tout  $\varphi : \mathbb{T} \rightarrow \mathbb{R}$  continue,  $T^n \varphi$  converge uniformément vers  $\int_{\mathbb{T}} \varphi d\mu$ .

*Démonstration.* Soit  $\mu$  une mesure stationnaire ergodique de  $f$ . Si l'exposant de Lyapunov  $\gamma$  associé est strictement négatif, alors le résultat est une conséquence du théorème précédent. Si ce n'est pas le cas, alors par le corollaire 1, quitte à conjuguer  $f$  on peut supposer que  $f$  est une rotation  $r_\alpha$ . Si  $\varphi(x) = \sum_{k=-d}^d c_k e^{2i\pi kx}$  est un polynôme trigonométrique, alors  $T^n \varphi(x) = \sum_{k=-d}^d c_k \mathbb{E}[e^{2ik\pi\alpha}]^n e^{2i\pi kx}$ , et pour  $k \neq 0$ ,  $|\mathbb{E}[e^{2ik\pi\alpha}]| < 1$  sauf si  $\alpha$  est presque sûrement de la forme  $\alpha = \alpha_0 + \frac{j}{k}$  où  $\alpha_0$  est non aléatoire et  $j$  est un entier, ce qu'on suppose ne pas arriver par hypothèse. On a donc que  $T^n \varphi$  tend uniformément vers  $c_0 = \int_{\mathbb{T}} \varphi(x) dx$ , et cela reste vrai si  $\varphi$  est continue par densité.  $\square$

## Chapitre 3

# Théorèmes de conjugaison locale, théorie KAM

Nous nous intéressons ici à des théorèmes de conjugaison locale sur  $\mathbb{T}^d$  : étant donné  $f$  un difféomorphisme, déterministe ou aléatoire, de  $\mathbb{T}^d$ , un moyen de comprendre sa dynamique est d'essayer de conjuguer  $f$  à un difféomorphisme plus simple, les difféomorphismes les plus simples étant les translations  $r_\alpha(x) = x + \alpha$ . Un tel problème de conjugaison est compliqué en toute généralité, même s'il existe beaucoup de travaux en ce sens dans le cas  $d = 1$ . Ce type de problème est cependant plus aisé à traiter dans un cadre perturbatif : supposant que  $f$  est proche d'une translation, on peut essayer de conjuguer  $f$  à cette translation en cherchant une conjugaison proche l'identité (appelée conjugaison locale), à l'aide de méthodes perturbatives. Ce type de théorème de conjugaison locale et les méthodes perturbatives utilisées forment une branche des systèmes dynamiques appelée théorie KAM, en l'honneur des mathématiciens Kolmogorov, Arnold et Moser qui furent les premiers à explorer cette voie.

### 3.1 Le théorème de conjugaison locale d'Arnold

Regardons dans un premier temps le cas d'un difféomorphisme  $f$  de  $\mathbb{T}$ , que l'on suppose suffisamment régulier, et proche d'une rotation en un sens suffisamment fort que l'on précisera plus tard. Le but est de conjuguer  $f$  à une rotation par un difféomorphisme proche de l'identité et régulier. Notant  $\alpha = \rho(f)$ , on peut aisément montrer qu'il existe  $x$  dans  $\mathbb{T}$  tel que  $f(x) = r_\alpha(x)$ , donc si le difféomorphisme  $f$  est proche d'une rotation, il est nécessairement proche de  $r_\alpha$ . De plus, si  $f$  est conjugué à une rotation, on sait que cette rotation est nécessairement  $r_\alpha$ .

On peut donc formaliser le problème comme suit. Ecrivant  $f(x) = x + \alpha + \zeta(x)$ ,  $\zeta$  étant une perturbation d'ordre  $\varepsilon$ , on cherche  $g(x) = x + \eta(x)$ ,  $\eta$  étant aussi d'ordre  $\varepsilon$ , tel que  $gfg^{-1} = r_\alpha$ . Voyant le terme de gauche comme une fonctionnelle en  $(f, g)$ , l'idée est de simplifier l'équation en linéarisant ce terme au voisinage de  $(r_\alpha, Id)$ , négligeant de la sorte les termes d'ordre  $\varepsilon^2$  : si  $U(x) = x + \alpha + u(x)$  et  $V(x) = x + \beta + v(x)$ , alors

$$U \circ V(x) = x + \alpha + \beta + u \circ V(x) + v(x),$$

et par le théorème des accroissements finis,

$$|u(V(x)) - u(x + \beta)| \leq \|u\|_{C^1(\mathbb{T})} \|v\|_{C^0(\mathbb{T})},$$

donc

$$|U \circ V(x) - (x + \alpha + \beta + u(x + \beta) + v(x))| \leq \|u\|_{C^1(\mathbb{T})} \|v\|_{C^0(\mathbb{T})}. \quad (3.1)$$

De plus, si  $\|u\|_{C^1(\mathbb{T})} \leq \frac{1}{2}$ , pour  $\alpha = \beta = 0$  et  $V = U^{-1}$  on obtient

$$|u(x) + v(x)| \leq \|u\|_{C^1(\mathbb{T})} \|v\|_{C^0(\mathbb{T})} \leq \frac{1}{2} \|v\|_{C^0(\mathbb{T})},$$

ce qui entraîne  $\|v\|_{C^0(\mathbb{T})} \leq \|u\|_{C^0(\mathbb{T})} + \frac{1}{2} \|v\|_{C^0(\mathbb{T})}$ , donc  $\|v\|_{C^0(\mathbb{T})} \leq 2\|u\|_{C^0(\mathbb{T})}$  et

$$|U^{-1}(x) - (x - u(x))| = |u(x) + v(x)| \leq 2\|u\|_{C^1(\mathbb{T})} \|u\|_{C^0(\mathbb{T})}. \quad (3.2)$$

En utilisant les estimées (3.1) et (3.2), on obtient la linéarisation

$$gfg^{-1}(x) = x + \alpha + \zeta(x) + \eta(x + \alpha) - \eta(x) + O(\|\zeta\|_{C^1(\mathbb{T})} \|\eta\|_{C^1(\mathbb{T})}), \quad (3.3)$$

sous la condition  $\|\zeta\|_{C^1(\mathbb{T})} \leq \frac{1}{2}$ . En particulier, si  $\|\zeta\|_{C^1(\mathbb{T})}, \|\eta\|_{C^1(\mathbb{T})} = O(\varepsilon)$  alors le reste est d'ordre  $\varepsilon^2$ . Ainsi, plutôt que de résoudre l'équation  $gfg^{-1}(x) = x + \alpha$ , on se ramène à résoudre l'équation approchée

$$\zeta(x) = \eta(x) - \eta(x + \alpha).$$

Cette équation implique que  $\int_{\mathbb{T}} \zeta(t) dt = 0$ , et n'a donc pas de solution si ce n'est pas le cas. Modifiant l'équation en conséquence, nous allons plutôt résoudre

$$\zeta(x) - \int_{\mathbb{T}} \zeta(t) dt = \eta(x) - \eta(x + \alpha). \quad (3.4)$$

Autrement dit, plutôt que de chercher une solution approchée à  $gfg^{-1} = r_\alpha$ , on cherche une solution approchée à  $gfg^{-1} = r_\beta$  où  $\beta = \alpha + \int_{\mathbb{T}} \zeta(t) dt$ , et du fait que  $\alpha = \rho(f)$ , on saura après coup que  $\beta$  est en fait proche de  $\alpha$ .

L'équation (3.4) peut se résoudre par développement en séries de Fourier. Si on impose  $\hat{\eta}(0) = 0$ , l'unique solution possible est donnée par

$$\eta(x) = \sum_{p \in \mathbb{Z}^*} \frac{\zeta(p)}{1 - e^{2i\pi p \alpha}} e^{2i\pi p x}.$$



On a alors à faire à un problème de "petits diviseurs". Afin que cette série soit bien définie, il faut évidemment que les dénominateurs  $1 - e^{2i\pi p\alpha}$  soient non nuls, ce qui est équivalent au fait que  $\alpha$  soit irrationnel, mais ce n'est pas suffisant pour assurer la convergence de la série. La fonction  $\zeta$  étant régulière, ses coefficients de Fourier de  $\zeta$  décroissent rapidement, mais une condition sur  $\alpha$  est nécessaire afin que les dénominateurs  $1 - e^{2i\pi p\alpha}$  ne soient pas trop petits et que les coefficients de Fourier de  $\eta$  décroissent eux aussi rapidement.

La condition demandée sur  $\alpha$  sera d'être diophantien. Cela signifie qu'il existe des réels strictement positifs  $A$  et  $\sigma$  tels que pour tout  $(p, q)$  appartenant à  $\mathbb{Z} \times \mathbb{N}^*$ ,

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{A}{q^\sigma}.$$

Cette condition implique

$$|e^{2i\pi p\alpha} - 1| \geq |\sin 2\pi p\alpha| \geq \frac{4|p|^\sigma}{A},$$

et en conséquence, si les coefficients de Fourier de  $\zeta$  décroissent rapidement, il en est de même de ceux de  $\eta$ . Par exemple, si  $j$  et  $k$  sont deux entiers tels que  $j \leq k - \sigma - 2$ ,

$$\|\eta\|_{C^j(\mathbb{T})} \leq \sum_{p \in \mathbb{Z}} |p|^j |\hat{\eta}(p)| \leq \frac{4}{A} \sum_{p \in \mathbb{Z}^*} |p|^{j+\sigma} |\hat{\zeta}(p)| \leq \left( \frac{4}{A} \sum_{p \in \mathbb{Z}^*} \frac{1}{|p|^{j-k+\sigma}} \right) \|\zeta\|_{C^k(\mathbb{T})}. \quad (3.5)$$

Cela permet ainsi de résoudre l'équation approchée (3.4) et donc de trouver  $g$  une perturbation de  $Id$  en  $\varepsilon$ , tel que  $gf g^{-1}$  soit égal à une rotation  $r_\beta$  à  $\varepsilon^2$  près et donc égal à  $r_\alpha$  à  $\varepsilon^2$  près puisque  $\rho(f) = \alpha$ .

Notant  $f_1 = gf g^{-1}$ , il suffit maintenant de réappliquer la méthode à  $f_1$  pour conjuguer celle-ci à un difféomorphisme  $f_2$  approchant  $r_\alpha$  avec un reste  $O(\varepsilon^4)$ . L'itération de cette méthode permet de récupérer une suite de conjugaison  $(g_k)$  telle que  $(g_{n-1} \circ \dots \circ g_0) \circ f \circ (g_0^{-1} \circ \dots \circ g_{n-1}^{-1}) = r_\alpha + O(\varepsilon^{2^n})$ . De plus,  $g_n = Id + O(\varepsilon^{2^n})$ , ce qui permet de montrer pour  $\varepsilon$  suffisamment petit la convergence de  $g_{n-1} \circ \dots \circ g_0$  vers un difféomorphisme  $h = Id + O(\varepsilon)$  satisfaisant  $hfh^{-1} = r_\alpha$ .

Hormis certains calculs délibérément omis, il reste un problème dans la démarche précédente dû à la perte de régularité : Une estimation initiale de  $\|f - r_\alpha\|_{C^k(\mathbb{T})}$  donne une estimation de  $\|f_1 - r_\alpha\|_{C^j(\mathbb{T})}$  pour un entier  $j$  strictement inférieur à  $k$ , et en conséquence la convergence de l'itération n'est pas assurée. Un moyen de remédier à cela est de remplacer la fonction  $\eta$  par l'approximation  $C^\infty$

$$\tilde{\eta}(x) = \sum_{|p| \leq \lambda} \hat{\eta}(p) e^{2i\pi p x}$$

où  $\lambda$  est bien choisi, ce qui permet d'avoir une estimation un peu moins bonne de  $f_1 - r_\alpha$ , mais pour une meilleure norme, ce qui suffit à assurer la convergence du procédé.

V. Arnold est le premier à avoir obtenu un théorème de conjugaison de ce type sur  $\mathbb{T}$ . Il ne considérait pas les espaces  $C^k(\mathbb{T})$  comme dans la description précédente, mais les espaces  $C_r^\omega(\mathbb{T})$  constitué des fonctions  $2\pi$ -périodiques se prolongeant analytiquement sur la bande  $\{z = x + iy | x \in \mathbb{R}, y \in [-r, r]\}$  et muni de la norme  $\|\cdot\|_{C_r^\omega}(\mathbb{T})$  de la borne supérieure sur cette bande. Il obtint le théorème suivant :

**Théorème 5.** [2] *Soit  $\alpha$  un nombre diophantien, et  $r > s > 0$  deux réels. Il existe un réel  $\varepsilon = \varepsilon(\alpha) > 0$  et une constante  $C$  tels que si  $f$  est un difféomorphisme dans  $C_r^\omega(\mathbb{T})$  vérifiant  $\|f - r_\alpha\|_{C_r^\omega}(\mathbb{T}) \leq \varepsilon$ , alors il existe un difféomorphisme  $g$  dans  $C_s^\omega(\mathbb{T})$  vérifiant  $\|g - Id\|_{C_s^\omega}(\mathbb{T}) \leq C\varepsilon$  et  $gf g^{-1} = r_\alpha$ .*

Le théorème de conjugaison dans les espaces  $C^k(\mathbb{T})$  est dû à J. Moser.

### 3.2 Une généralisation d'un théorème de Moser

Nous nous proposons dans cette section d'aborder un problème de conjugaison locale dans un cadre plus rigoureux et plus général. Premièrement, nous regardons des difféomorphismes de  $\mathbb{T}^d$ , l'hypothèse sur le nombre de rotation étant remplacé par une hypothèse sur l'ensemble de rotation. D'autre part, nous traiterons le cas d'un difféomorphisme de  $\mathbb{T}^d$  aléatoire.

Il nous faut généraliser la notion de nombre diophantien dans ce cadre. Nous dirons qu'une variable aléatoire  $\alpha$  à valeurs dans  $\mathbb{R}^d$  est diophantienne si il existe des réels strictement positifs  $A$  et  $\sigma$  tels que pour tout  $q$  dans  $\mathbb{Z}^d - \{0\}$ ,

$$\| \{q \cdot x\} \|_{L^2(\Omega)} \geq \frac{A}{|q|^\sigma}$$

où  $u \cdot v$  désigne le produit scalaire de  $u$  et  $v$ , et  $\{t\}$  la distance de  $t$  à  $\mathbb{Z}$ . On peut vérifier que si  $\alpha$  a une loi non singulière par rapport à la mesure de Lebesgue, alors  $\alpha$  est diophantien, et si  $\alpha$  a une loi quelconque,  $\alpha + x$  est diophantien pour Lebesgue presque tout  $x$  de  $\mathbb{R}^d$ .

Nous allons établir le théorème suivant :

**Théorème 6.** *Soit  $\alpha$  une variable aléatoire sur  $\mathbb{R}^d$  diophantienne pour certaines constantes  $A$  et  $\sigma$ . Pour tout entier  $k, l$  existe  $\varepsilon = \varepsilon(A, \sigma) > 0$  et un entier  $K$  tel que pour tout difféomorphisme aléatoire  $f$  vérifiant :*

*i)  $\alpha \in \mathcal{R}(f)$  presque sûrement,*

*ii) Si  $\tilde{f}$  est une copie indépendante de  $f$ ,  $\tilde{f} \circ f = f \circ \tilde{f}$  presque sûrement,*

*iii)  $\|f - r_\alpha\|_{C^0} \leq 1$  presque sûrement et  $\mathbb{E}[\|f - r_\alpha\|_{C^k}^2]^{\frac{1}{2}} \leq \varepsilon$ ,*

il existe un difféomorphisme  $g$  (non aléatoire) tel que  $\|g - Id\|_{C^k(\mathbb{T}^d)} \leq C\varepsilon$  et  $gf g^{-1} = r_\alpha$  presque sûrement.

Ce résultat étend un résultat de Moser [14] traitant le cas où  $d = 1$  et où l'espace probabilisé associé à  $f$  est fini.

Commençons par énoncer une proposition donnant certaines inégalités fonctionnelles sur les espaces  $C^k(\mathbb{T}^d)$ .

**Proposition 3.**

i) Pour tout  $\varphi$  dans  $C^k(\mathbb{T}^d)$  et tous  $k_0 < k_\lambda < k_1$ ,

$$\|\varphi\|_{C^{k_\lambda}} \leq \|\varphi\|_{C^{k_0}}^{\frac{k_1 - k_\lambda}{k_1 - k_0}} \|\varphi\|_{C^{k_1}}^{\frac{k_\lambda - k_0}{k_1 - k_0}}.$$

ii) Pour tous  $\phi, \psi$  dans  $C^k(\mathbb{T}^d)$

$$\|\phi\psi\|_k \leq C(k)(\|\phi\|_{C^k}\|\psi\|_{C^0} + \|\phi\|_{C^0}\|\psi\|_{C^k})$$

iii) Pour tout  $\varphi$  dans  $C^k(\mathbb{T}^d)$  et  $g = Id + \psi$  dans  $Dif f_0^k(\mathbb{T}^d)$  tel que  $\|\psi\|_{C^1} \leq 1$ ,

$$\|\varphi \circ g\|_{C^k(\mathbb{T}^d)} \leq C(k)\|\varphi\|_{C^k}(1 + \|\psi\|_{C^k}).$$

iv) Pour tous  $f = Id + \phi$  et  $g = Id + \psi$  dans  $Dif f_0^k(\mathbb{T}^d)$  tel que  $\|\psi\|_{C^1} \leq 1$ ,

$$\|f \circ g - Id\|_{C^k} \leq C(k)(\|\psi\|_k + \|\phi\|_k)$$

v) Pour tout  $f = Id + \phi$  dans  $Dif f^k(\mathbb{T}^2)$  tel que  $\|\phi\|_{C^1} \leq \frac{1}{2}$ ,

$$\|f^{-1} - Id\|_{C^k} \leq C(k)\|f - Id\|_{C^k}$$

( $C(k)$  désigne une constante ne dépendant que de  $k$ )

La démarche pour prouver le théorème 6 est identique à celle du théorème d'Arnold. On pose  $f(x) = x + \alpha + \zeta(x)$ ,  $g(x) = x + \eta(x)$  et  $\varepsilon_k = \mathbb{E}[\|\zeta\|_{C^k(\mathbb{T}^d)}^2]^{\frac{1}{2}}$ . Si  $\frac{1}{2} > \varepsilon > 0$ , supposant que l'on a  $\max(\varepsilon_1, \|\eta\|_{C^1(\mathbb{T}^d)}) \leq \varepsilon$ , on peut effectuer par (3.3) la linéarisation

$$gf g^{-1}(x) = x + \alpha + \zeta(x) + \eta(x + \alpha) - \eta(x) + O_{L^2(\Omega)}(\varepsilon^2), \quad (3.6)$$

où  $O_{L^2(\Omega)}(\varepsilon^2)$  est un raccourci pour désigner un terme dont la norme dans  $L^2(\Omega)$  est majoré par  $C\varepsilon^2$  pour une certaine constante  $C$  ne dépendant que de  $A$  et  $\sigma$  (ici,  $C$  est en fait absolue). Il s'agit de choisir  $\eta$  de sorte que  $\zeta(x + \alpha) + \eta(x + \alpha) - \eta(x)$  soit proche d'une fonction constante avec grande probabilité. Cela donne donc un aléa d'équations fonctionnelles que  $\eta$  doit satisfaire. En général, il n'y aurait pas de solutions  $\eta$  à un tel problème, mais dans le contexte actuel, l'hypothèse de commutativité du système entraîne des relations entre les différentes équations fonctionnelles à résoudre, ce qui permet de ramener le problème à la résolution d'une seule équation fonctionnelle. Plus précisément,

on va définir  $\eta$  de sorte que  $\mathbb{E}[\zeta(x + \alpha) + \eta(x + \alpha) - \eta(x)]$  soit proche d'une constante, puis vérifier que  $\eta$  satisfait nos exigences.

Définissons  $\eta$  par

$$\eta(x) = \sum_{0 < |p| < \lambda} \frac{\mathbb{E}[\hat{\zeta}(p)]}{1 - \mathbb{E}[e^{2i\pi p\alpha}]} e^{2i\pi p x},$$

où  $\lambda > 1$  sera défini plus tard. Le fait que  $\alpha$  est diophantien donne la minoration des dénominateurs

$$|1 - \mathbb{E}[e^{2i\pi p\alpha}]| \geq \mathbb{E}[1 - \cos(2\pi p\alpha)] \geq 2\pi^2 \mathbb{E}[|p\alpha|^2] \geq \frac{2A^2\pi^2}{|p|^{2\sigma}},$$

et on en déduit que, posant  $r = 2\sigma + d + 1$ , on a pour tous entiers  $j, k$

$$\|\eta\|_{C^j(\mathbb{T}^d)} \leq \sum_{0 < |p| < \lambda} |p|^j |\hat{\eta}(p)| \leq \frac{1}{2A^2\pi^2} \sum_{0 < |p| < \lambda} |p|^{j+2\sigma} |\mathbb{E}[\hat{\zeta}(p)]| \leq \left( \frac{1}{2A^2\pi^2} \sum_{p \in \mathbb{Z}^d - \{0\}} \frac{1}{|p|^{d+1}} \right) \lambda^{j-k+r} \varepsilon_k.$$

Posons maintenant  $f_1 = g f g^{-1} = r_\alpha + \zeta_1$  et  $\varepsilon_k^{(1)} = \mathbb{E}[\|\zeta_1\|_{C^k(\mathbb{T}^d)}^2]^{1/2}$ . Nous allons chercher des majorations aux coefficients de Fourier de  $\zeta_1$  afin d'estimer  $\varepsilon_k^{(1)}$ . Par la proposition 3 et l'estimée ci-dessus, on a

$$\varepsilon_k^{(1)} \leq C(\varepsilon_k + \|\eta\|_{C^k(\mathbb{T}^d)}) \leq C\lambda^r \varepsilon_k. \quad (3.7)$$

pour une certaine constante  $C = C(k)$ . D'autre part, si  $\tilde{f} = Id + \tilde{\alpha} + \tilde{\zeta}$  est une copie indépendante de  $f$ , l'égalité  $\tilde{f} \circ f = f \circ \tilde{f}$  p.s. se linéarise en

$$\tilde{\zeta}(x + \alpha) - \tilde{\zeta}(x) = \zeta(x + \tilde{\alpha}) - \zeta(x) + O_{L^2(\Omega)}(\varepsilon_1^2).$$

Un développement en séries de Fourier donne alors

$$\hat{\zeta}(p)(1 - e^{2i\pi p\alpha}) = \hat{\zeta}(p)(1 - e^{2i\pi p\tilde{\alpha}}) + O_{L^2(\Omega)}(\varepsilon_1^2),$$

ce qui entraîne pour  $0 < |p| < \lambda$

$$\hat{\zeta}(p) = \frac{\mathbb{E}[\hat{\zeta}(p)]}{1 - \mathbb{E}[e^{2i\pi p\tilde{\alpha}}]} (1 - e^{2i\pi p\alpha}) + O_{L^2(\Omega)} \left( \frac{\varepsilon_1^2}{|1 - \mathbb{E}[e^{2i\pi p\tilde{\alpha}}]|} \right) = \hat{\eta}(p)(1 - e^{2i\pi p\alpha}) + O_{L^2(\Omega)}(|p|^{2\sigma} \varepsilon_1^2).$$

Ainsi, en utilisant de plus que

$$\zeta_1(x) = \zeta(x) + \eta(x + \alpha) - \eta(x) + O_{L^2(\Omega)}(\varepsilon_{r+1}^2)$$

(par la formule de linéarisation (3.6) et l'estimée  $\|\eta\|_{C^1(\mathbb{T}^d)} \leq C\varepsilon_{r+1}$ ), on obtient

$$\hat{\zeta}_1(p) = \hat{\zeta}(p) + (\hat{\eta}(p) - e^{2i\pi p\alpha} \hat{\eta}(p)) + O_{L^2(\Omega)}(\varepsilon_{r+1}^2) = O_{L^2(\Omega)}(|p|^{2\sigma} \varepsilon_{r+1}^2)$$

pour  $0 < |p| < \lambda$ . On en déduit que pour  $K \geq k \geq r + 1$ ,

$$\mathbb{E}[\|\zeta_1 - \hat{\zeta}_1(0)\|_{C^k(\mathbb{T}^d)}^2]^{\frac{1}{2}} \leq \sum_{p \in \mathbb{Z}^d - \{0\}} |p|^k \|\hat{\zeta}_1(p)\|_{L^2(\Omega)} \leq \sum_{|p| < \lambda} |p|^k \varepsilon_k^2 + \sum_{|p| \geq \lambda} \frac{\lambda^r \varepsilon_K}{|p|^{K-k}} \leq C \left( \lambda^{k+d} \varepsilon_k^2 + \frac{\varepsilon_K}{\lambda^{K-k-d-r}} \right)$$

pour une certaine constante  $C = C(K)$ . Egalisant les deux termes du membre de

droite, on choisit  $\lambda = \left( \frac{\varepsilon_K}{\varepsilon_k} \right)^{\frac{1}{K+2k+2d+r}}$ , on obtient  $\mathbb{E}[\|\zeta_1 - \hat{\zeta}_1(0)\|_{C^k(\mathbb{T}^d)}^2]^{\frac{1}{2}} \leq C \varepsilon_K^\delta \varepsilon_k^{2(1-\delta)}$

où on a posé  $\delta = \frac{2k+2d+r}{K+2k+2d+r}$ . Puisque presque sûrement  $\alpha \in \mathcal{R}(f) = \mathcal{R}(f_1)$ , il existe une mesure  $\nu$  invariante pour  $f$  telle que  $\int_{\mathbb{T}^d} \zeta_1 d\nu = 0$ , donc  $|\hat{\zeta}_1(0)| = |\hat{\zeta}_1(0) - \int_{\mathbb{T}^d} \zeta_1 d\nu| \leq \|\zeta_1 - \hat{\zeta}_1(0)\|_{C^k(\mathbb{T}^d)}$ , et on obtient ainsi

$$\varepsilon_k^{(1)} \leq C \varepsilon_K^\delta \varepsilon_k^{2(1-\delta)}. \quad (3.8)$$

Les inégalités (3.7) (avec  $K$  à la place de  $k$ ) et (3.8) mènent ainsi au système

$$\begin{cases} \varepsilon_K^{(1)} \leq C \varepsilon_K^{1+\delta_1} \varepsilon_k^{-2\delta_1} \\ \varepsilon_k^{(1)} \leq C \varepsilon_K^\delta \varepsilon_k^{2(1-\delta)} \end{cases},$$

où  $\delta_1 = \frac{r}{K+2k+2d+r}$ . Itérons cette démarche et construisons ainsi une suite de difféomorphismes  $f_1, f_2, \dots, f_n, \dots$  conjugués à  $f$ . On note  $f_n = r_\alpha + \zeta_n$  et  $\varepsilon_j^{(n)} = \mathbb{E}[\|\zeta_n\|_{C^j(\mathbb{T}^d)}^2]^{\frac{1}{2}}$ . Si  $f_n$  est construite, on peut construire  $f_{n+1} = g_n f_n g_n^{-1}$  par la démarche précédente à condition que  $\varepsilon_k^{(n)}$  soit assez petit, disons inférieur à  $c$ , de sorte que  $\|g_n - Id\|_{C^1(\mathbb{T}^d)} \leq \frac{1}{2}$ . Supposons  $(f_n)$  définie pour  $n < N$ :

$$\begin{cases} \varepsilon_K^{(n+1)} \leq C (\varepsilon_K^{(n)})^{1+\delta_1} (\varepsilon_k^{(n)})^{-2\delta_1} \\ \varepsilon_k^{(n+1)} \leq C (\varepsilon_K^{(n)})^\delta (\varepsilon_k^{(n)})^{2(1-\delta)}. \end{cases}$$

Vérifions que si  $K$  est assez grand, et  $\varepsilon_K^{(0)} = \varepsilon_K$  est assez petit, alors  $\varepsilon_k^{(n)}$  tend rapidement vers 0 quand  $n$  grandit.

Quand  $K$  est grand,  $\delta$  et  $\delta_1$  sont proches de 0, et la matrice  $\begin{pmatrix} 1 + \delta_1 & \delta \\ -2\delta_1 & 2(1 - \delta) \end{pmatrix}$  a donc une valeur propre  $q$  proche de 2, et un vecteur propre  $\begin{pmatrix} u \\ v \end{pmatrix}$  associé, normalisé par  $u + v = 1$ , proche de  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . En particulier,  $v$  est positif, et comme  $(1 + \delta_1)u + \delta v = qu$ ,  $u$  est aussi positif. La suite  $r_n = (\varepsilon_K^{(n)})^u (\varepsilon_k^{(n)})^v$  vérifie donc la relation de récurrence  $r_{n+1} \leq C r_n^q$ , donc si  $r_0$  est assez petit,  $r_n \leq C r_0 2^{-q^n}$ . Puisque  $\varepsilon_k^{(n)} \leq r_n$  et  $r_0 \leq \varepsilon_K^{(0)}$ , on en déduit donc que si  $\varepsilon_K^{(0)}$  est assez petit,  $\varepsilon_k^{(n)} \leq C \varepsilon_K^{(0)} 2^{-q^n}$ . En particulier, la condition  $\varepsilon_k^{(n)} \leq c$  est vérifiée pour tout  $n < N$ , ce qui montre par récurrence que l'on peut définir  $f_n$  pour tout entier  $n$ , et l'estimer précédente montre que  $f_n$  converge rapidement vers  $r_\alpha$  quand  $n$  tend

vers  $+\infty$ .

Pour finir, notons  $h_n = Id + \xi_n$  le difféomorphisme conjuguant  $f$  à  $f_n$ . Il est de la forme  $h_n = g_{n-1} \circ \dots \circ g_0$ , où  $g_m = Id + \eta_m$  vérifie  $\|\eta_m\|_{C^{k-r}(\mathbb{T}^d)} \leq C\varepsilon_k^{(m)} \leq C\varepsilon_K 2^{-q^m}$ . En particulier,  $\sum_m \|\eta_m\|_{C^{k-r}(\mathbb{T}^d)}$  converge, de somme inférieure à  $C\varepsilon_K$ . Ecrivant  $\xi_n = \sum_{m=0}^{n-1} \eta_m \circ h_m$ , on obtient par la proposition 3

$$\begin{aligned} \|\xi_n\|_{C^{k-r}(\mathbb{T}^d)} &\leq \sum_{m=0}^{n-1} \|\eta_m \circ h_m\|_{C^{k-r}(\mathbb{T}^d)} \\ &\leq C \sum_{m=0}^{n-1} \|\eta_m\|_{C^{k-r}(\mathbb{T}^d)} (1 + \|\xi_m\|_{C^{k-r}(\mathbb{T}^d)}) \\ &\leq C\varepsilon_K \sup_{m < n} (1 + \|\xi_m\|_{C^{k-r}(\mathbb{T}^d)}). \end{aligned}$$

On en déduit que  $\|\xi_n\|_{C^{k-r}(\mathbb{T}^d)} \leq C\varepsilon_K$ , puis que la série  $\sum_m \eta_m \circ h_m$  converge normalement dans  $C^{k-r}(\mathbb{T}^d)$ , et donc que  $(h_n)$  converge dans  $C^{k-r}(\mathbb{T}^d)$  vers un difféomorphisme  $h = Id + \xi$  vérifiant  $\|\xi\|_{C^{k-r}(\mathbb{T}^d)} \leq C\varepsilon_K$ , et  $h$  conjugue  $f$  à  $r_\alpha$ .

Nous proposons maintenant de démontrer une autre généralisation du théorème d'Arnold. Pour  $\alpha$  dans  $\mathbb{T}$ , on peut voir la propriété " $\alpha$  est diophantien" comme une propriété de répartition des points des orbites de  $r_\alpha$ . Notant  $d$  la métrique sur  $\mathbb{T}$  définie par  $d(x, y) = \{x - y\}$  (que l'on rappelle être la distance de  $x - y$  à  $\mathbb{Z}$ , qui ne dépend pas des représentants  $x$  et  $y$  choisis),  $\alpha$  est diophantien si et seulement il existe  $A$  et  $\sigma$  tel que pour tout  $n$  dans  $\mathbb{N}$  et pour tout  $x$  dans  $\mathbb{T}$ ,  $d(r_\alpha^n(x), x) \geq \frac{A}{n^\sigma}$ . On en déduit que si  $f$  est  $C^1$ -conjuguée à une rotation diophantienne, alors  $f$  vérifie pour tout  $x$  dans  $\mathbb{T}$  une inégalité du type  $d(f^n(x), x) \geq \frac{A}{n^\sigma}$ . On s'intéresse, dans un cadre perturbatif, à la réciproque. On va montrer que si  $f$  est proche d'une rotation et vérifie de telles inégalités, alors  $f$  est conjugquée à une rotation diophantienne (ce qui équivaut à  $\rho(f)$  diophantien par le théorème d'Arnold).

**Théorème 7.** Soient  $\sigma$  et  $A_0 > A_1 > A_2$  quatre réels strictement positifs. Il existe un entier  $k$  et un réel  $\varepsilon > 0$  tel que pour tout difféomorphisme  $f$  de  $\mathbb{T}$  préservant l'orientation, si  $\|f - r_\alpha\|_k \leq \varepsilon$  pour une certaine rotation  $r_\alpha$ , alors les propriétés suivantes sont équivalentes :

- i) Pour tout entier  $n$ ,  $\sup_{x \in \mathbb{T}} d(f^n(x), x) \geq \frac{A_0}{n^\sigma}$ .
- ii)  $\rho(f)$  est  $(A_1, \sigma)$ -diophantien.
- iii) Pour tout entier  $n$ ,  $\inf_{x \in \mathbb{T}} d(f^n(x), x) \geq \frac{A_2}{n^\sigma}$ .

**Remarque :** contrairement au théorème d'Arnold, ce théorème n'a pas d'analogue global. En effet, si  $f$  est un difféomorphisme de  $\mathbb{T}$  différent de l'identité et ayant des points fixes, on vérifie aisément que si  $x$  n'est pas un point fixe de  $f$  alors la suite  $(d(f^n(x), x))_n$  est minorée, ce qui montre que  $f$  vérifie i) pour certaines constantes  $A$  et  $\sigma$  tandis que  $\rho(f) = 0$ , donc n'est pas diophantien.

*Démonstration.* Notons  $\alpha = \rho(f)$ ,  $m_n(f) = \inf_{x \in \mathbb{T}} d(f^n(x), x)$  et  $M_n(f) = \sup_{x \in \mathbb{T}} d(f^n(x), x)$ .

Pour tout  $n$ ,  $\rho(f^n) = n\alpha \bmod 1$  donc il existe  $x$  tel que  $f^n(x) = x + n\alpha \bmod 1$ , donc tel que  $d(f^n(x), x) = \{n\alpha\}$ , ce qui implique les inégalités

$$m_n(f) \leq \{n\alpha\} \leq M_n(f),$$

d'où on déduit iii)  $\Rightarrow$  ii)  $\Rightarrow$  i). On suppose maintenant que  $f$  vérifie i), et nous allons montrer que si  $\varepsilon$  est assez petit, il existe une conjugaison  $g = Id + \eta$  vérifiant  $gfg^{-1} = r_\alpha$  et  $\|\eta\|_{C^1(\mathbb{T})} \leq 1 + C\varepsilon$  pour une certaine constante  $C$ , d'où on déduit  $(\inf_{\mathbb{T}} |g'|)M_n(f) \leq M_n(r_\alpha) = \{n\alpha\} = m_n(r_\alpha) \leq (\sup_{\mathbb{T}} |g'|)m_n(f)$  et donc

$$m_n(f) \geq \frac{1 - C\varepsilon}{1 + C\varepsilon} M_n(f),$$

ce qui entraîne i)  $\Rightarrow$  iii).

L'idée de la preuve est d'utiliser que si  $n$  n'est pas trop grand,  $f^n(x)$  est proche de  $r_\alpha^n(x)$ , et donc  $M_n(f)$  est proche de  $M_n(r_\alpha) = \{n\alpha\}$ . Ainsi, la condition iii) permet de minorer  $\{n\alpha\}$  et donc  $|e^{2i\pi n\alpha} - 1|$  si  $n$  n'est pas trop grand, ce qui permet de définir une première conjugaison  $g$  comme dans le théorème d'Arnold de sorte que  $gfg^{-1}$  soit plus proche de  $r_\alpha$  que  $f$ . On utilisera alors que  $gfg^{-1}$  vérifie encore iii), quitte à remplacer  $A_0$  par un réel un peu plus petit, et on pourra alors itérer le procédé.

Posons  $f(x) = x + \alpha + \zeta(x)$ ,  $\varepsilon = \|\zeta\|_{C^1(\mathbb{T})}$ , et  $\gamma = 1 + \|\zeta\|_{C^k(\mathbb{T})}$ , où  $k$  est un entier que nous choisirons plus tard. Soit  $g = Id + \eta$ , où

$$\eta(x) = \sum_{0 < |p| < \varepsilon^{-\delta}} \frac{\hat{\zeta}(p)}{1 - e^{2i\pi p\alpha}} e^{2i\pi px},$$

$\delta > 0$  étant un réel proche de 0.

De  $|f^n(x) - r_\alpha^n(x)| \leq n\varepsilon$ , on déduit  $\{n\alpha\} \geq d(f^n(x), x) - n\varepsilon$ , donc en prenant le sup sur  $x$ , pour  $n \leq \varepsilon^{-\delta}$  on a

$$\{n\alpha\} \geq M_n(f) - n\varepsilon \geq \frac{A_0}{n^\sigma} - n\varepsilon \geq A_0\varepsilon^{\sigma\delta} - \varepsilon^{1-\delta} \geq \frac{A_0}{2}\varepsilon^{\sigma\delta},$$

et donc

$$|1 - e^{2i\pi n\alpha}| \geq 4\{n\alpha\} \geq 2A_0\varepsilon^{\sigma\delta}.$$

On en déduit que pour tout entier  $j$ ,

$$\|\eta\|_{C^j(\mathbb{T})} \leq \frac{C}{A_0} \sum_{0 < |p| < \varepsilon^{-\delta}} |p|^j \varepsilon^{-\sigma\delta} |\hat{\zeta}(p)| \leq \frac{C}{A_0} \varepsilon^{-(1+\sigma)\delta} \|\zeta\|_{C^j(\mathbb{T})},$$

où  $C = C(j, \sigma)$ .

Posons maintenant  $f_1 = gf g^{-1} = r_{\alpha_1} + \zeta_1$ , où  $\alpha_1 = \alpha + \hat{\zeta}(0)$ , et  $\varepsilon_1 = \|\zeta_1\|_{C^1(\mathbb{T})}$ ,  $\gamma_1 = 1 + \|\zeta_1\|_{C^k(\mathbb{T})}$ . Par la proposition 3,

$$\gamma_1 \leq C(\gamma + \|\eta\|_{C^k(\mathbb{T})}) \leq C(1 + A_0^{-1})\varepsilon^{-(1+\sigma)\delta}\gamma \quad (3.9)$$

où  $C = C(k, \sigma)$ . D'autre part, par la formule de linéarisation (3.3) et l'estimée  $\|\eta\|_{C^1(\mathbb{T})} \leq CA_0^{-1}\varepsilon^{1-(1+\sigma)\delta}$  on a

$$\zeta_1(x) = (gf g^{-1})(x) - \alpha_1 = (\zeta(x) - \hat{\zeta}(0)) + (\eta(x + \alpha) - \eta(x)) + O(A_0^{-1}\varepsilon^{2-(1+\sigma)\delta}),$$

donc pour  $0 < |p| < \varepsilon^{-\delta}$ ,

$$\begin{aligned} \hat{\zeta}_1(p) &= \hat{\zeta}(p) + (e^{2i\pi p\alpha} - 1)\hat{\eta}(p) + O(A_0^{-1}\varepsilon^{2-(1+\sigma)\delta}) \\ &= \hat{\zeta}(p) + (e^{2i\pi p\alpha} - 1)\left(\frac{\hat{\zeta}(p)}{1 - e^{2i\pi p\alpha}}\right) + O(A_0^{-1}\varepsilon^{2-(1+\sigma)\delta}) \\ &= O(A_0^{-1}\varepsilon^{2-(1+\sigma)\delta}). \end{aligned}$$

Ainsi,

$$\begin{aligned} \varepsilon_1 &\leq 2\pi \sum_{p \neq 0} |p| |\hat{\zeta}_1(p)| \\ &\leq CA_0^{-1} \sum_{0 < |p| < \varepsilon^{-\delta}} |p| \varepsilon^2 + C \sum_{|p| \geq \varepsilon^{-\delta}} \frac{\gamma_1}{|p|^{k-1}} \\ &\leq C(1 + A_0^{-1})(\varepsilon^{2-(3+\sigma)\delta} + \varepsilon^{(k-3-\sigma)\delta}\gamma), \end{aligned} \quad (3.10)$$

où  $C = C(k, \sigma)$ . Supposant avoir choisi  $k = 1 + [2\delta^{-1}]$ , de sorte que  $(k - 3 - \sigma)\delta \geq 2 - (3 + \sigma)\delta$ , les inégalités (3.9) et (3.10) mènent au système

$$\begin{cases} \gamma_1 \leq C(1 + A_0^{-1})\gamma\varepsilon^{-(1+\sigma)\delta} \\ \varepsilon_1 \leq C(1 + A_0^{-1})\gamma\varepsilon^{2-(3+\sigma)\delta} \end{cases}$$

Notons de plus que

$$M_n(f_1) \geq \inf_{\mathbb{T}} |g'| \frac{A_0}{n^\sigma} \geq (1 - C\varepsilon) \frac{A_0}{n^\sigma}.$$

Par itération du procédé, on peut trouver une suite  $(f_n)$  de conjuguées de  $f$  telles que, notant  $f_n = r_{\alpha_n} + \zeta_n$ ,  $\varepsilon_n = \|\zeta_n\|_{C^1(\mathbb{T})}$  et  $\gamma_n = 1 + \|\zeta_n\|_{C^k(\mathbb{T})}$ ,  $f_{n+1}$  est défini si  $\varepsilon_n$  est assez petit et les suites  $\varepsilon_n, \gamma_n$  vérifient

$$\begin{cases} \gamma_{n+1} \leq C_n \gamma_n \varepsilon_n^{-(1+\sigma)\delta} \\ \varepsilon_{n+1} \leq C_n \gamma_n \varepsilon_n^{2-(3+\sigma)\delta}, \end{cases}$$

où  $C_n = C(1 + \prod_{k=0}^{n-1} (1 - C\varepsilon_k)^{-1})$ ,  $C$  étant une constante ne dépendant que de  $\delta$  et  $\sigma$ . On applique maintenant la même méthode que pour le théorème 6. Si



$\delta$  a été choisi assez proche de 0, la matrice  $\begin{pmatrix} 1 & 1 \\ -(1+\sigma)\delta & 2-(3+\sigma)\delta \end{pmatrix}$  admet une valeur propre  $q$  proche de 2, avec un vecteur propre associé  $\begin{pmatrix} u \\ v \end{pmatrix}$  tel que  $u, v \geq 0, u + v = 1$ . La suite  $r_n = \gamma_n^u \varepsilon_n^v$  vérifie  $r_{n+1} \leq C_n r_n^q$ . Si on suppose  $C_k \leq 2C_0$  pour  $k = 0, \dots, n$ , et si  $r_0 \leq \gamma_0$  est assez petit, alors  $\varepsilon_k \leq r_k \leq C\gamma_0 2^{-q^k}$  pour une constante  $C$  ne dépendant pas de  $k$ , et ceci implique  $C_{n+1} \leq 2C_0$  si  $\gamma_0$  est assez petit. Ainsi, on obtient par récurrence  $C_n \leq 2$  et  $\varepsilon_n \leq C\gamma_0 2^{-q^n}$ .

En procédant toujours comme pour le théorème 6, écrivant  $f_{n+1} = g_n f_n g_n^{-1}$  où  $g_n = Id + \eta_n$  on a  $\|\eta_n\|_{C^1(\mathbb{T})} \leq C\varepsilon_n^{1-(1+\sigma)\delta}$ , donc  $\|\eta_n\|_{C^1(\mathbb{T})}$  décroît rapidement vers 0, et on en déduit finalement que la suite de conjugaisons  $h_n = g_{n-1} \circ \dots \circ g_0$  converge dans  $C^1(\mathbb{T})$  vers un difféomorphisme  $h$  vérifiant  $\|h - Id\|_{C^1(\mathbb{T})} \leq C\varepsilon$  et  $hfh^{-1} = r_\alpha$ .

□



## Chapitre 4

# Théorèmes de conjugaison lisse simultanée, cocycles d'Anzai

Dans les 2 chapitres suivants, nous montrerons des théorèmes de rigidité analogues à ceux du chapitre 2 dans un cadre perturbatif, ce qui permet l'utilisation de techniques KAM. Nous allons décrire dans ce chapitre certaines techniques et certains objets que nous utiliserons à cette fin.

Le premier résultat de rigidité lisse utilisant des techniques KAM est dû à D. Dolgopyat et R. Krikorian [6] et dit que si  $d > 1$ , si  $\omega$  est une variable aléatoire uniforme sur  $\{1, \dots, p\}$ , si  $R_1, \dots, R_p$  sont des rotations de  $S^d$  telles que l'action de  $R_\omega$  est minimale, et si  $f_1, \dots, f_p$  sont des difféomorphismes de  $S^d$  suffisamment  $C^\infty$ -proches de  $R_1, \dots, R_p$  tels que les  $d$  exposants de Lyapunov de  $f_\omega$  associés à une mesure stationnaire sont positifs, alors  $f_\omega$  est conjugué à une rotation par une conjugaison  $C^\infty$ -proche de l'identité.

Nous établirons donc par la suite un résultat analogue sur  $\mathbb{T}^d$  pour  $d = 1$  et 2. Si  $f$  est une perturbation d'un difféomorphisme aléatoire  $f_0$  d'une variété  $V$  dont la mesure de Lebesgue est l'unique mesure stationnaire, une première étape consiste à approximer les mesures stationnaires de  $f$ , afin d'en déduire ensuite une approximation des exposants de Lyapunov associés. Comme  $f$  est une perturbation de  $f_0$ , il est naturel de penser que ses mesures stationnaires sont proches de la mesure de Lebesgue en un certain sens, et ceci est effectivement le cas si, notant  $T_0$  l'opérateur de transfert de  $f_0$ , on a un bon contrôle de la vitesse de convergence des sommes de Birkhoff de  $T_0$  vers la mesure de Lebesgue. Dans le cadre du théorème de Dolgopyat et Krikorian où  $f_0 = R_\omega$ , une telle estimation est effectivement établie (en fait, il est même établi dans ce cadre une estimation de la vitesse de convergence de  $T_0^n$  vers la mesure de Lebesgue). Dans le cas où  $f_0$  est une translation aléatoire de  $\mathbb{T}^d$ , l'étude des sommes de Birkhoff de  $T_0$  fait

naturellement apparaître des petits diviseurs, et il est nécessaire d'avoir une condition diophantienne pour pouvoir estimer cette vitesse de convergence.

## 4.1 Mesures invariantes de perturbations de translations

**Proposition 4.** Soit  $\alpha$  un  $d$ -uplet  $(A, \sigma)$ -diophantien, dans le sens où pour tout  $p$  dans  $\mathbb{Z}$  et tout  $q$  dans  $\mathbb{Z}^d - \{0\}$ ,

$$|q \cdot \alpha - p| \geq \frac{A}{|q|^\sigma},$$

et soit  $r_\alpha$  la translation  $Id + \alpha$ . Si  $k$  est supérieur à  $\sigma + d + 1$ ,

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} \varphi \circ r_\alpha^n - \int_{\mathbb{T}^d} \varphi(x) dx \right| \leq \frac{C}{AN} \|\varphi\|_{C^k(\mathbb{T}^d)},$$

où  $C$  est une constante ne dépendant que de  $d$ .

*Démonstration.* Posons  $\varphi_N = \frac{1}{N} \sum_{n=0}^{N-1} \varphi \circ r_\alpha^n$ . Pour tout  $p$  dans  $\mathbb{Z}^d - \{0\}$ ,

$$\hat{\varphi}_N(p) = \left( \frac{1}{N} \sum_{n=0}^{N-1} e^{2i\pi n p \cdot \alpha} \right) \hat{\varphi}(p) = \frac{1}{N} \frac{1 - e^{2i\pi N p \cdot \alpha}}{1 - e^{2i\pi p \cdot \alpha}} \hat{\varphi}(p),$$

et la condition diophantienne sur  $\alpha$  donne la minoration

$$|1 - e^{2i\pi p \cdot \alpha}| \geq |\sin(2\pi p \cdot \alpha)| \geq \frac{2}{\pi} (2\pi \{p \cdot \alpha\}) \geq \frac{4A}{|p|^\sigma},$$

donc

$$|\hat{\varphi}_N(p)| \leq \frac{1}{2AN} |p|^\sigma |\hat{\varphi}(p)|.$$

On en déduit

$$|\varphi_N - \hat{\varphi}(0)| \leq \sum_{p \in \mathbb{Z}^d - \{0\}} |\hat{\varphi}_N(p)| \leq \frac{1}{2AN} \sum_{p \in \mathbb{Z}^d - \{0\}} |p|^\sigma |\hat{\varphi}(p)| \leq \left( \frac{1}{2AN} \sum_{p \in \mathbb{Z}^d - \{0\}} \frac{1}{|p|^{d+1}} \right) \|\varphi\|_{C^k(\mathbb{T}^d)}$$

□

Avec cette estimée sur la vitesse de convergence des sommes de Birkhoff d'une translation diophantienne  $r_\alpha$ , on peut déduire que les mesures invariantes d'un difféomorphisme proche de  $r_\alpha$  sont proches de la mesure de Lebesgue.

**Corollaire 4.** Soit  $\alpha$  un  $d$ -uplet  $(A, \sigma)$  diophantien,  $f$  un difféomorphisme de  $\mathbb{T}^d$  et  $\mu$  une mesure  $f$ -invariante. En posant  $\varepsilon = \|f - r_\alpha\|_{C^k(\mathbb{T}^d)}$  où  $k \geq \sigma + d + 1$ , on a pour toute fonction  $\varphi \in C^k$  sur  $\mathbb{T}^d$

$$\left| \int_{\mathbb{T}^d} \varphi(x) d\mu(x) - \int_{\mathbb{T}^d} \varphi(x) dx \right| \leq C \sqrt{\frac{\varepsilon}{A}} \|\varphi\|_{C^k(\mathbb{T}^d)}$$

où  $C$  est une constante ne dépendant que de  $d$ .

#### 4.1. MESURES INVARIANTES DE PERTURBATIONS DE TRANSLATIONS 43

*Démonstration.* Par récurrence immédiate,  $|f^n - r_\alpha^n| \leq n\varepsilon$ , donc  $|\varphi \circ f^n - \varphi \circ r_\alpha^n| \leq n\varepsilon \|\varphi\|_{C^1(\mathbb{T}^d)}$ . Ainsi, en utilisant cette inégalité et le lemme précédent,

$$\left| \frac{1}{N} \sum_{n=0}^{N-1} \varphi \circ f^n - \int_{\mathbb{T}^d} \varphi(x) dx \right| \leq \frac{C}{AN} \|\varphi\|_{C^k(\mathbb{T}^d)} + N\varepsilon \|\varphi\|_{C^1(\mathbb{T}^d)}.$$

On intègre cette inégalité selon  $\mu$ , ce qui donne par  $f$ -invariance de  $\mu$

$$\left| \int_{\mathbb{T}^d} \varphi(x) d\mu(x) - \int_{\mathbb{T}^d} \varphi(x) dx \right| \leq \left( \frac{C}{AN} + N\varepsilon \right) \|\varphi\|_{C^k(\mathbb{T}^d)},$$

et on obtient le résultat voulu en choisissant  $N = \left\lceil \frac{1}{\sqrt{A\varepsilon}} \right\rceil + 1$ . □

**Remarque :** Comme application immédiate de ce résultat, on obtient par exemple que le diamètre de l'ensemble de rotation de  $f$  est  $O(\varepsilon^{\frac{3}{2}})$  quand  $\varepsilon$  tends vers 0.

Considérons  $f$  un difféomorphisme aléatoire lisse de  $\mathbb{T}$ ,  $\mu$  une mesure stationnaire ergodique de  $f$ , et  $\gamma$  l'exposant de Lyapunov correspondant. Dans ce contexte, on a

$$\gamma = \mathbb{E} \int_{\mathbb{T}} \ln f'(x) d\mu(x).$$

On veut justifier qu'en général,  $\gamma \leq 0$  et étudier le cas d'égalité. Supposons que la mesure stationnaire  $\mu$  est la mesure de Lebesgue. Alors on a directement par l'inégalité de Jensen

$$\gamma = \mathbb{E} \int_{\mathbb{T}} \ln(f'(x)) dx \leq \ln \left( \mathbb{E} \int_{\mathbb{T}} f'(x) dx \right) \leq \ln 1 = 0,$$

avec égalité si et seulement si  $f' = 1$  presque sûrement, donc si et seulement si  $f$  est presque sûrement une rotation. Supposons maintenant que  $\mu$  est une densité  $g(x)dx$  où  $g$  est lisse et ne s'annule pas. Si on note  $G(x) = \int_0^x g(t)dt$ , alors  $G_*\mu$  est la mesure de Lebesgue, donc  $GfG^{-1}$  admet la mesure de Lebesgue comme mesure stationnaire, et en utilisant l'invariance des exposants de Lyapunov par conjugaison, on obtient de nouveau que  $\gamma \leq 0$ , avec égalité si et seulement si le difféomorphisme lisse  $G$  conjugue  $f$  à une rotation.

Il est en général difficile d'obtenir la régularité d'une mesure stationnaire. Cependant, si  $f$  est proche d'une rotation aléatoire vérifiant une condition diophantienne, on peut montrer de manière analogue au corollaire 4 que  $\mu$  est proche de la mesure de Lebesgue, et en affinant la méthode on peut approcher  $\mu$  par une densité  $g(x)dx$  avec une précision supérieure,  $g$  étant une densité lisse proche de 1. Adaptant alors l'argument du paragraphe ci-dessus, on peut montrer qu'en notant  $G(x) = \int_0^x g(t)dt$ ,  $\gamma$  est strictement

négatif excepté si  $GfG^{-1}$  est très proche d'une rotation. On peut alors à l'instar du théorème d'Arnold et ses dérivés présentés au chapitre 2 itérer cette démarche pour conjuguer  $f$  à des difféomorphismes de plus en plus proches d'une rotation, démarche que l'on peut itérer à l'infini pour conjuguer  $f$  à une rotation, à moins que  $\gamma$  ne soit strictement négatif. L'énoncé et la démonstration précise de ce fait heuristique sera l'objet du prochain chapitre.

Le cas de la dimension 2, c'est à dire l'étude des exposants de Lyapunov de difféomorphismes proches de translations, et plus précisément l'étude du cas critique où tous les exposants de Lyapunov sont nuls, fait naturellement intervenir une classe de difféomorphismes plus large que celle des translations, que nous allons présenter dans les deux prochaines sections.

## 4.2 Cocycles d'Anzai, unique ergodicité

Nous noterons  $\mathcal{G}$  le groupe des homéomorphismes sur  $\mathbb{T}^2$  de la forme  $(x, y) \rightarrow (x + \alpha, y + \beta(x))$ , où  $\alpha$  est dans  $\mathbb{T}$  et  $\beta$  est une fonction continue sur  $\mathbb{T}$ . Nous appellerons cocycles d'Anzai de tels homéomorphismes. Nous allons étudier quelques propriétés dynamiques des éléments de  $\mathcal{G}$ .

Si  $R(x, y) \rightarrow (x + \alpha, y + \beta(x))$  et  $U(x, y) = (x, y + u(x))$ , alors

$$URU^{-1}(x, y) = (x + \alpha, y + \beta(x) + u(x + \alpha) - u(x)),$$

donc s'il existe une fonction  $u$  sur  $\mathbb{T}$  tel que pour tout  $x$  dans  $\mathbb{T}$   $u(x + \alpha) - u(x) = \beta(x) - \int_{\mathbb{T}} \beta(t) dt$  (ce qui est par exemple le cas si  $\alpha$  est diophantien et  $\beta$  suffisamment lisse), alors  $R$  est conjugué à une translation. Ainsi un cocycle d'Anzai suffisamment régulier hérite des propriétés dynamiques d'une translation. Il est en revanche plus rare qu'un cocycle d'Anzai aléatoire soit conjugué à une translation aléatoire car ceci implique que presque toutes les réalisations de ce cocycle commutent, ce qui n'est en général pas le cas.

Il est trivial que la mesure de Lebesgue est stationnaire pour tout cocycle d'Anzai aléatoire. La proposition suivante donne un critère pour que ce soit la seule.

**Proposition 5.** *Soit  $R : (x, y) \mapsto (x + \alpha, y + \beta(x))$  un cocycle d'Anzai aléatoire, et  $\hat{\beta} = \int_{\mathbb{T}} \beta(x) dx$ . On suppose que pour tout entier  $p$ ,  $p\alpha$  n'est pas presque sûrement un entier. Les assertions suivantes sont alors équivalentes :*

- i) La mesure de Lebesgue est l'unique mesure stationnaire de  $R$ .*
- ii)  $R$  est ergodique pour la mesure de Lebesgue.*
- iii) Il n'existe pas d'entier  $p$  non nul et de fonction mesurable  $u : \mathbb{T} \rightarrow \mathbb{R}$  ( $p$  et  $u$  non aléatoires) tel que presque sûrement, pour tout  $x$  dans  $\mathbb{T}$ ,*

$$p\beta(x) + u(x + \alpha) - u(x) \in \mathbb{Z}.$$

*Démonstration.* i)  $\Rightarrow$  ii) est un fait trivial classique, et iii) équivaut à la non existence d'une fonction  $u$  et d'un entier  $p$  tels que  $(x, y) \mapsto e^{2i\pi(u(x)+py)}$  soit presque sûrement  $R$ -invariante, donc ii)  $\Rightarrow$  iii). Supposons maintenant iii), et soit  $\mu$  une mesure stationnaire de  $R$ . On va montrer que  $\mu$  est la mesure de Lebesgue en suivant une méthode due à Anzai [1]. Pour  $q$  dans  $\mathbb{Z}$ , posons  $\mu_q$  la mesure complexe sur  $\mathbb{T}$  définie par

$$\int_{\mathbb{T}} \varphi(x) d\mu_q(x) = \int_{\mathbb{T}} \varphi(x) e^{-2i\pi qy} d\mu(x, y).$$

L'invariance de  $\mu$  par  $R$  donne

$$\int_{\mathbb{T}} \varphi(x) d\mu_q(x) = \mathbb{E} \int_{\mathbb{T}} \varphi(x + \alpha) e^{-2i\pi q\beta(x)} d\mu_q(x), \quad (4.1)$$

Nous allons montrer que  $\mu_0 = dx$  et que  $\mu_q = 0$  pour  $q \neq 0$ , ce qui suffit pour conclure puisque cela implique  $\hat{\mu}(p, q) = 0$  pour  $(p, q) \neq (0, 0)$ .

Pour  $q = 0$ ,  $\mu_0$  est invariante par l'opérateur de transfert  $T\varphi(x) = \mathbb{E}[\varphi(x + \alpha)]$ , ce qui implique  $\hat{\mu}_0(p) = \mathbb{E}[e^{ip\alpha}] \hat{\mu}_0(p)$  pour tout entier  $p$ , mais par hypothèse si  $p \neq 0$ ,  $|\mathbb{E}[e^{2i\pi p\alpha}]| < 1$  donc  $\hat{\mu}_0(p) = 0$ , et  $\hat{\mu}_0(0) = \int 1 d\mu = 1$ , donc  $\mu_0$  est la mesure de Lebesgue sur  $\mathbb{T}$ .

On suppose maintenant  $q \neq 0$ . On écrit  $\mu_q$  sous la forme  $d\mu_q(x) = e^{i\theta_q(x)} d|\mu_q(x)|$ , où  $|\mu_q|$  est une mesure positive finie et  $\theta_q$  une fonction borelienne à valeurs dans  $]-\pi, \pi]$ . La relation (4.1) implique

$$\int_{\mathbb{T}} \varphi(x) d\mu_q(x) \leq \int_{\mathbb{T}} T\varphi(x) d|\mu_q|(x),$$

ce qui en remplaçant  $\varphi$  par  $\varphi/g$  donne pour  $\varphi \geq 0$

$$\left| \int_{\mathbb{T}} \varphi(x) d\mu_q(x) \right| \leq \int_{\mathbb{T}} |T\varphi(x)| d|\mu_q|(x).$$

Par suite, si  $|\varphi| \leq M$ , appliquant la formule précédente aux fonctions positives  $M + \varphi$  et  $M - \varphi$  donne par double inégalité

$$\int_{\mathbb{T}} \varphi(x) d|\mu_q|(x) = \int_{\mathbb{T}} T\varphi(x) d|\mu_q|(x).$$

Ainsi  $|\mu_q|$  est invariante par  $T$ , donc le même raisonnement que pour l'étude de  $\mu_0$  montre que  $d|\mu_q| = c_q dx$  pour un certain réel positif  $c_q$ . Il nous reste à montrer que  $c_q = 0$ . Pour cela remplaçons  $\mu_q$  par  $c_q e^{i\theta_q(x)} dx$  dans (4.1) :

$$\begin{aligned} c_q \int_{\mathbb{T}} \varphi(x) e^{i\theta_q(x)} dx &= c_q \mathbb{E} \int_{\mathbb{T}} \varphi(x + \alpha) e^{i(-2\pi q\beta(x) + \theta_q(x))} dx \\ &= c_q \int_{\mathbb{T}} \varphi(x) \mathbb{E}[e^{i(-2\pi\beta(x-\alpha) + \theta_q(x-\alpha))}] dx. \end{aligned}$$

La fonction  $\varphi$  étant arbitraire, si  $c_q \neq 0$ , cela entraîne que pour presque tout  $x$ ,  $e^{i\theta_q(x)} = \mathbb{E}[e^{i(-2\pi q\beta(x-\alpha) + \theta_q(x-\alpha))}]$ , et donc que

$$-q\beta(x-\alpha) + \frac{1}{2\pi}(\theta_q(x-\alpha) - \theta_q(x)) \in \mathbb{Z},$$

ce qui contredit iii) □

Grâce à ce critère, nous pouvons montrer maintenant qu'en général, un cocycle d'Anzai aléatoire est uniquement ergodique :

**Corollaire 5.** Soit  $R : (x, y) \mapsto (x + \alpha, y + \beta(x))$  un cocycle d'Anzai aléatoire tel que :

- i) Il n'existe pas d'entier  $p$  non nul tel que  $p\alpha \in \mathbb{Z}$  presque sûrement.
- ii) Si  $\tilde{R}$  est une copie indépendante de  $R$ , alors  $R \circ \tilde{R}$  n'est pas presque sûrement égal à  $\tilde{R} \circ R$ .

Alors  $R$  est ergodique pour la mesure de Lebesgue.

*Démonstration.* Supposons  $R$  non ergodique. Il existe par la proposition précédente un entier  $p$  non nul et une fonction mesurable  $u$  telle que presque sûrement, pour presque tout  $x$  dans  $\mathbb{T}$ ,

$$p\beta(x) = u(x) - u(x + \alpha) \pmod{1}$$

De même, si  $\tilde{R} : (x, y) \mapsto (x + \tilde{\alpha}, y + \tilde{\beta}(x))$  est une copie indépendante de  $R$ , alors presque sûrement, pour presque tout  $x$  dans  $\mathbb{T}$ ,

$$p\tilde{\beta}(x) = u(x) - u(x + \tilde{\alpha}) \pmod{1},$$

et on en déduit que la fonction  $x \mapsto \beta(x + \tilde{\alpha}) - \tilde{\beta}(x) + \beta(x) - \tilde{\beta}(x + \alpha)$  est presque sûrement à valeurs dans l'ensemble discret  $\frac{1}{p}\mathbb{Z}$ , donc constante par continuité, donc nulle car d'intégrale nulle. Ceci entraîne que  $R$  et  $\tilde{R}$  commutent presque sûrement. □

Si l'on suppose certaines conditions diophantiennes, alors l'assertion  $R \circ \tilde{R} = \tilde{R} \circ R$  presque sûrement entraîne une conjugaison de  $R$  à une translation, et donc là encore l'unique ergodicité.

Nous réutiliserons au chapitre 6 des techniques analogues pour montrer que les mesures stationnaires d'un difféomorphisme aléatoire proche d'un cocycle d'Anzai aléatoire diophantien sont proches de la mesure de Lebesgue.

### 4.3 Une caractérisation dynamique des cocycles d'Anzai généralisés

Nous nous intéressons aux cocycles d'Anzai car les exposants de Lyapunov d'un cocycle d'Anzai aléatoire sont tous nuls. Au chapitre 6, nous montrerons



#### 4.3. UNE CARACTÉRISATION DYNAMIQUE DES COCYCLES D'ANZAI GÉNÉRALISÉS 47

que dans un contexte perturbatif, les cocycles d'Anzai caractérisent, à conjugaision près, les difféomorphismes aléatoires dont les exposants de Lyapunov sont nuls. Nous montrons dans cette section un résultat allant en ce sens dans un autre contexte : nous allons montrer ici, dans un contexte global, que les cocycles d'Anzai caractérisent les difféomorphismes aléatoires ayant probabilité positive d'être une translation minimale et dont les exposants de Lyapunov sont nuls. Nous allons démontrer ce résultat dans un tore de dimension quelconque.

Commençons par généraliser la notion de cocycle d'Anzai en dimension supérieure. On pose  $\mathcal{G}_0$  le sous-groupe de  $\text{Hom}_0(\mathbb{T}^d)$  dont les éléments  $R$  sont de la forme  $R(x_1, x_2, \dots, x_d) = (x_1 + \alpha_1, x_2 + \alpha_2(x_1), \dots, x_d + \alpha_d(x_1, \dots, x_{d-1}))$  où  $\alpha_k$  est une fonction continue sur  $\mathbb{T}^{k-1}$ .

Si  $f : \omega \rightarrow f^\omega$  est un difféomorphisme aléatoire de  $\mathbb{T}^d$  sur un espace probabilisé  $(\Omega, \mathcal{F}, \mathbb{P})$ , on appellera restriction de  $f$  tout difféomorphisme aléatoire  $g$  de  $\mathbb{T}^d$  de la forme  $g = f|_A$  où  $A$  est un sous ensemble mesurable de  $\Omega$  de probabilité positive muni de la probabilité induite  $\frac{1}{\mathbb{P}(A)}\mathbb{P}$

**Théorème 8.** *Soit  $f$  dans  $\text{Diff}_0(\mathbb{T}^d)$  un difféomorphisme aléatoire  $C^1$ ,  $\mu$  une mesure stationnaire de  $f$ , et supposant  $\mathbb{E} \int_{\mathbb{T}^d} \ln^+ \|df(x)\| + \ln^+ \|df(x)^{-1}\| d\mu(x) < +\infty$ , soit  $\gamma_1 \geq \dots \geq \gamma_d$  les exposants de Lyapunov associés à  $\mu$ .*

*On suppose qu'il existe une restriction  $R$  de  $f$  qui soit une translation aléatoire minimale. Les assertions suivantes sont alors équivalentes :*

- i) Il existe une action algébrique (déterministe)  $P$  dans  $Sl_d(\mathbb{Z})$  telle que  $PfP^{-1}$  appartient presque sûrement à  $\mathcal{G}_0$ .*
- ii)  $f$  appartient presque sûrement à un sous groupe résoluble (déterministe)  $\mathcal{G}$  de  $\text{Diff}_0(\mathbb{T}^d)$ .*
- iii)  $\gamma_d(\mu) \geq 0$ .*

*(remarque : l'hypothèse d'intégrabilité servant à définir les exposants de Lyapunov n'est pas nécessaire pour l'équivalence i)  $\Leftrightarrow$  ii.)*

*Lorsque ces assertions équivalentes sont vérifiées, on a en fait plus précisément que  $\gamma_i = 0$  pour  $i = 1, \dots, d$ , et  $f$  appartient presque sûrement à un sous groupe résoluble déterministe  $\mathcal{G}$  d'ordre de résolubilité au plus  $d$  (i.e. le  $d$ -ième groupe dérivé de  $\mathcal{G}$  est réduit à un singleton).*

Une récurrence directe montre que les éléments du  $k$ -ième sous groupe dérivé de  $\mathcal{G}_0$  sont de la forme  $(x_1, \dots, x_d) \rightarrow (x_1, \dots, x_k, x_{k+1} + \alpha_{k+1}(x_1, \dots, x_k), \dots, x_d + \alpha_d(x_1, \dots, x_{d-1}))$ , donc en particulier le  $d$ -ième sous-groupe dérivé de  $\mathcal{G}_0$  est trivial, donc  $\mathcal{G}_0$  est résoluble, ce qui montre i)  $\Rightarrow$  ii).

Démontrons i)  $\Rightarrow$  iii). Soit  $r : (x_1, x_2, \dots, x_d) \mapsto (x_1 + \alpha_1, x_2 + \alpha_2(x_1), \dots, x_d + \alpha_d(x_1, \dots, x_{d-1}))$  un difféomorphisme aléatoire de  $\mathcal{G}_0$  et  $\mu$  une mesure stationnaire ergodique,  $(r_k)$  une suite i.i.d. de copies de  $r$  et  $R_n = r_{n-1} \circ \dots \circ r_0$ . Les endomorphismes  $N_k = dr_k(R_k(x)) - Id$  sont triangulaires sur la base canonique de  $\mathbb{R}^d$ , donc le produit de  $d$  endomorphismes quelconques parmi la famille  $(N_k)$

est nul. En conséquence :

$$dR_n(x) = (Id + N_{n-1}) \cdots (Id + N_0) = Id + \sum_{j=1}^d \sum_{0 \leq k_1 < \cdots < k_d \leq n} N_{k_d} \cdots N_{k_1},$$

et, notant  $M_n = \sup_{k \leq n} \|dr_k(R_k(x))\|$  de sorte que pour  $k \leq n$ ,  $\|N_k\| \leq 1 + M_n$ , on a

$$\|dR_n(x)\| \leq dn^d(1 + M_n)^d,$$

et donc

$$\frac{1}{n} \ln \|dR_n(x)\| \leq \frac{1}{n} \ln(dn^d) + \frac{d}{n} \ln(1 + M_n)$$

Le premier terme du membre de droite tend trivialement vers 0, et le second terme tend presque sûrement vers 0 pour  $\mu$  presque tout choix de  $x$ , par exemple en appliquant le théorème de Birkhoff à  $\ln(1 + \|dr(\cdot)\|)$ . En conséquence, le plus grand exposant de Lyapunov associé à  $\mu$  est négatif, et on montre de même que le plus petit exposant de Lyapunov est positif.

Nous avons démontré à ce stade i)  $\Rightarrow$  ii) et i)  $\Rightarrow$  iii), mais aussi que l'hypothèse i) entraîne les assertions plus précises énoncées à la fin du théorème.

Nous allons démontrer maintenant iii)  $\Rightarrow$  i), grâce au théorème de Avila-Viana. Appliquer le théorème directement à  $f$  ne donnera pas assez d'informations pour conclure, nous aurons besoin d'étendre  $f$  en un cocycle mettant en évidence l'action de  $df$  sur les droites de  $\mathbb{R}^d$ . A ces fins nous utiliserons les notations suivantes :

-Si  $h$  est un vecteur de  $\mathbb{R}^d$ , on pose  $\hat{h} = \frac{h}{|h|}$ .

-Si  $u$  est un isomorphisme de  $\mathbb{R}^d$ , on définit  $\hat{u} : S^{d-1} \rightarrow S^{d-1}$  par  $\hat{u}(h) = \widehat{u(h)}$ .

-Si  $f$  est un difféomorphisme de  $\mathbb{T}^d$ , on définit  $\hat{f} : \mathbb{T}^d \times S^{d-1} \rightarrow \mathbb{T}^d \times S^{d-1}$  par  $\hat{f}(x, h) = (f(x), \widehat{df(x)h})$ .

En plus du théorème d'Avila-Viana, la démonstration reposera sur le lemme suivant :

**Lemme 5.** Soit  $f$  dans  $\text{Diff}_0(\mathbb{T}^d)$  un difféomorphisme  $C^1$  préservant la mesure de Lebesgue et soit  $\nu$  une probabilité sur la sphere  $S^{d-1}$  invariante par  $\widehat{df(x)}$  pour tout  $x$  de  $\mathbb{T}^d$ , alors

$$\text{supp}(\nu) \subset \bigcap_{x \in \mathbb{T}^d} \ker(df(x) - Id).$$

Admettons ce lemme pour le moment. Donnons nous un difféomorphisme aléatoire  $f$  et une mesure stationnaire  $\mu$  comme dans l'énoncé du théorème, et supposons que iii) est vérifié, c'est à dire  $\gamma_d$  positif.

Par le théorème d'Avila-Viana,  $\mu$  est presque sûrement  $f$ -invariante. En particulier,  $\mu$  est presque sûrement  $R$ -invariante, donc  $\mu$  est la mesure de

#### 4.3. UNE CARACTÉRISATION DYNAMIQUE DES COCYCLES D'ANZAI GÉNÉRALISÉS 49

Lebesgue sur  $\mathbb{T}^d$ . Ainsi  $0 = \int_{\mathbb{T}^d} \ln \det df(x) dx = \gamma_1 + \dots + \gamma_d$ , et puisque le plus petit exposant  $\gamma_d$  est positif, on a en fait  $\gamma_1 = \dots = \gamma_d = 0$ .

Si  $\hat{\mu}$  est une mesure stationnaire de  $\hat{f}$  sur  $\mathbb{T}^d \times S^{d-1}$ , alors les exposants de Lyapunov associés sont encore tous nuls, donc  $\hat{f}_* \hat{\mu} = \hat{\mu}$  presque sûrement. En particulier,  $\hat{\mu}$  est presque sûrement  $\hat{f}$ -invariante, et on en déduit que  $\hat{\mu}$  est une mesure produit : en effet, notant,  $\nu$  la projection de  $\hat{\mu}$  sur  $S^{d-1}$ , si  $\varphi$  est une fonction continue sur  $\mathbb{T}^d \times S^{d-1}$ ,

$$\begin{aligned} \int_{\mathbb{T}^d \times S^{d-1}} \varphi(x, h) d\hat{\mu}(x, h) &= \lim_{n \rightarrow +\infty} \mathbb{E} \int_{\mathbb{T}^d \times S^{d-1}} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(R_k(x), h) d\hat{\mu}(x, h) \\ &= \int_{\mathbb{T}^d \times S^{d-1}} \varphi(x, h) dx d\nu(h). \end{aligned}$$

Presque sûrement, la mesure produit  $dx d\nu(h)$  est invariante par le cocycle  $\hat{f} : (x, h) \rightarrow (f(x), \widehat{df(x)h})$ , donc on déduit que  $\nu$  est  $\widehat{df(x)}$ -invariante pour tout  $x$  dans  $\mathbb{T}^d$  : en effet, si  $u$  est une fonction continue sur  $\mathbb{T}^d$ , et  $v$  une fonction continue sur  $S^{d-1}$ , l'intégrale  $\int_{\mathbb{T}^d \times S^{d-1}} u(x)v(h) dx d\nu(h)$  est égale d'une part à  $\int_{\mathbb{T}^d \times S^{d-1}} u(f(x))v(\widehat{df(x)h}) dx d\nu(h)$  car  $\hat{f}$  est invariant par  $\hat{\mu} = dx d\nu(h)$ , et d'autre part à  $\int_{\mathbb{T}^d \times S^{d-1}} u(f(x))v(h) dx d\nu(h)$  car  $f$  est Lebesgue-invariant. Ainsi, l'application continue  $x \mapsto \int_{S^{d-1}} (v(\widehat{df(x)h}) - v(h)) d\nu(h)$  est orthogonal à  $u \circ f$  pour toute application continue  $u$ , et est donc identiquement nulle par inversibilité de  $f$ .

Définissons maintenant le sous espace  $V$  des vecteurs  $h$  dans  $\mathbb{R}^d$  tel que presque sûrement, pour tout  $x$  dans  $\mathbb{R}^d$ ,  $df(x)h = h$ . Par le lemme 5,  $V$  est non trivial. Nous allons montrer qu'il existe dans  $V$  des vecteurs de  $\mathbb{Z}^d$  non triviaux. Notant  $\phi = f - Id = (\phi_1, \dots, \phi_d)$ , pour  $p = (p_1, \dots, p_d)$  dans  $\mathbb{Z}^d$  la matrice de

$$\widehat{d\phi}(p) = \int_{\mathbb{T}^d} e^{-2i\pi p \cdot x} d\phi(x) dx$$

dans la base canonique de  $\mathbb{R}^d$  s'écrit

$$\begin{pmatrix} p_1 \hat{\phi}_1(p) & \cdots & p_d \hat{\phi}_1(p) \\ \vdots & & \vdots \\ p_1 \hat{\phi}_d(p) & \cdots & p_d \hat{\phi}_d(p) \end{pmatrix} = \begin{pmatrix} \hat{\phi}_1(p) \\ \vdots \\ \hat{\phi}_d(p) \end{pmatrix} \cdot (p_1 \cdots p_d),$$

donc pour  $h$  dans  $\mathbb{R}^d$ ,  $\widehat{d\phi}(p)h = p \cdot h \hat{\phi}(p)$ . Si on note  $A$  l'ensemble des  $p$  tels que  $\hat{\phi}(p)$  n'est pas presque sûrement nul,  $h$  est dans  $V$  si et seulement si presque sûrement, pour tout  $p$  dans  $\mathbb{Z}^d$ ,  $\widehat{d\phi}(p)h = p \cdot h \hat{\phi}(p) = 0$ , donc si et seulement si  $p \cdot h = 0$  pour tout  $p$  dans  $A$ . L'espace  $V$  est donc régi par le système d'équations

$$h = (h_1, \dots, h_d) \in V \Leftrightarrow \forall p \in A, \sum_{k=1}^d h_k (p \cdot e_k) = 0,$$

où  $e_1, \dots, e_d$  est la base canonique de  $\mathbb{R}^d$ . Les coefficients  $p.e_k$  étant dans  $\mathbb{Q}^d$ , ce système admet des solutions non triviales dans  $\mathbb{Q}^d$  si et seulement pour tous  $p^{(1)}, \dots, p^{(d)}$  appartenant à  $A$ , le déterminant de  $(p^{(i)}.e_j)_{1 \leq i, j \leq d}$  est nul, donc si et seulement si il admet des solutions non triviales dans  $\mathbb{R}^d$ . Ceci prouve donc qu'il existe un vecteur  $h$  non trivial dans  $\mathbb{Q}^d \cap V$ , et donc aussi dans  $\mathbb{Z}^d \cap V$  en multipliant  $h$  par un dénominateur commun.

Quitte à diviser ce vecteur  $h$  par le pgcd de ses coordonnées, on peut supposer que celui ci vaut 1. Il existe alors  $P$  dans  $Sl_d(\mathbb{Z})$  tel que  $Pe_d = h$ . Le difféomorphisme aléatoire  $f_1 = P^{-1}fP$  vérifie alors  $df_1(x)e_d = e_d$  pour tout  $x$  de  $\mathbb{T}^d$ , presque sûrement. On en déduit que  $f_1$  est presque sûrement de la forme

$$\forall (x, t) \in \mathbb{T}^{d-1} \times \mathbb{T}, f_1(x, t) = (g(x), t + h(x)). \quad (4.2)$$

Cette première réduction faite, les exposants de Lyapunov de  $g$  associés à la mesure de Lebesgue sur  $\mathbb{T}^{d-1}$  sont tous nuls, ce qui permet de conclure par récurrence sur  $d$ .

Notons que pour démontrer la première réduction, à savoir l'existence de  $P$  dans  $Sl_d(\mathbb{Z})$  tel que  $f_1 = P^{-1}fP$  soit de la forme (4.2), nous n'avons pas utilisé l'assertion iii) mais seulement sa conséquence "il existe une probabilité  $\hat{\mu}$  presque sûrement invariante par  $\hat{f}$ ". Ceci nous permet de montrer l'implication ii)  $\Rightarrow$  i) d'une manière analogue. A la place du théorème d'Avila-Viana, on utilise la proposition suivante, qui sera démontrée au chapitre 6 (Proposition 1) :

**Proposition 6.** *Soit  $X$  un espace topologique compact, et  $G$  un sous groupe résoluble d'homéomorphismes de  $X$ . Alors il existe une mesure de probabilité  $\mu$  sur  $X$  invariante par  $G$ , c'est-à-dire telle que  $f_*\mu = \mu$  pour tout  $f$  dans  $G$ .*

Comme  $f \mapsto \hat{f}$  est un morphisme de groupe, la proposition précédente donne une mesure  $\hat{\mu}$  presque sûrement invariante par  $\hat{f}$ . En conséquence il existe  $P$  (déterministe) appartenant à  $Sl_d(\mathbb{Z})$  tel que  $f_1 = PfP^{-1}$  soit presque sûrement de la forme (4.2), et le difféomorphisme aléatoire  $g$  de  $\mathbb{T}^{d-1}$  appartient presque sûrement à un sous-groupe résoluble de  $\mathbb{T}^{d-1}$ , on peut là encore conclure par récurrence.

Reste maintenant à démontrer le lemme 5.

*Démonstration.* (Lemme 5) Posons  $E$  le sous-espace de  $\mathbb{R}^d$  engendré par  $\text{supp}(\mu)$ , stable par  $df(x)$  pour tout  $x$  de  $\mathbb{R}^d$ . Nous allons commencer par montrer que pour tout  $x$ , toutes les valeurs propres de  $df(x)|_E$  sont de module 1. Fixons  $x_0$  dans  $\mathbb{T}^d$  et notons  $r_1 < \dots < r_m$  les différents modules des valeurs propres de  $df(x_0)$  et pour  $1 \geq k \geq m$  posons  $E_k = \{h \in E \mid \lim_{n \rightarrow \pm\infty} \frac{1}{n} \ln |df(x_0)^n| = \ln r_k\} \cup \{0\}$  le sous espace de  $E$  associé à l'exposant de Lyapunov  $\ln r_k$  de  $df(x_0)|_E$ .

#### 4.3. UNE CARACTÉRISATION DYNAMIQUE DES COCYCLES D'ANZAI GÉNÉRALISÉS 51

Si  $h$  est un vecteur non nul de  $E$ , écrivons  $h = h_1 + \dots + h_p$  avec  $h_k$  dans  $E_k$  ( $k = 1, \dots, p$ ),  $h_p \neq 0$ . On sait alors que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \ln |df(x_0)^n h| = \ln r_p,$$

donc en particulier si  $k < p$ ,

$$\lim_{n \rightarrow +\infty} \frac{1}{|df(x_0)^n h|} df(x_0)^n h_k = 0,$$

et toute valeur d'adhérence de  $\widehat{df(x_0)}^n h$  appartient donc à  $E_p$ . En conséquence, si  $B$  est une boule fermée de  $S^{d-1}$  ne rencontrant pas  $E_1, \dots, E_m$ , pour tout  $h$  de  $B$ ,  $\widehat{df(x_0)}^n h$  n'appartient à  $B$  que pour un nombre fini d'entiers  $n$ , donc  $\nu(B) = 0$  par le théorème de récurrence de Poincaré, et donc  $\text{supp}(\nu) \subset E_1 \cup \dots \cup E_m$ . L'ensemble  $E_1 \cup \dots \cup E_m$  est donc stable par  $df(x)$  pour tout  $x$  dans  $\mathbb{T}^d$ .

Soit  $i \leq m$  et  $h$  dans  $E_i$ ,  $h \neq 0$ . Pour tout  $x$  dans  $\mathbb{T}^d$  il existe  $j \leq m$  tel que  $df(x)h \in E_j - \{0\}$ . Par continuité de  $x \mapsto df(x)$ ,  $j$  ne dépend pas de  $x$ , et puisque  $\int_{\mathbb{T}^d} df(x)h dx = h$ ,  $j$  est nécessairement égal à  $i$ . Ainsi pour tout  $i \leq m$  et tout  $x$  dans  $\mathbb{T}^d$ ,  $E_i$  est stable par  $df(x)$ .

Admettons temporairement le lemme suivant :

**Lemme 6.** *Si  $f = Id + \Phi$  est un difféomorphisme  $C^1$  de  $\mathbb{T}^d$  préservant le volume, et  $V$  un sous espace de  $\mathbb{R}^d$  stable par  $df(x)$  pour tout  $x$  dans  $\mathbb{T}^d$ , alors*

$$\forall x \in \mathbb{R}^d, \det df(x)|_V = 1$$

Ce lemme implique donc que pour tout  $i \leq m$ ,  $\det df(x_0)|_{E_i} = 1$ . Or toutes les valeurs propres de  $df(x_0)|_{E_i}$  sont de module  $r_i$ , donc  $r_i = 1$ , et toutes les valeurs propres de  $df(x_0)|_E$  sont donc de module 1.

On a montré que pour tout  $x$  de  $\mathbb{T}^d$ , les valeurs propres de  $df(x)|_E$  sont de module 1. En particulier, pour tout  $x$  de  $\mathbb{T}^d$ ,  $\text{Tr}(df(x)|_E) \leq \dim E$ , mais puisque

$$\int_{\mathbb{T}^d} \text{Tr}(df(x)|_E) dx = \int_{\mathbb{T}^d} \text{Tr}(Id_E) dx = \dim E,$$

on a en fait pour tout  $x$  de  $\mathbb{T}^d$   $\text{Tr}(df(x)|_E) = \dim E$ , et donc toutes les valeurs propres de  $df(x)|_E$  sont égales à 1.

Fixant  $x$  dans  $\mathbb{T}^d$ , l'endomorphisme  $N = df(x)|_E - Id_E$  est nilpotent. Utilisant la formule du binôme de Newton, si  $h$  est un vecteur de  $E$ , pour tout entier  $n$ ,  $df(x)^n h = \sum_{k=0}^p C_n^k N^k h$ , où  $p = \max(k \leq 0 | N^k h \neq 0)$  et  $C_n^k$  est le coefficient binomial "k parmi n". On a donc

$$\lim_{n \rightarrow +\infty} \frac{1}{C_n^p} df(x)^n h = N^p h \in \ker N - \{0\},$$

et la suite  $(\widehat{df(x)}^n h)_n$  converge donc vers un élément dans  $\ker N = \ker(df(x)|_E - Id_E) \subset \ker(df(x) - Id)$ , et donc  $\text{supp}(\mu) \subset \ker(df(x) - Id)$ . Cela achève la preuve du lemme 5, et donc de l'implication iii)  $\Rightarrow$  i).  $\square$

Démontrons maintenant le lemme 6 :

*Démonstration.* (Lemme 6)

Traisons dans un premier temps le cas  $V = \mathbb{R}^k \times \{0\}$ , où  $k = \dim V$ . Notant  $f = (f_1, \dots, f_d)$ , on a par hypothèse  $\partial_i f_j = 0$  si  $1 \leq i \leq k, k+1 \leq j \leq d$ , donc  $f$  est de la forme  $f(x, y) = (g(x, y), h(y))$  pour  $(x, y)$  dans  $\mathbb{T}^k \times \mathbb{T}^{d-k}$ . Sa différentielle s'écrit alors

$$df(x, y) = \begin{pmatrix} d_x g(x, y) & d_y g(x, y) \\ 0 & dh(y) \end{pmatrix},$$

donc  $\det df(x, y) = \det d_x g(x, y) \det dh(y)$ . Or  $d_x g(x, y) = df(x, y)|_V$ , et puisque  $f$  préserve le volume sur  $\mathbb{T}^d$ ,  $h$  préserve également le volume sur  $\mathbb{T}^{d-k}$ , donc  $\det df(x, y) = \det dh(y) = 1$ , et on obtient donc  $|\det df(x, y)|_V = 1$ .

Traisons maintenant le cas général, l'idée étant de se ramener au cas précédent en conjuguant  $f$  par une action linéaire  $P$ . Mais il faut pour cela choisir  $P$  dans  $Gl_d(\mathbb{Z})$ , il va donc s'agir de coder en un certain sens l'espace  $V$  par des équations linéaires à coefficients entiers, ce que nous allons faire grâce à une décomposition de  $\Phi$  en séries de Fourier.

Remarquons que si  $h$  appartient à  $V$ , alors, pour tout  $(p, x)$  dans  $\mathbb{Z}^d \times \mathbb{T}^d$ ,  $e^{-2i\pi p \cdot x} d\phi(x)h$  appartient à  $V + iV$ , donc

$$\widehat{d\phi}(p)h = \int_{\mathbb{T}^d} e^{-2i\pi p \cdot x} d\phi(x)h dx \in V + iV.$$

Par la relation  $\widehat{d\phi}(p)h = (p \cdot h)\hat{\phi}(p)$ , on en déduit que pour tout  $p$  dans  $\mathbb{Z}^d$ , soit  $\hat{\phi}(p) \in V + iV$ , soit  $p \cdot h = 0$ . Notons maintenant

$$A = \{p \in \mathbb{Z}^d \mid \hat{\phi}(p) \notin V + iV\}$$

et

$$W = \{h \in \mathbb{R}^d \mid h \perp A\}$$

Du raisonnement précédent on déduit que  $V \subset W$ . De plus, si  $h$  appartient à  $W$ , alors pour tout  $p$  dans  $\mathbb{Z}^d$ ,  $\widehat{d\phi}(p)h = (p \cdot h)\hat{\phi}(p)$  est soit nul (si  $p \in A$ ), soit un multiple d'un vecteur de  $V + iV$  (si  $p \notin A$ ), donc appartient à  $V + iV$  dans tous les cas. En conséquence, pour tout  $x$  dans  $\mathbb{T}^d$ ,  $d\Phi(x)h$  appartient à  $(V + iV) \cap \mathbb{R}^d = V$ . En particulier,  $W$  est stable par  $df(x)$  pour tout  $x$  dans  $\mathbb{T}^d$ .

Le théorème de structure des sous-groupes de  $\mathbb{Z}^d$  permet de trouver une  $\mathbb{Z}$ -base de  $\mathbb{Z}^d$ , disons  $u_1, \dots, u_d$ , un entier  $k$  et des entiers non nuls  $d_{k+1}, \dots, d_n$

#### 4.3. UNE CARACTÉRISATION DYNAMIQUE DES COCYCLES D'ANZAI GÉNÉRALISÉS 53

tels que  $d_{k+1}u_{k+1}, \dots, d_n u_n$  engendrent le même sous groupe additif que  $A$ . Définissant alors  $P$  dans  $Gl_d(\mathbb{Z})$  par  $Pe_i = u_i, i = 1, \dots, n$ , on a alors

$$P^*(W) = \text{vect}(e_1, \dots, e_k),$$

où  $P^*$  est la matrice transposée de  $P$ . En appliquant le cas particulier du lemme traité plus haut au difféomorphisme  $P^* f (P^*)^{-1}$ , on déduit que

$$\forall x \in W, \det(df(x)|_W) = 1.$$

Finalement, pour  $x$  dans  $\mathbb{T}^d$ , puisque  $d\Phi(x)W \subset V$  la matrice de  $df(x)|_W = Id + d\Phi(x)|_W$  dans une base de  $W$  associée à  $V$  est de la forme

$$df(x)|_W = \begin{pmatrix} df(x)|_V & \times \\ 0 & Id \end{pmatrix}$$

et donc

$$\det(df(x)|_V) = \det(df(x)|_W) = 1$$

□





## Chapter 5

# Lyapunov exponents of dynamical systems on the circle

### 5.1 Introduction

Arnold has proved in 1961 that a smooth orientation preserving diffeomorphism of  $\mathbb{T}$  close enough to a rotation in analytic topology, and whose rotation number is diophantine, is conjugated to a rotation, and that the conjugation is analytic and close to Identity (see [2]). In this paper, we intend to study the case where we have several such functions we want to simultaneously conjugate to rotations. More precisely, if  $f_1, \dots, f_m$  are diffeomorphisms of  $\mathbb{T}$  close enough to rotations whose rotation numbers  $\rho(f_i)$  ( $i = 1, \dots, m$ ) satisfy some diophantine condition, we want to find an assumption which insure that  $f_1, \dots, f_m$  are simultaneously conjugated to rotations. An obvious necessary condition is that these  $m$  functions commute, and under the hypothesis that at least one of the  $\rho(f_i)$  for  $i$  in  $\{1, \dots, m\}$ , is diophantine, this condition is sufficient by Arnold's theorem: indeed, up to a simultaneous conjugation, we can assume that this function  $f_i$  is an irrational rotation, and the commutation hypothesis then implies that all the  $f_j$ 's are rotations. Moser [14] proved in 1989 that under the commutation assumption, the diophantine hypothesis can be weakened in a simultaneous diophantine condition on  $\rho(f_1), \dots, \rho(f_m)$  which basically says that for each denominator  $q$ , there is at least one number  $\rho(f_i)$  which has a bad approximation by rationals of the form  $\frac{p}{q}$ . An other obvious necessary condition to simultaneous conjugation is that Lyapunov exponents of the Markov system  $x_{n+1} = f_{i_n}(x_n)$ , where  $(i_n)$  is an i.i.d sequence in  $\{1, \dots, m\}$ , are equal to 0, and we want to prove in this paper that this is also a sufficient condition under the same diophantine condition as Moser's. And instead of considering a finite number of diffeomorphisms, we will consider

the more general case of a Markov system generated by any probability on  $\text{Diff}_+^\infty(\mathbb{T})$ . This result is not a surprising fact, since there exists many other contexts where the nullity of Lyapunov exponent implies some rigidity. See for example Furstenberg's result [8] on products of i.i.d. matrices, and Ledrappier's generalisation [13] on product of matrices over a cocycle. See also the Dolgopyat and Krikorian's paper [6] which prove an analog of our result on  $S^d$ ,  $d \geq 2$ , and which initially inspired this paper.

Let's also cite some global theorems (i.e. without perturbation assumptions) on the conjugation of diffeomorphisms of  $\mathbb{T}$ . Herman [9] and Yoccoz [19] have proved that a  $C^\infty$  orientation preserving diffeomorphism of  $\mathbb{T}$  whose rotation number is diophantine is  $C^\infty$  conjugated to a rotation. B. Fayad and K. Khanin [7] have proved that Moser's result is still true without assuming the diffeomorphisms are close to rotations. Avila and Viana [4] have recently proved that nullity of Lyapunov exponent implies simultaneous conjugation. Together, these two last results give a global version of our results. But the perturbative case still remains interesting since it gives a good control in the construction of the conjugation, allowing for example to study the case where the Lyapunov exponent is small but not null, and moreover, the proof might be adapted in other contexts.

Here is the scheme of this paper. In a first part, we compute an explicit approximation of the Lyapunov exponent of the system, studying first its stationary measure. This approximation is interesting in itself, and can for example be applied to a random product of i.i.d.  $2 \times 2$  matrices, naturally acting on the projective space  $\mathcal{P}(\mathbb{R}^2)$  identified to  $\mathbb{T}$ . In a second part, we use this approximation to show that smallness of the Lyapunov exponent implies that the system is conjugated to another closer to rotations, and when Lyapunov exponent is null, we will see that we can iterate this process to get a simultaneous conjugation to rotations.

### 5.1.1 Notations and main results

We denote by  $\mathbb{T}$  the 1-dimensional torus  $\mathbb{R}/\mathbb{Z}$ . We will identify functions on  $\mathbb{T}$  with their lifting on  $\mathbb{R}$  by the application  $x \mapsto x + \mathbb{Z}$ . We denote by  $C^k(\mathbb{T})$  the space of the functions on  $\mathbb{R}$  which are  $C^k$  and 1-periodic, and by  $\text{Diff}_+^k(\mathbb{T})$  the space of increasing functions on  $\mathbb{R}$  on the form  $Id + \varphi$  with  $\varphi$  in  $C^k(\mathbb{T})$ . Setting for  $k$  in  $\mathbb{N}$

$$\|\varphi\|_k = \sup_{x \in \mathbb{R}, 0 \leq j \leq k} |\varphi^{(j)}(x)|,$$

we equip  $C^k(\mathbb{T})$  and  $Diff_+^k(\mathbb{T})$  with the metric

$$d_k(\varphi, \psi) = \|\varphi - \psi\|_k$$

if  $k < +\infty$ , and

$$d_\infty(\varphi, \psi) = \sum_{k=0}^{+\infty} \frac{1}{2^k} \inf(\|\varphi - \psi\|_k, 1)$$

if  $k = +\infty$ . Sometimes, for convenience, we will use the notations  $C^k$ ,  $\|\cdot\|_k$ ,  $d_k$  with  $k$  non integer to denote the corresponding quantity replacing  $k$  by its integer part.

If  $\varphi$  is a random  $C^k$  application on  $\mathbb{T}$  (i.e. a random variable  $\omega \mapsto \varphi^\omega$  valued in  $C^k(\mathbb{T})$ ), we denote

$$\|\varphi\|_k = \mathbb{E} \left[ \|\varphi\|_k^2 \right]^{\frac{1}{2}}$$

and if  $\varphi$  and  $\psi$  are random elements of  $C^k(\mathbb{T})$  or  $Diff_+^k(\mathbb{T})$ , we denote

$$\delta_k(\varphi, \psi) = \mathbb{E} \left[ d_k(\varphi, \psi)^2 \right]^{\frac{1}{2}}.$$

If  $f$  is a random diffeomorphism of  $\mathbb{T}$ , a probability measure  $\mu$  on  $\mathbb{T}$  is said to be *stationary* for  $f$  if it is invariant for the transition operator  $T\varphi = \mathbb{E}[\varphi \circ f]$ . The associated *Lyapunov exponent* is

$$\gamma(f, \mu) = \mathbb{E} \int_{\mathbb{T}} \ln f'(x) d\mu(x).$$

This Lyapunov exponent is invariant by conjugation in the sense that if  $G$  belongs to  $Diff_+(\mathbb{T})$ , then  $\gamma(f, \mu) = \gamma(GfG^{-1}, G_*\mu)$ .

A real valued random variable  $\alpha$  will be said *diophantine* if it satisfies for some  $\sigma > 1$  and  $A > 0$  the condition:

$$\forall q \in \mathbb{N}^*, \left\| \inf_{p \in \mathbb{Z}} \left( \alpha - \frac{p}{q} \right) \right\|_{L^2(\Omega)} \geq \frac{1}{Aq^\sigma}. \quad (5.1)$$

This is for exemple the case if the law of  $\alpha$  charges a diophantine number, or if it is not Lebesgue-singular.

**Theorem 1.** *Let  $\alpha$  a diophantine random variable and  $r_\alpha = Id + \alpha$ . There exists  $\varepsilon_0 = \varepsilon_0(\alpha)$  such that for every random  $C^\infty$  diffeomorphism  $f$  of  $\mathbb{T}$  satisfying  $\rho(f) = \alpha$ ,  $\delta_\infty(f, r_\alpha) < \varepsilon_0$  and  $\|f - r_\alpha\|_1 \leq 1$  a.s., and for every stationary measure  $\mu$  of  $f$ , the associated Lyapunov exponent  $\gamma = \gamma(f, \mu)$  is non positive and there exists a (non random) diffeomorphism  $H \in Diff_+^\infty(\mathbb{T})$  such that*

$$\delta_0(HfH^{-1}, r_\alpha) \leq \sqrt{-5\gamma}.$$

In particular, when  $\gamma = 0$ ,  $H$  is a simultaneous conjugation of almost every realisations of  $f$  to rotations. Moreover,  $H$  can be chosen close to identity in the sense that  $d_\infty(H, Id)$  tends to 0 when  $\varepsilon_0$  tends to 0.

**Corollary 1.** *Under the same assumptions as Theorem 1 ( $\alpha = \rho(f)$  diophantine,  $\|f - r_\alpha\|_1 \leq 1$  a.s. and  $\delta_\infty(f, r_\alpha)$  small enough), if  $\tilde{f}$  is an independent copy of  $f$ , then*

$$\delta_0(f \circ \tilde{f}, \tilde{f} \circ f) \leq 10 \sqrt{-\gamma}.$$

*Proof.* If  $\delta_\infty(f, r_\alpha)$  is small enough, then by Theorem 1 we can find  $H$  in  $Dif_+(\mathbb{T})$  with  $9/10 < H' < 11/10$  such that  $f_1 = HfH^{-1}$  satisfies  $\delta_0(f_1, r_\alpha) \leq \sqrt{-5\gamma}$ . Obviously, setting  $\tilde{f}_1 = H\tilde{f}H^{-1}$  and  $\tilde{\alpha} = \rho(\tilde{f})$ , we also have  $\delta_0(\tilde{f}_1, r_{\tilde{\alpha}}) \leq \sqrt{-5\gamma}$ . Consequently,  $\delta_0(\tilde{f}_1 \circ f_1, r_{\alpha+\tilde{\alpha}}) \leq \delta_0(f_1, r_\alpha) + \delta_0(\tilde{f}_1, r_{\tilde{\alpha}}) \leq 2\sqrt{-5\gamma}$  and  $\delta_0(\tilde{f}_1 \circ f_1, r_{\alpha+\tilde{\alpha}}) \leq 2\sqrt{-5\gamma}$ , hence  $\delta_0(f_1 \circ \tilde{f}_1, \tilde{f}_1 \circ f_1) \leq 9\sqrt{-\gamma}$ , which implies by the mean value inequality that  $\delta_0(f \circ \tilde{f}, \tilde{f} \circ f) \leq 10\sqrt{-\gamma}$ .  $\square$

**Remark :** In the same way, using Moser's ideas we could get an analogous result of theorem 1, which would quantitatively say that " $f$  and  $\tilde{f}$  almost commutes  $\Rightarrow f$  is almost conjugated to a rotation", and then deduce a reciprocal inequality of Corollary 1, of the type  $\sqrt{-\gamma} \leq C\delta_k(f \circ \tilde{f}, \tilde{f} \circ f)$ . Thus the square root of  $-\gamma$  seems to quantitatively represents the "default of commutativity" of realisations of  $f$ .

## 5.2 Proof of the main Theorem

In all this section, we fix a random increasing diffeomorphism  $f$  of  $\mathbb{T}$  and a random real variable  $\alpha$  satisfying (5.1) for some  $\sigma$  and  $A$ , and up to replacing  $A$  by  $\min(1, A)$ , we will assume  $A \geq 1$ . We set  $\zeta = f - r_\alpha$ . We also define on  $C^0(\mathbb{T})$  transfer operators  $T$  and  $T_0$  by  $T\varphi = \mathbb{E}[\varphi \circ f]$  and  $T_0\varphi = \mathbb{E}[\varphi \circ r_\alpha]$ . We will denote  $dx$  the Lebesgue measure on  $\mathbb{T}$ .

### 5.2.1 Cohomological equation

**Lemma 1.** *For any  $\psi \in C^k(\mathbb{T})$  with  $k \geq 2\sigma$ , the equation*

$$\varphi - T_0\varphi = \psi - \int_{\mathbb{T}} \psi dx$$

*has an unique solution  $\varphi$  such that  $\int_{\mathbb{T}} \varphi dx = 0$ . Moreover, this solution satisfies the inequality  $\|\varphi\|_{k-2\sigma} \leq \frac{\|\psi\|_k}{A^2}$*

*Proof.* The equality  $\varphi - T_0\varphi = \psi - \hat{\psi}(0)$  is equivalent to the fact that for any  $p$  in  $\mathbb{Z}^*$ ,  $\hat{\varphi}(p)(1 - \mathbb{E}[e^{2i\pi p\alpha}]) = \hat{\psi}(p)$ , or still  $\hat{\varphi}(p) = \frac{\hat{\psi}(p)}{1 - \mathbb{E}[e^{2i\pi p\alpha}]}$  ( $\mathbb{E}[e^{2i\pi p\alpha}]$  can not be equal

to 1 thanks to diophantine condition on  $\alpha$ ). Thus the uniqueness is clear and to justify the existence, it is sufficient to prove that we can define the function

$$\varphi(x) = \sum_p \frac{\hat{\psi}(q)}{1 - \mathbb{E}[e^{2i\pi q\alpha}]} e^{2i\pi qx},$$

and hence we need to bound the Fourier coefficients  $\frac{\hat{\psi}(q)}{1 - \mathbb{E}[e^{2i\pi q\alpha}]}$ . The modulus of the numerator  $\hat{\psi}(q)$  is less than  $\frac{\|\psi\|_k}{(2\pi|q|)^k}$ . Using the diophantine condition (5.1), we can estimate the modulus of the denominator  $1 - \mathbb{E}[e^{2i\pi q\alpha}]$  as follows :

$$\begin{aligned} |1 - \mathbb{E}[e^{2i\pi q\alpha}]| &\geq 1 - \mathbb{E}[\cos(2\pi q\alpha)] \\ &\geq \mathbb{E} \left[ \inf_{q \in \mathbb{Z}} \left( \frac{(2\pi(q\alpha - p))^2}{2} \right) \right] \\ &\geq \frac{2\pi^2 A^2}{|q|^{2\sigma-2}}. \end{aligned}$$

Thus finally,

$$\left| \frac{\hat{\psi}(q)}{1 - \mathbb{E}[e^{2i\pi q\alpha}]} \right| \leq \frac{2\|\psi\|_k}{(2\pi)^{k+2} A^2 |q|^{k-2\sigma+2}}.$$

So  $\varphi$  is well defined, is  $C^{k-2\sigma}$  and satisfies

$$\|\varphi\|_{k-2\sigma} \leq (2\pi)^{k-2\sigma} \sum_{q \in \mathbb{Z}} |q|^{k-2\sigma} |\hat{\varphi}(q)| \leq \frac{2}{(2\pi)^{2\sigma+2} A^2} \left( \sum_{q \in \mathbb{Z}^*} \frac{1}{|q|^2} \right) \|\psi\|_k \leq \frac{\|\psi\|_k}{A^2}$$

□

## 5.2.2 Estimation of $T$ -invariant measure

We define  $U$  the operator which associate to a function  $\psi$  in  $C^k(\mathbb{T})$  the solution  $\varphi$  of the cohomological equation  $\varphi - T_0\varphi = \psi - \int_{\mathbb{T}} \psi dx$ . Thus :

$$U\psi(x) = \sum_{q \neq 0} \frac{\hat{\psi}(q)}{1 - \mathbb{E}[e^{2i\pi q\alpha}]} e^{2i\pi qx}$$

By Lemma 1,  $U$  is well defined on  $C^k(\mathbb{T})$  for  $k \geq 2\sigma$  and satisfies the bound  $\|U\psi\|_{k-2\sigma} \leq A^{-2} \|\psi\|_k$ . We also define  $U^*$  the adjoint of  $U$  in  $L^2(\mathbb{T})$  ( $U^*\psi(x) = \sum_{q \neq 0} \frac{\hat{\psi}(q)}{1 - \mathbb{E}[e^{-2i\pi q\alpha}]} e^{2i\pi qx}$ ). Operator  $U$  will allow us to obtain an approximation of stationary measures of  $f$ , and next of the associated Lyapunov exponents.

**Proposition 1.** *If  $\mu$  is a  $T$ -invariant probability measure, then:*

$$\int_{\mathbb{T}} \varphi d\mu = \int_{\mathbb{T}} \varphi dx + O(\varepsilon \|\varphi\|_{2\sigma+1}) = \int_{\mathbb{T}} \varphi dx + \int_{\mathbb{T}} (U^* \bar{\zeta}) \varphi' dx + O(\varepsilon^2 \|\varphi\|_{4\sigma+2})$$

where  $\bar{\zeta} = \mathbb{E}[\zeta \circ r_{-\alpha}]$  and  $\varepsilon = \|\zeta\|_{4\sigma+2}$ .

(Here and in the sequel,  $O(M)$  is a quantity bounded by  $CM$  where  $C$  is a constant depending only on  $\alpha$ )

*Proof.* First, we will establish an expansion formula of  $\mu$  at order 0. We start from a Taylor formula at order 0 :

$$\varphi \circ f = \varphi \circ r_\alpha + O(\|\zeta\|_0 \|\varphi\|_1).$$

Taking the expectation, this becomes

$$T\varphi = T_0\varphi + O(\varepsilon \|\varphi\|_1).$$

Then, thanks to the invariance of  $\mu$  :

$$\int_{\mathbb{T}} (\varphi - T_0\varphi) d\mu = O(\varepsilon \|\varphi\|_1).$$

For  $\psi$  in  $C^{1+2\sigma}(\mathbb{T})$ , we apply the previous formula to  $\varphi = U\psi$  and we get :

$$\int_{\mathbb{T}} \psi d\mu = \int_{\mathbb{T}} \psi dx + O(\varepsilon \|\psi\|_{2\sigma+1}). \quad (5.2)$$

Now, to obtain an expansion of  $\mu$  at order 1, we use this time a Taylor formula at order 1 :

$$T\varphi = T_0\varphi + \mathbb{E}[(\varphi' \circ r_\alpha)\zeta] + O(\varepsilon^2 \|\varphi\|_2).$$

By  $\mu$ -invariance and our first estimation of  $\mu$  :

$$\begin{aligned} \int_{\mathbb{T}} (\varphi - T_0\varphi) d\mu &= \int_{\mathbb{T}} \mathbb{E}[(\varphi' \circ r_\alpha)\zeta] d\mu + O(\varepsilon^2 \|\varphi\|_2) \\ &= \int_{\mathbb{T}} \mathbb{E}[(\varphi' \circ r_\alpha)\zeta] dx + O(\varepsilon^2 \|\varphi\|_2 + \varepsilon \|\mathbb{E}[(\varphi' \circ r_\alpha)\zeta]\|_{2\sigma+1}) \\ &= \int_{\mathbb{T}} \varphi' \bar{\zeta} dx + O(\varepsilon^2 \|\varphi\|_{2\sigma+2}) \end{aligned}$$

As previously, for  $\psi$  in  $C^{4\sigma+2}(\mathbb{T})$  we take  $\varphi = U\psi$  to get

$$\begin{aligned} \int_{\mathbb{T}} \psi d\mu &= \int_{\mathbb{T}} \psi dx + \int_{\mathbb{T}} (U\psi)' \bar{\zeta} dx + O(\varepsilon^2 \|U\psi\|_{2\sigma+2}) \\ &= \int_{\mathbb{T}} \psi dx + \int_{\mathbb{T}} \psi' (U^* \bar{\zeta}) dx + O(\varepsilon^2 \|\psi\|_{4\sigma+2}) \end{aligned}$$

□

### 5.2.3 Estimate of the Lyapunov exponent

Thanks to Proposition 1 we can deduce estimates of the Lyapunov exponent of  $f$ :

**Proposition 2.**

$$\gamma(f, \mu) = -\frac{1}{2} \mathbb{E} \int_{\mathbb{T}} (\zeta' - (U^* \bar{\zeta})' \circ r_\alpha + (U^* \bar{\zeta})')^2 dx + O(\varepsilon^3)$$

where  $\varepsilon^3 = \mathbb{E}[|\zeta|^3]$ .

*Proof.* Let  $g = U^* \bar{\zeta}$ ,  $G = Id - g$ ,  $f_1 = GfG^{-1}$ ,  $\zeta_1 = f_1 - r_\alpha$  and  $\mu_1 = G_* \mu$  ( $G$  is invertible if  $\|g\|_1 \leq \frac{1}{2}$ , which is the case if  $\varepsilon$  is small enough). If  $\varphi$  is in  $C^{4\sigma+2}(\mathbb{T})$ , then thanks to Proposition 1:

$$\begin{aligned} \int_{\mathbb{T}} \varphi d\mu_1 &= \int_{\mathbb{T}} \varphi \circ G d\mu \\ &= \int_{\mathbb{T}} \varphi d\mu - \int_{\mathbb{T}} \varphi' g d\mu + O(\varepsilon^2 \|\varphi\|_2) \\ &= \left( \int_{\mathbb{T}} \varphi dx + \int_{\mathbb{T}} \varphi' g dx \right) - \int_{\mathbb{T}} \varphi' g dx + O(\varepsilon^2 \|\varphi\|_{4\sigma+2} + \varepsilon \|g\|_{2\sigma+1} \|\varphi\|_{2\sigma+1}) \\ &= \int_{\mathbb{T}} \varphi dx + O(\varepsilon^2 \|\varphi\|_{4\sigma+2}). \end{aligned}$$

Now, using invariance of Lyapunov exponent by conjugation and the estimates  $\|\zeta_1\|_1 = O(\varepsilon)$  and  $\zeta_1 = \zeta - g \circ r_\alpha + g + O(\varepsilon^2)$ , we can compute  $\gamma(f, \mu)$ :

$$\begin{aligned} \gamma(f, \mu) &= \mathbb{E} \int_{\mathbb{T}} \ln(1 + \zeta_1') d\mu_1 \\ &= \mathbb{E} \int_{\mathbb{T}} (\zeta_1' - \zeta_1'^2/2) d\mu_1 + O(\varepsilon^3) \\ &= -\frac{1}{2} \mathbb{E} \int_{\mathbb{T}} \zeta_1'^2 dx + O(\varepsilon^3) \\ &= -\frac{1}{2} \int_{\mathbb{T}} \mathbb{E}[(\zeta' - g' \circ r_\alpha + g')^2] dx + O(\varepsilon^3). \end{aligned}$$

□

**Remark :** We could avoid the conjugation by  $G$  to estimate  $\gamma$  and directly expand  $\mathbb{E} \int \ln f'(x) d\mu(x)$  using Proposition 1, but the method we have used has the advantage to make appear a main term clearly non-positive in the expansion of  $\gamma$ . Moreover, in the context of Theorem 1 this conjugation  $G$  will correspond to the first step in order to conjugate  $f$  to a diffeomorphism closer to rotations.

Here is an application of Proposition 2, independent of the proof of the main Theorem, which allows to estimate the characteristic Lyapunov exponent of a product of i.i.d.  $2 \times 2$  matrices close to rotation matrices.

**Corollary 2.** Let  $(M_n)$  an i.i.d sequence of random matrix of  $SL_2(\mathbb{R})$  on the form  $M_n = R_{\alpha_n} + N_n$  where  $R_{\alpha_n} = \begin{pmatrix} \cos \pi \alpha_n & -\sin \pi \alpha_n \\ \sin \pi \alpha_n & \cos \pi \alpha_n \end{pmatrix}$  and  $N_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix}$ , where  $2\alpha_n$  is not almost surely an integer. Then the Lyapunov exponent of the sequence  $(M_n)$ , given by  $\Gamma = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{\ln \|M_{n-1} \cdots M_0\|}{n} \right]$  satisfies, setting  $\varepsilon^3 = \mathbb{E}[\|N_0\|^3]$ ,  $z = e^{i\pi\alpha_0}$  and  $s = a_0i - b_0 + c_0 - d_0i$ :

$$\Gamma = \frac{1}{8|1 - \mathbb{E}[\bar{z}^2]|^2} \mathbb{E} \left[ |sz(1 - \mathbb{E}[\bar{z}^2]) + \mathbb{E}[s\bar{z}](1 - z)|^2 \right] + O(\varepsilon^3).$$

In the particular case where  $\alpha_0$  is not random, this formula becomes

$$\Gamma = \frac{1}{8} \mathbb{E} \left[ |s - \mathbb{E}[s]|^2 \right] + O(\varepsilon^3) = \frac{\text{Var}(s)}{8} + O(\varepsilon^3).$$

(  $O(\varepsilon^3)$  represents here a quantity bounded by  $C\varepsilon^3$  where  $C$  depends only on  $\alpha_0$  )

*Proof.* We identify  $\mathbb{C}$  and  $\mathbb{R}^2$ . Let  $f = r_\alpha + \zeta \in \text{Diff}_+(\mathbb{T})$  defined by

$$e^{if(x)} = \frac{M_0(e^{i\pi x})}{|M_0(e^{i\pi x})|}$$

$f$  can be seen as a perturbation of  $r_{\alpha_0}$ , and thanks to the formula of area conservation  $|\sin(\pi(\theta - \theta'))| = |M_0 e^{i\pi\theta}| |M_0 e^{i\pi\theta'}| |\sin(\pi(f(\theta) - f(\theta')))|$ , leading to  $1 = |M_0 e^{i\pi\theta}|^2 |f'(\theta)|$ , we deduce that the Lyapunov exponent of the system generated by  $f$  is  $\gamma = -2\Gamma$ , and we can estimate it by Proposition 2. We first need to estimate  $\zeta$ : writing that

$$i\pi\zeta(x) = \ln \left( \frac{1 + N_0(e^{i\pi x})e^{-i\pi(x+\alpha_0)}}{|M_0(e^{-i\pi x})|} \right),$$

and using that  $|M_0(e^{i\pi x})|^{-1} = \left(1 + 2\text{Re}(N_0(e^{i\pi x})e^{-i\pi(x+\alpha_0)}) + |N_0(e^{i\pi x})|^2\right)^{-1/2}$  and usual Taylor expansions, we get that  $\zeta = \zeta_0 + \zeta_1 + \zeta_2$  where :

- $\zeta_0(x) = \text{Im}(N_0(e^{i\pi x})e^{-i\pi(x+\alpha_0)})$  is a trigonometrical polynomial of degree 1 satisfying  $\|\zeta_0\|_0 = O(\|N\|)$
- $\zeta_1$  is a trigonometrical polynomial of degree 2 satisfying  $\|\zeta_1\|_0 = O(\|N\|^2)$
- $\|\zeta_2\|_0 = O(\|N\|^3)$



Now, if we assume that  $\alpha_0$  satisfies the diophantine condition (5.1), then by Proposition 2,

$$\Gamma = \frac{1}{2} \int_{\mathbb{T}} \mathbb{E}[(\zeta'_0 - (U^* \bar{\zeta}_0)' \circ r_\alpha + (U^* \bar{\zeta}_0)')^2] dx + O(\varepsilon^3).$$

A simple computation gives us

$$\begin{aligned} \zeta_0(x) &= -(a_0 \cos(\pi x) + b_0 \sin(\pi x)) \sin(\pi(x + \alpha_0)) \\ &\quad + (c_0 \cos(\pi x) + d_0 \sin(\pi x)) \cos(\pi(x + \alpha_0)) \\ &= \operatorname{Re} \left( \frac{s}{4} e^{i\pi(2x + \alpha_0)} \right) + \text{constant} \\ U^* \bar{\zeta}_0(x) &= \operatorname{Re} \left( \frac{\mathbb{E}[s\bar{z}]}{4(1 - \mathbb{E}[\bar{z}^2])} e^{2i\pi x} \right), \end{aligned}$$

hence using our estimation of  $\Gamma$ , the end of the proof is straightforward.

Now, in order to treat the case where  $\alpha_0$  does not satisfy (5.1), we must copy the proof of Proposition 2, remarking that we only need in this case to estimate  $\mu$  on trigonometrical polynomials of small degrees, and so we do not need a uniform estimate on small divisors appearing in the definition of  $U\varphi$ . More precisely, if  $T_0$  is transition operator associated to  $r_{\alpha_0}$ , the cohomological equation  $\varphi - T_0\varphi = \psi - \hat{\psi}(0)$  can be solve in the  $\mathbb{R}_2[e^{2i\pi x}]$  space of trigonometrical polynomials of degree 2, allowing to define as before the operator  $U$  on this space, and  $U$  is necessarily continuous since this space has a finite dimension. Next, we successively deduce that:

- For every  $\psi$  in  $\mathbb{R}_2[e^{2i\pi x}]$ , denoting  $\varphi = U\psi$  we have

$$\int_{\mathbb{T}} \psi d\mu - \int_{\mathbb{T}} \psi dx = \int_{\mathbb{T}} (\varphi - T_0\varphi) d\mu + O(\varepsilon \|\varphi\|_0) = O(\varepsilon \|\psi\|_0)$$

- For every  $\psi$  in  $\mathbb{R}_1[e^{2i\pi x}]$ ,

$$\begin{aligned} \int_{\mathbb{T}} \psi d\mu - \int_{\mathbb{T}} \psi(x) dx &= \int_{\mathbb{T}} (\varphi - T_0\varphi) d\mu + \int_{\mathbb{T}} \mathbb{E}[\varphi'(x + \alpha) \zeta_0(x)] d\mu(x) + O(\varepsilon^2 \|\varphi\|_0) \\ &= \int_{\mathbb{T}} \psi'(x) U^* \bar{\zeta}_0(x) dx + O(\varepsilon^2 \|\psi\|_0) \end{aligned}$$

- Denoting  $\mu_1 = (Id - U^* \bar{\zeta}_0)_* \mu$ , we have for  $\psi$  in  $\mathbb{R}_2[e^{2i\pi x}]$

$$\int_{\mathbb{T}} \psi d\mu_1 = \int_{\mathbb{T}} \psi d\mu + O(\varepsilon \|\psi\|_0) = \int_{\mathbb{T}} \psi dx + O(\varepsilon \|\psi\|_0)$$

and for  $\psi$  in  $\mathbb{R}_1[e^{2i\pi x}]$ ,

$$\begin{aligned} \int_{\mathbb{T}} \psi d\mu_1 &= \int_{\mathbb{T}} \psi d\mu - \int_{\mathbb{T}} \psi' U^* \bar{\zeta}_0 d\mu + O(\varepsilon^2 \|\psi\|_0) \\ &= \int_{\mathbb{T}} \psi dx + O(\varepsilon^2 \|\psi\|_0) \end{aligned}$$

- Denoting  $\tilde{f} = (Id - U^* \tilde{\zeta}_0) \circ f \circ (Id - U^* \tilde{\zeta}_0)^{-1} = Id + \alpha_0 + \tilde{\zeta}$ , we can write  $\tilde{\zeta} = \tilde{\zeta}_0 + \tilde{\zeta}_1 + \tilde{\zeta}_2$  with

$$\begin{cases} \tilde{\zeta}_0 = \zeta - U^*(\zeta_0) \circ r_{\alpha_0} + U^* \zeta_0 \in \mathbb{R}_1[e^{2i\pi x}], \|\tilde{\zeta}_0\|_0 = O(\varepsilon) \\ \tilde{\zeta}_1 \in \mathbb{R}_2[e^{2i\pi x}], \|\tilde{\zeta}_1\|_0 = O(\varepsilon^2) \\ \|\tilde{\zeta}_2\|_0 = O(\varepsilon^3) \end{cases}$$

and next

$$\begin{aligned} \gamma &= \int_{\mathbb{T}} \ln f'_1 d\mu_1 \\ &= \int_{\mathbb{T}} \tilde{\zeta}'_0 d\mu_1 + \int_{\mathbb{T}} \tilde{\zeta}'_1 d\mu_1 - \frac{1}{2} \int_{\mathbb{T}} \tilde{\zeta}_0'^2 d\mu_1 + O(\varepsilon^3) \\ &= -\frac{1}{2} \int_{\mathbb{T}} \tilde{\zeta}_0'^2 dx + O(\varepsilon^3) \end{aligned}$$

Thus, we can conclude that the claimed estimate on  $\Gamma$  is valid without diophantine assumptions on  $\alpha_0$ .  $\square$

**Remarks :**

1) If  $(M_n)$  is not assumed to belong to  $Sl_2(\mathbb{R})$  but only to  $Gl_2(\mathbb{R})$ , we still can obtain an asymptotic formula for the Lyapunov exponent of  $(M_n)$  applying previous statement to  $M_n / \sqrt{\det(M_n)}$ .

2) If  $M_n = \begin{pmatrix} E - gv_n & -1 \\ 1 & 0 \end{pmatrix}$ , where  $E = 2 \cos(\theta) \in ]-2, 2[ - \{0\}$  and  $(v_n)$  is an i.i.d centered sequence of reals, then  $M_n$  is conjugated to  $R_\theta + gv_n \begin{pmatrix} 1 & \cot \theta \\ 0 & 0 \end{pmatrix}$  and previous result gives the Lyapunov exponent estimate of Figotin and Pastur [16] when  $g$  tends to 0 :

$$\Gamma = \frac{g^2 V}{8 \sin^2 \theta} + O(g^3) = \frac{g^2 V}{2(4 - E^2)} + O(g^3)$$

( $V = \text{Var}(v_0)$ )

### 5.3 Simultaneous conjugation

We summarize tools we need to make functional calculus on  $C^k$  spaces in the following proposition :

**Proposition 3.** *We can find constants  $B_k$  depending only on  $k$  such that for any  $k$ , the following assertions are satisfied :*

1) *There exists a family of linear operators  $(P_\lambda)_{\lambda > 0}$  such that for any  $\varphi$  in  $C^k(\mathbb{T})$  and for any  $j < k$  :*

$$\begin{cases} \|P_\lambda \varphi\|_k \leq B_k \lambda^{k-j} \|\varphi\|_j \\ \|\varphi - P_\lambda \varphi\|_j \leq \frac{B_k \|\varphi\|_k}{\lambda^{k-j}}. \end{cases} \quad (5.3)$$

2)(Kolmogorov inequality) For any integer  $j \leq k$  and for any  $\varphi$  in  $C^k(\mathbb{T})$ ,

$$\|\varphi\|_j \leq B_k \|\varphi\|_k^{j/k} \|\varphi\|_0^{1-j/k}. \quad (5.4)$$

3) For any  $\varphi, \psi$  in  $C^k(\mathbb{T})$ , and any integer  $j \leq k$ ,

$$\|\varphi\|_j \|\psi\|_{k-j} \leq B_k (\|\varphi\|_k \|\psi\|_0 + \|\varphi\|_0 \|\psi\|_k) \quad (5.5)$$

and

$$\|\varphi\psi\|_k \leq B_k (\|\varphi\|_k \|\psi\|_0 + \|\varphi\|_0 \|\psi\|_k). \quad (5.6)$$

4) For any  $\Phi, \Psi$  in  $\text{Diff}_+^k(\mathbb{T})$  such that  $\max(d_1(\Phi, r_\alpha), d_1(\Psi, r_\beta)) \leq 3$ , and for any real numbers  $\alpha$  and  $\beta$ ,

$$d_k(\Phi \circ \Psi, r_{\alpha+\beta}) \leq B_k (d_k(\Phi, r_\alpha) + d_k(\Psi, r_\beta)) \quad (5.7)$$

and when  $d_1(\Phi, r_\alpha) \leq 1/2$ ,

$$d_k(\Phi^{-1}, r_{-\alpha}) \leq B_k d_k(\Phi, r_\alpha) \quad (5.8)$$

*Proof.* Points 1), 2) and 3) are classical:

- $P_\lambda$  is constructed using a convolution by an appropriate kernel.
- (5.4) is a consequence of (5.3), using the decomposition  $\varphi = P_\lambda \varphi + (\varphi - P_\lambda \varphi)$  for an appropriate choice of  $\lambda$ .
- (5.5) is a direct consequence of (5.4) and the convexity inequality  $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b$ , and (5.6) is a direct consequence of Leibniz formula and (5.5).

Let's now prove inequalities (5.7) and (5.8): writing  $\Phi = r_\alpha + \varphi$  and  $\Psi = r_\beta + \psi$ , we have  $d_k(\Phi \circ \Psi, r_{\alpha+\beta}) = \|\psi + \varphi \circ \Psi\|_k \leq \|\psi\|_k + \|\varphi \circ \Psi\|_k$ . And

$$\|\varphi \circ \Psi\|_k \leq \|\varphi\|_0 + \|\Psi' \varphi' \circ \Psi\|_{k-1} \leq \|\varphi\|_0 + B_{k-1} (\|\Psi'\|_0 \|\varphi' \circ \Psi\|_{k-1} + \|\Psi'\|_{k-1} \|\varphi'\|_0)$$

thanks to (5.6). Now, using that  $\|\Psi'\|_0 \leq 4$ , an iteration of this inequality gives the existence of  $C = C(k)$  such that

$$\|\varphi \circ \Psi\|_k \leq C \left( \|\varphi\|_k + \sum_{j=0}^{k-1} \|\Psi'\|_j \|\varphi'\|_{k-1-j} \right)$$

and so we get inequality (5.7) using that

$$\|\Psi'\|_j \|\varphi'\|_{k-1-j} \leq B_{k-1} (\|\Psi'\|_0 \|\varphi'\|_{k-1} + \|\Psi'\|_{k-1} \|\varphi'\|_0) \leq 3B_{k-1} (\|\varphi\|_k + \|\psi\|_k).$$

For inequality (5.8), we set  $\Psi = \Phi^{-1} = r_{-\alpha} + \psi$ , and we make the induction assumption that for every  $j < k$ ,  $\|\psi\|_j \leq B_j \|\varphi\|_j$  for some constant  $B_j$ . Noticing that  $\psi = -\varphi \circ \Psi$ , we get

$$\|\psi\|_k = \|\varphi \circ \Psi\|_k \leq \|\varphi\|_0 + \|\Psi' \varphi' \circ \Psi\|_{k-1} \leq \|\varphi\|_0 + \sum_{j=0}^{k-1} C_{k-1}^j \|\varphi' \circ \Psi\|_j \|\Psi'\|_{k-1-j}.$$

Then, for  $j = 0$ ,  $\|\varphi' \circ \Psi\|_j \|\Psi'\|_{k-1-j}$  is bounded by above by  $\frac{1}{2} \|\psi\|_k + \|\varphi\|_1$ , and for  $j \neq 0$ ,  $\|\varphi' \circ \Psi\|_j \|\Psi'\|_{k-1-j}$  is bounded by above up to a multiplicative constant by  $\|\varphi\|_k$  thanks to the inequality (5.7), induction assumption and inequality (5.5). Thus, we get an inequality of the type  $\|\psi\|_k \leq \frac{1}{2} \|\psi\|_k + C \|\varphi\|_k$ , which implies that  $\|\psi\|_k \leq 2C \|\varphi\|_k$ .  $\square$

The following lemma is an easy consequence of Proposition 3 which will give us useful  $C^k$ -estimates of a conjugation :

**Lemma 2.** *Let  $U = r_\alpha + u$ ,  $V = Id + v$  and  $W = Id + w$  in  $Diff^+(\mathbb{T})$  with  $\|u\|_1 \leq 2$ ,  $\max(\|v\|_1, \|w\|_1) \leq 1/2$ . Then :*

$$\begin{aligned} 1) d_k(VUV^{-1}, r_\alpha) &\leq B_k(\|u\|_k + \|v\|_k) \\ 2) d_0(VUV^{-1}, WUW^{-1}) &\leq B_0 d_0(V, W) \end{aligned}$$

where  $B_k$  is a constant depending only on  $k$ .

*Proof.* 1) is a direct consequence of inequalities (5.7) and (5.8) of Proposition 3. To prove 2), we write that

$$\begin{aligned} d_0(VUV^{-1}, WUW^{-1}) &\leq d_0(VUV^{-1}, WUW^{-1}) + d_0(WUW^{-1}, WUW^{-1}) \\ &\leq d_0(V, W) + d_0(WU, WUW^{-1}V) \\ &\leq (1 + \|(WUW^{-1})'\|_0) d_0(V, W) \end{aligned}$$

and we bound  $\|(WUW^{-1})'\|_0$  by the chain rule.  $\square$

For the sequel we assume that  $f$  satisfies assumptions of Theorem 1, that is  $\rho(f) = \alpha$ ,  $d_1(f, r_\alpha) \leq 1$  a.s. and  $\delta_\infty(f, r_\alpha)$  small enough. We fix  $\gamma$  a Lyapunov exponent of  $f$  associated to some stationary measure.

### 5.3.1 First conjugation

**Lemma 3.** *If  $\|\zeta\|_{4\sigma+2}$  is small enough,  $f$  is conjugated by some  $G = Id - g$  to some  $\tilde{f} = GfG^{-1} = r_\alpha + \tilde{\zeta}$  such that for any  $K \geq 2\sigma$  :*

- $\|g\|_{K-2\sigma} \leq C \|\zeta\|_K$
- $\|\tilde{\zeta}\|_0 \leq \sqrt{-4\gamma + C_0 \|\zeta\|_{4\sigma+2}^3}$

where  $C$  depends only on  $\alpha$  and  $K$ , and  $C_0$  depends only on  $\alpha$ .

*Proof.* We proceed in a same way as in Proposition 2. We set  $g = U^*\tilde{\zeta}$  which satisfies the first claimed inequality by Lemma 1,  $G = Id - g$  and  $\tilde{f} = GfG^{-1} = r_\alpha + \tilde{\zeta}$ . Noticing that  $\|\tilde{\zeta}\|_1 \leq \|\zeta\|_1(1 + O(\varepsilon)) \leq \frac{3}{2}$  if  $\varepsilon$  is small enough and that  $\ln(1+t) \leq t - \frac{t^2}{4}$  for  $t$  in  $[0, \frac{3}{2}]$ , we have

$$\gamma \leq \mathbb{E} \int_{\mathbb{T}} \left( \tilde{\zeta}' - \frac{1}{4} \tilde{\zeta}'^2 \right) d\mu_1 = -\frac{1}{4} \int_{\mathbb{T}} \tilde{\zeta}'^2 dx + O(\varepsilon^3)$$

where  $\varepsilon = \|\zeta\|_{4\sigma+2}$ . Thus there exists  $C_0$  depending only on  $\alpha$  such that

$$\mathbb{E} \int_{\mathbb{T}} \tilde{\zeta}'^2 dx \leq -4\gamma + C_0\varepsilon^3.$$

Next, we notice that for a fixed event, for every  $a, b$ ,  $|\tilde{\zeta}(a) - \tilde{\zeta}(b)| \leq \int_{\mathbb{T}} |\tilde{\zeta}'| dx$ , and since  $\rho(\tilde{f}) = \rho(f) = \alpha$ , we have  $\tilde{\zeta}(b) = 0$  for some  $b$ , and so  $\|\tilde{\zeta}\|_0 \leq \int_{\mathbb{T}} |\tilde{\zeta}'| dx$ . Thus, by Cauchy-Schwarz,  $\|\tilde{\zeta}\|_0^2 \leq \int_{\mathbb{T}} \tilde{\zeta}'^2 dx$ , and taking the expectation and using previous inequality gives the result.  $\square$

**Lemma 4.** *Let  $k = 4\sigma + 2$ . If  $\gamma \geq -C_0\|\zeta\|_k^3$  and if  $\|\zeta\|_{4\sigma+2}$  is small enough,  $f$  is conjugated by some  $G_0 = Id - g_0$  to some  $f_1 = G_0fG_0^{-1} = r_\alpha + \zeta_1$  such that, for every  $K \geq k$ :*

- $\|g_0\|_{K-2\sigma} \leq C_1\|\zeta\|_K$
- $\|\zeta_1\|_K \leq C_1\|\zeta\|_K\|\zeta\|_k^{-a}$
- $\|\zeta_1\|_k \leq C_1\|\zeta_1\|_K^\delta\|\zeta\|_k^{\frac{3}{2}(1-\delta)}$

where  $a = \frac{2\sigma}{k-2\sigma}$ ,  $\delta = k/K$  and  $C_1$  is a constant which depends only of  $K$  and  $\alpha$ .

*Proof.* If  $g$  is the function given by the previous lemma, we set  $g_0 = P_\lambda g$  for some  $\lambda > 1$  we will choose later (where  $P_\lambda$  is defined in proposition 3). The first inequality simply comes from  $\|g_0\|_{K-2\sigma} \leq \|g\|_{K-2\sigma}$  and Lemma 3. Next, by Lemma 2 and 3,

$$\|\zeta_1\|_K \leq B_K(\|\zeta\|_K + \|g_0\|_K) \leq B_K(\|\zeta\|_K + B_K\lambda^{2\sigma}\|g\|_{K-2\sigma}) \leq C\lambda^{2\sigma}\|\zeta\|_K \quad (5.9)$$

with  $C = B_K(1 + B_K C_0)$ . On another hand, still by Lemma 3,

$$\|GFG^{-1} - r_\alpha\|_0 \leq C'\|\zeta\|_k^{3/2}$$

with  $C' = \sqrt{5C_0}$ , and by Lemma 2,

$$\|G_0fG_0 - GfG^{-1}\|_0 \leq B_0\|G_0 - G\|_0 = B_0\|g - P_\lambda g\|_0 \leq B_0B_{k-2\sigma} \frac{\|g\|_{k-2\sigma}}{\lambda^{k-2\sigma}} \leq \frac{C''\|\zeta\|_k}{\lambda^{k-2\sigma}}$$

with  $C'' = B_0C_0B_{k-2\sigma}$ . Combining the two last inequalities we get

$$\|\tilde{\zeta}\|_0 = \|G_0fG_0^{-1} - r_\alpha\|_0 \leq \max(C', C'') \left( \|\zeta\|_k^{3/2} + \frac{\|\zeta\|_k}{\lambda^{k-2\sigma}} \right),$$

and by Kolmogorov inequality (5.4),

$$\|\zeta_1\|_k \leq C''' \|\zeta_1\|_k^\delta \left( \|\zeta\|_k^{3/2} + \frac{\|\zeta\|_k}{\lambda^{k-2\sigma}} \right)^{1-\delta} \quad (5.10)$$

with  $C''' = B_K \max(C'', C''')$ . Taking  $\lambda^{k-2\sigma} = \|\zeta\|_k^{-1/2}$ , (5.9) and (5.10) give the result.  $\square$

### 5.3.2 KAM iteration

Now we assume that  $\gamma \geq 0$ , and we will iterate the constructed conjugation, and verify the convergence of the scheme to get a conjugation of  $f$  to  $r_\alpha$ , which will imply in particular that  $\gamma = 0$ . Thus, we fix  $k = 4\sigma - 2$ , we define  $f_0 = f$ ,  $\zeta_0 = \zeta$ , and once constructed  $f_n = r_\alpha + \zeta_n$ , if Lemma 4 applies we set  $G_n = Id - g_n$  the conjugation obtained and  $f_{n+1} = G_n f_n G_n^{-1} = r_\alpha + \zeta_{n+1}$ . We also fix a large integer  $K$ , and we set  $\varepsilon_n = \|\zeta_n\|_k$  and  $\gamma_n = \|\zeta_n\|_K$ .

**Lemma 5.** *There exists an integer  $K_0$  and  $b > 0$  depending only on  $\alpha$  such that if  $K \geq K_0$ , then for some constant  $C$  depending on  $\alpha$ ,  $K$  and  $\gamma_0$ , we have  $\varepsilon_{n+1} \leq C^{n^2} \varepsilon_n^{\frac{4}{3}}$  and  $\gamma_n \leq C^{n^2} \varepsilon_n^{-b}$  as long as  $f_n$  can be defined.*

*Proof.* Using notations of Lemma 4

$$\gamma_{n+1} \leq C_1 \gamma_n \varepsilon_n^{-a} \quad (5.11)$$

$$\varepsilon_{n+1} \leq C_1 \gamma_{n+1}^\delta \varepsilon_n^{\frac{3}{2}(1-\delta)} \quad (5.12)$$

We will assume  $K$  large enough so that  $\frac{3}{2}(1-\delta) - 4a\delta > \frac{4}{3}$  where  $\delta = \frac{k}{K}$ . Let  $P_n = \varepsilon_n \cdots \varepsilon_0$ . By (5.11), we have  $\gamma_{n+1} \leq C_1^n \gamma_0 P_n^{-a}$ . So (5.12) becomes

$$\varepsilon_{n+1} \leq C_2^n \varepsilon_n^{\frac{3}{2}(1-\delta)} P_n^{-a\delta}. \quad (5.13)$$

with  $C_2 = ((1 + \gamma_0)C_1)^\delta$ . If for some integer  $n$  and real  $M > 1$ , we have  $\varepsilon_n \leq M P_n^{\frac{1}{4}}$ , then the previous inequality becomes

$$\varepsilon_{n+1} \leq M C_2^n P_n^{\frac{3}{2}(1-\delta)-a\delta} \leq M C_1^n P_n^{\frac{1}{3}},$$

and multiplying left and right side by  $\varepsilon_{n+1}^{\frac{1}{3}}$  and then taking the power  $\frac{3}{4}$ , we obtain  $\varepsilon_{n+1} \leq M C_2^n P_{n+1}^{\frac{1}{4}}$ . Thus by induction, we have  $\varepsilon_n \leq C_2^{n^2} P_n^{\frac{1}{4}}$ , and (5.13) becomes

$$\varepsilon_{n+1} \leq C_2^{n^2} \varepsilon_n^{\frac{3}{2}(1-\delta)-4a\delta}$$

and since  $\frac{3}{2}(1-\delta) - 4a\delta > \frac{4}{3}$ , this prove the first inequality. And the second inequality is then a direct consequence of  $\gamma_{n+1} \leq C_1^n \gamma_0 P_n^{-a}$  and  $\varepsilon_n \leq C_2^{n^2} P_n^{\frac{1}{4}}$ .  $\square$

Now we can finish the proof of Theorem 1 :  
 First we apply the lemma with  $K = K_0$ . Thus, if  $\varepsilon_0 = \|\zeta\|_{K_0}$  is small enough, then  $\varepsilon_n = O(\frac{\varepsilon_0}{2^{(4/3)^n}})$ . Next, we apply the lemma with a larger  $K$ : if  $l = \theta K$  be an integer such that  $\theta < \frac{1}{2(5a+1)}$ , then

$$\|\zeta_n\|_l \leq B_l \varepsilon_n^{1-\theta} \gamma_n^\theta \leq C \varepsilon_n^{1-\theta(5a+1)} \leq C \varepsilon_n^{\frac{1}{2}}$$

and so  $\|\zeta_n\|_l$  quickly decrease to 0. Since  $K$  can be chosen arbitrary large, this holds for every  $l$ .

Now we set  $H_n = G_{n-1} \cdots G_0$ , so that  $f_n = H_n f H_n^{-1}$ . We have  $H_n = Id + h_n$  with

$$h_n = \sum_{j=0}^{n-1} g_j \circ H_{j-1}.$$

For every  $l$ , we have for some  $C = C(l)$

$$\begin{aligned} \|h_n\|_l &\leq \sum_{j=0}^{n-1} \|g_j \circ H_j\|_l \\ &\leq C \sum_{j=0}^{n-1} \|g_j\|_l (1 + \|h_j\|_l) \\ &\leq C \varepsilon_0 \sup_{j < n} (1 + \|h_j\|_l). \end{aligned}$$

Using an induction we deduce that if  $\|h_n\|_l \leq C \varepsilon_0$ , and next that

$$\sum_{j=0}^{n-1} \|g_j \circ H_j\|_l \leq C \varepsilon_0$$

for some  $C = C(l)$ . In consequence,  $h_n$  normally converges in  $C^l(\mathbb{T}^2)$  to some limit  $h$  satisfying  $\|h\|_l \leq C \varepsilon_0$ , and  $H = P + h$  is thus  $C^\infty$ , close to  $Id$ , invertible if  $\varepsilon_0$  is small enough, and satisfies  $HfH^{-1} = r_\alpha$  almost surely.

Now, if we assume that  $\gamma < 0$ , then we notice that in the previous construction, we can define  $f_{n+1}$  from  $f_n$  as long as  $\gamma \geq -C_0 \varepsilon_n^3$  since we can use Lemma 4. Thus we can conjugate  $f$  to  $f_n$  satisfying  $\gamma \geq -C_0 \varepsilon_n^3$ . Then Lemma 3 allows us to conjugate  $f_n$  to some  $\tilde{f} = r_\alpha + \tilde{\zeta}$  satisfying  $\|\tilde{\zeta}\|_0 \leq \sqrt{-4\gamma + C_0 \varepsilon_n^3} \leq \sqrt{-5\gamma}$ . This completes the proof of Theorem 1.





## Chapter 6

# On simultaneous linearization of diffeomorphisms of $\mathbb{T}^2$

D.Dolgopyat and R.Krikorian studied in [6] Markov system on  $S^d$  ( $d \geq 2$ ) of the form  $x_{n+1} = f_{i_n}(x_n)$ , where  $(i_n)$  is an i.i.d sequence on  $\{1, \dots, m\}$ , and  $f_1, \dots, f_m$  are smooth diffeomorphisms  $C^\infty$ -close to rotations  $R_1, \dots, R_m$  whose action on  $S^d$  is minimal, and proved that for each stationary measure of this system, at least one associated Lyapunov exponent is non-positive unless  $f_1, \dots, f_m$  are simultaneously conjugated to rotations, with a conjugation  $C^\infty$ -close to Identity.

Our purpose is to obtain an analogous result on  $\mathbb{T}^d$ : if  $f_1, \dots, f_m$  are diffeomorphisms of  $\mathbb{T}^d$  close to rotations, does nullity of Lyapunov exponents imply simultaneous conjugation? We proved in the previous chapter that the answer is yes when  $d = 1$ , under some diophantine condition on rotations numbers  $\rho(f_1), \dots, \rho(f_m)$ . If  $d \geq 2$ , it does not hold anymore. Indeed, if each function  $f_i$  is on the form

$$f_i(x_1, \dots, x_d) = (x_1 + \alpha_1, x_2 + \alpha_2(x_1), x_3 + \alpha_3(x_1, x_2), \dots, x_d + \alpha_d(x_1, \dots, x_{d-1})) \quad (6.1)$$

where  $\alpha_k$  is a smooth map on  $\mathbb{T}^{k-1}$ , then it is not difficult to show that all the Lyapunov exponents of the system are null, though there is not in general simultaneous conjugation to rotations since in general, the  $f_i$ 's do not commute. Thus, the natural extension of the result would be rather:

if  $f_1, \dots, f_m$  are diffeomorphisms of  $\mathbb{T}^d$   $C^\infty$ -close to diffeomorphisms of the form (6.1), does nullity of Lyapunov exponents implies simultaneous conjugation of  $f_1, \dots, f_m$  to diffeomorphisms of the form (6.1)? In this paper we will prove this assertion, under some diophantine condition, in the case  $d = 2$ . Moreover, we will also prove that simultaneous conjugation is equivalent to some algebraic relation between  $f_1, \dots, f_m$  (precisely: the group of diffeomorphisms generated

by  $f_1, \dots, f_m$  is solvable of order 2, that is, for every  $i, j, k, l$ ,  $f_i f_j f_i^{-1} f_j^{-1}$  and  $f_k f_l f_k^{-1} f_l^{-1}$  commute. This can be viewed as an analog of Moser's theorem [14] which states that in a perturbative context, conjugation to translation is equivalent to the fact that the diffeomorphisms commute).

## 6.1 Notations and results

Let  $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$  be the  $d$ -dimensional torus. We will identify functions on  $\mathbb{T}^d$  with their liftings on  $\mathbb{R}^d$  by the map  $x \mapsto x + \mathbb{Z}^d$ . Thus, in this context, we will denote:

- $C^k(\mathbb{T}^d, \mathbb{R}^n)$  the space of the  $C^k$  functions on  $\mathbb{R}^d$  into  $\mathbb{R}^n$ , 1-periodic in each variable, or only  $C^k(\mathbb{T}^d)$  if  $n = 1$ . If  $k < +\infty$ , we endow  $C^k(\mathbb{T}^d)$  with the norm  $\|\varphi\|_k = \sup_{|\alpha| \leq k, x \in \mathbb{R}^d} |\partial^\alpha \varphi(x)|$ , and we endow  $C^\infty(\mathbb{T}^2)$  with the metric

$$d_\infty(\varphi, \psi) = \sum_{k \in \mathbb{N}} \frac{\min(1, \|\varphi - \psi\|_k)}{2^k}$$

- $Dif f^k(\mathbb{T}^d)$  the space of the functions which induce a  $C^k$  diffeomorphism of  $\mathbb{T}^d$ , that is the bijective functions  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$  of the form  $f(x) = Ax + \phi(x)$  where  $A \in Sl_d(\mathbb{Z})$ ,  $\phi \in C^k(\mathbb{T}^d, \mathbb{R}^d)$ .  $A$  is the *linear part* of  $f$ , and  $\phi$  its *periodic part*.
- $Dif f_0^k(\mathbb{T}^d)$  the connected component of  $Id$  in  $Dif f^k(\mathbb{T}^d)$ , that is functions of  $Dif f^k(\mathbb{T}^d)$  whose linear part is  $Id$ .

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a finite probability space. If  $E$  is one of the function spaces described above, we will denote by  $\bar{E}$  the space of random variables  $f : \omega \mapsto f^\omega$  from  $\Omega$  into  $E$ . Such a random variable  $f$  will be called random function/diffeomorphism, and if  $x$  belongs to  $\mathbb{T}^d$ ,  $f(x)$  will denote the random variable  $\omega \mapsto f^\omega(x)$ . If  $\|\cdot\|_\alpha$  is some functional norm and  $\varphi$  a random map, we will denote

$$\|\|\varphi\|\|_\alpha = \mathbb{E} \left[ \|\varphi\|_\alpha^2 \right]^{\frac{1}{2}}.$$

A probability measure  $\mu$  is said *stationary* for  $f \in \widetilde{Dif f^0}(\mathbb{T}^d)$  if  $\mathbb{E}[f_* \mu] = \mu$ . A stationary measure is said *ergodic* if it is extremal in the convex set of stationary measures.

If  $f = Id + \Phi \in \widetilde{Dif f_0^0}(\mathbb{T}^d)$ , we will call *random rotation vector* of  $f$  a random variable  $X$  on  $\mathbb{R}^d$  of the form  $X = \int_{\mathbb{T}^d} \phi d\mu$  where  $\mu$  is a stationary measure of  $f$ .

If  $A, \sigma$  are positive reals, a random variable  $\Theta$  in  $\mathbb{R}^d$  will be said  $(A, \sigma)$ -diophantine if

$$\forall q \in \mathbb{Z}^d, \mathbb{E} \left[ \inf_{p \in \mathbb{Z}} |q \cdot \Theta - p| \right] \geq \frac{A}{|q|^\sigma}.$$

From now on we will essentially work in  $\mathbb{T}^2$ :

We will call Anzai cocycles the maps of  $\text{Dif}f^\infty(\mathbb{T}^2)$  of the form  $R(x, y) = (x + \alpha, y + \beta(x))$ . We will write  $\hat{\beta}$  as a shortcut for  $\hat{\beta}(0) = \int_{\mathbb{T}} \beta(x) dx$ .

If  $f$  belongs to  $\widetilde{\text{Dif}}f^1(\mathbb{T}^2)$  and  $\mu$  is a stationary ergodic measure of  $f$ , then letting  $(f_n)$  be a sequence of independant copies of  $f$  and  $F_n = f_{n-1} \circ \dots \circ f_0$ , we set

$$\begin{cases} \gamma_1(f, \mu) = \lim_{n \rightarrow +\infty} \frac{1}{n} \mathbb{E} \int_{\mathbb{T}^2} \ln \|dF_n(x)\| d\mu(x) \\ \gamma_2(f, \mu) = \lim_{n \rightarrow +\infty} \frac{1}{n} \mathbb{E} \int_{\mathbb{T}^2} \ln \|dF_n(x)^{-1}\|^{-1} d\mu(x) \end{cases}$$

the Lyapunov exponents of  $f$  associated to  $\mu$ .

**Theorem 1.** *Let  $\Omega$  a finite probability space and let  $A, \sigma, M_0 > 0$ . There exists  $\varepsilon_0 > 0$  such that for every random Anzai cocycle  $R$  satisfying  $d_\infty(R, Id) \leq M_0$  a.s., and for every random diffeomorphism  $f : \Omega \rightarrow \text{Dif}f_0^\infty(\mathbb{T}^2)$  satisfying  $d_\infty(f, R) \leq \varepsilon_0$  a.s. and admitting a  $(A, \sigma)$ -diophantine random rotation vector, the following assertions are equivalent:*

- (I)  $\gamma_2(f, \mu) \geq 0$  for some probability measure  $\mu$  stationary for  $f$ .
- (II) There exists a solvable subgroup  $\mathcal{G}$  of  $\text{Dif}f_0^\infty(\mathbb{T}^2)$  such that  $f \in \mathcal{G}$  almost surely.
- (III)  $f$  is smoothly conjugated to a random Anzai cocycle, by a conjugation  $C^\infty$ -close to a linear automorphism of  $\mathbb{T}^2$ .

**Remark:** Random rotation vectors are not in general invariant by conjugation. But it is when the stationary measure of  $f$  is almost surely invariant by  $f$ , which is the case when  $f$  is conjugated to a random Anzai cocycle. (Anzai cocycles are Lebesgue invariant).

(III) trivially implies (I) and (II) (and in fact implies more precisely that  $\gamma_i(f, \mu) = 0$  ( $i = 1, 2$ ) and that  $f$  almost surely belongs to some solvable subgroup of  $\text{Dif}f_0(\mathbb{T}^2)$  with index 2), thus we will be interest with the converse implications. The way we will use assumption (II) relies on the following proposition:

**Proposition 1.** *Let  $X$  a compact space, and  $\mathcal{G}$  a solvable group of homeomorphisms of  $X$ . Then there exists a  $\mathcal{G}$ -invariant probability measure  $\mu$  (that is  $g_*\mu = \mu$  for every  $g$  in  $\mathcal{G}$ ).*

*Proof.* Let  $\mathcal{G}$  a solvable group of homeomorphisms of  $X$ . Thus, there exists a chain of subgroup  $\mathcal{G}_0 = \{Id\}, \mathcal{G}_1, \dots, \mathcal{G}_{n-1}, \mathcal{G}_n = \mathcal{G}$  such that for  $0 \leq i \leq n-1$ ,  $\mathcal{G}_i \triangleleft \mathcal{G}_{i+1}$  and  $\mathcal{G}_{i+1}/\mathcal{G}_i$  is abelian. Let  $\Pi(X) = C^0(X)'$  the set of probability measures on  $X$  endowed with the weak topology, and  $\mathcal{M}_i$  the set of  $\mathcal{G}_i$ -invariant probability measures, which is convex and closed in  $\Pi(X)$ . We assume that  $\mathcal{M}_i \neq \emptyset$  for some  $i$ , and then we define for  $f \in \mathcal{G}_{i+1}$  the linear operator  $T_f$  on  $\mathcal{M}_i$  by  $T_f(\mu) = f_*\mu$ . Since  $\mathcal{G}_i \triangleleft \mathcal{G}_{i+1}$ , for every  $f$  in  $\mathcal{G}_{i+1}$ , every  $g$  in  $\mathcal{G}_i$  and every  $\mu \in \mathcal{M}_i$  we have  $(f^{-1}gf)_*\mu = \mu$  and so  $g_*(f_*\mu) = f_*\mu$ , thus  $T_f(\mu) = f_*\mu \in \mathcal{M}_i$ . Moreover, if  $f, g \in \mathcal{G}_{i+1}$ , then  $f^{-1}g^{-1}fg \in \mathcal{G}_i$ , so for every  $\mu$  in  $\mathcal{M}_i$ ,  $(f^{-1}g^{-1}fg)_*\mu = \mu$  and

so  $T_f \circ T_g(\mu) = T_g \circ T_f(\mu)$ . Thus,  $(T_f)_{f \in \mathcal{G}_{i+1}}$  is a commuting family of linear operators of  $\mathcal{M}_i$  into itself. By Kakutani-Markov fixed point theorem, there exists a common fixed point of these operators in  $\mathcal{M}_i$ , that is there exists  $\mu$  in  $\mathcal{M}_i$  such that  $f_*\mu = \mu$  for every  $f$  in  $\mathcal{G}_{i+1}$ , so  $\mathcal{M}_{i+1} \neq \emptyset$ . By immediate induction,  $\mathcal{M}_n \neq \emptyset$ . □

We recall here properties of  $C^k$ -calculus we will need in the sequel:

**Proposition 2.** See [6]

i) For every  $\phi, \psi$  in  $C^k(\mathbb{T}^2)$ ,

$$\|\phi\psi\|_k \leq C(k)(\|\phi\|_k\|\psi\|_0 + \|\phi\|_0\|\psi\|_k)$$

ii) For every  $f = P + \phi$  and  $g = Q + \psi$  in  $\text{Dif} f^k(\mathbb{T}^2)$  ( $P$  and  $Q$  are the respective linear parts of  $f$  and  $g$ ) such that  $\|\psi\|_k \leq M_0$ ,

$$\|f \circ g - P \circ Q\| \leq C(k, M_0)(\|P\| \|\psi\|_k + \|Q\| \|\phi\|_k + \|\phi\|_k \|\psi\|_k)$$

iii) For every  $f = P + \phi$  in  $\text{Dif} f^k(\mathbb{T}^2)$  such that  $\|\phi\|_1 \leq \frac{1}{2\|P\|}$ ,

$$\|f^{-1} - P^{-1}\|_k \leq C(k)\|f - P\|_k$$

### 6.1.1 Scheme of the proof

We describe here the different steps of the proof:

- **Step 1:** If  $R$  is a random Anzai map whose random rotation vector  $(\alpha, \hat{\beta})$  satisfies a diophantine condition, and if we denote  $T_0 : \varphi \mapsto \mathbb{E}[\varphi \circ R]$  the associated transition operator, for  $\psi$  in  $C^\infty(\mathbb{T}^2)$  we will prove that the cohomological equation

$$\varphi - T_0\varphi = \psi - \int_{\mathbb{T}^2} \psi(z) dz$$

admits a solution  $\varphi$  in  $C^\infty(\mathbb{T}^2)$  whose norms of derivatives are controlled by derivatives of  $\psi$ .

- **Step 2:** Using estimates of step 1, we prove that a random diffeomorphism of  $\mathbb{T}^2$  close to  $R$  has stationary measures close to Lebesgue measure, noticing that such a stationary measure will be small on functions of the form  $\varphi - T_0\varphi$ . Moreover, by a finer analysis we will obtain a more precise approximation of stationary measures by some density measure  $(1 + h(z))dz$ , where  $h$  is small.

Next, we conjugate the diffeomorphism by a conjugation close to Identity in order to vanish the term  $h(z)dz$ . Thus, the new diffeomorphism has stationary measures closer to Lebesgue measure.

- **Step 3:** Using the formula

$$\gamma_1(f, \mu) + \gamma_2(f, \mu) = \mathbb{E} \int_{\mathbb{T}^2} \ln \det d_z f d\mu(z)$$

( $d_z f$  is the differential of  $f$  at point  $z$ ), when  $\mu$  is close to Lebesgue measure, we verify that the condition  $\gamma_2(f, \mu) \geq 0$  implies by Jensen inequality some uniform estimates from which we deduce that Lebesgue measure is close to be invariant by realisations of  $f$ , and that in fact,  $\gamma_1(f, \mu)$  and  $\gamma_2(f, \mu)$  are close to 0.

- **Step 4:** We study the action of the differential of  $f$  on the space of vectorial lines of  $\mathbb{R}^2$  denoted  $\mathcal{P}(\mathbb{R}^2)$  (which can be identified to  $\mathbb{T}$ ): we define a natural extension of  $f$  on  $\mathbb{T}^2 \times \mathcal{P}(\mathbb{R}^2)$  by  $\hat{f}(z, h) = (f(z), d_z f(h))$ . Using step 2, we can assume, up to conjugating the system, that stationary measures of the cocycle have a projection on  $\mathbb{T}^2$  close to Lebesgue measure. Moreover, using analog arguments we can also assume that stationary measures of  $\hat{f}$  are close to product measures.

When  $f$  is close to a translation,  $d_z f$  is close to the identity, and by Taylor expansions the transition operator  $S$  of the cocycle has on  $C^\infty(\mathbb{T})$  the form, up to negligibles terms:

$$S\varphi(h) \approx \varphi(h) + (a(h)\varphi'(h) + b(h)\varphi''(h)),$$

with  $b \geq 0$ . Next we prove a dichotomy:

-If  $b$  has a bound by below not too small, this implies some ellipticity on the operator  $S$  and then stationary measures can be approximated by some density  $dz \otimes g(h)dh$ , where  $g$  is smooth and has a controlled size.

-If there exists  $h$  such that  $b(h)$  is very small, then this direction  $h$  is close to be invariant by  $d_z f$ , uniformly in  $z$ .

If the first case of this alternative occurs, then using that the Lyapunov exponent of the cocycle is close to 0 (consequence of the fact that Lyapunov exponents of  $f$  are close to 0), we prove that the cocycle is close to be linearizable, that is conjugated to one on the form  $(z, h) \mapsto (f(z), e^{i\omega(z)}h)$  up to negligible terms. And we prove next that such a linearization is impossible in this context, and so that the first case cannot occur. Thus, the second case of the alternative necessarily occurs and there exists a direction  $h$  close to be invariant by the cocycle. It is then not difficult to deduce that  $f$  is linearly conjugated to a diffeomorphism close to an Anzai cocycle.

When  $f$  is not close to a translation, using a conjugation of the cocycle by  $(z, h) \mapsto (z, D(h))$  where  $D$  is some diagonal matrix, we can obtain a

cocycle whose second term is close to the identity, and thus use previous arguments and prove the same result in this case, and more precisely in this case the diffeomorphism  $f$  is close to an Anzai cocycle without algebraic conjugation.

- **Step 5** At this point we have proved that a random diffeomorphism close to a diophantine Anzai cocycle whose Lyapunov exponents are positive is conjugated to a random diffeomorphism closer to an Anzai map, and in general, the algebraic part of the conjugation is the identity. Iterating this fact, by the use of classical KAM methods we prove that the random diffeomorphism is conjugated to an Anzai cocycle.

**Remark:** It is natural to wonder if the result holds in a torus of dimension  $d > 2$ . It seems that a large part of the proof can be adapted, and the main difficulty seems to be the step 4: the projective space  $\mathcal{P}(\mathbb{R}^d)$  has dimension  $d - 1$ , and the approximation

$$S\varphi \approx \varphi + (a\varphi' + b\varphi''),$$

becomes on the form

$$S\varphi \approx \varphi + \left( \sum_i a_i \partial_i \varphi + \sum_{i,j} b_{i,j} \partial_{ij} \varphi \right),$$

where the right operator is elliptic. In consequence the study of the equation  $S^* \mu = \mu$  becomes much harder, and the classical results of elliptic theory does not suffice to obtain needed estimates on  $\mu$ . It is not clear if this new difficulty is deep or only technical.

## 6.2 Cohomological equation

For all the sequel, we will assume  $\sigma$  and  $M_0$  are fixed, and we will not explicit the dependence of constants which will appear in  $\sigma$  and  $M_0$ . We will neither explicit their dependence on  $A$ , but we will need the easily checkable fact that this dependence is polynomial (there are bounded by  $CA^{-m}$  for some  $m > 0$  with  $C$  not depending on  $A$ ) We will use the notation  $O(Z)$  to represent a quantity whose absolute value is bounded by  $CZ$  for some constant  $C$ , and if  $E$  is some normed vectorial space,  $O_E(Z)$  will represent an element of  $E$  whose norm is bounded by  $CZ$ .

Let  $R : (x, y) \mapsto (x + \alpha, y + \beta(x))$  be a random Anzai cocycle,  $T_0$  the associated transition operator defined by  $T_0\varphi = \mathbb{E}[\varphi \circ R]$ ,  $\hat{\beta} = \int_{\mathbb{T}} \beta(x) dx$  and  $\Delta = \|R \circ \tilde{R} - \tilde{R} \circ R\|_{L^2(\mathbb{T}^2)}$  where  $\tilde{R}$  is an independant copy of  $R$ . We will assume that  $(\alpha, \hat{\beta})$  is  $(A, \sigma)$ -diophantine and that  $d_\infty(R, Id) \leq M_0$ .

If  $k_1$  and  $k_2$  are two integers and  $\varphi$  is a  $C^\infty$  function on  $\mathbb{T}^2$ , we will denote

$$\|\varphi\|_{k_1, k_2} = \sup_{i_1, i_2 \in \mathbb{N}, 0 \leq \frac{i_1}{k_1} + \frac{i_2}{k_2} \leq 1} \left\| \frac{\partial^{i_1+i_2} \varphi}{\partial x^{i_1} \partial y^{i_2}} \right\|_{L^2(\mathbb{T}^2)}$$

(with the convention  $\frac{1}{0} = +\infty$ ,  $\frac{0}{0} = 0$ ).

**Proposition 3.** *The operator  $I - T_0$  is a bijection from the space*

$$E_0 = \left\{ \varphi \in C^\infty(\mathbb{T}^2) \mid \int_{\mathbb{T}^2} \varphi(x, y) dx dy = 0 \right\}$$

onto itself. Moreover:

i) *For every integer  $k$  there exists an integer  $K$  such that for every  $\psi$  in  $E_0$ ,*

$$\|(I - T_0)^{-1} \psi\|_k \leq C \|\psi\|_K$$

for some  $C = C(k)$ .

ii) *If  $\Delta \neq 0$ , there exists an integer  $a$  such that for every integers  $k_1, k_2 \geq 1$  and every  $\psi$  in  $E_0$ ,*

$$\|(I - T_0)^{-1} \psi\|_{k_1, k_2} \leq \frac{C}{\Delta^{8k_1}} \|\psi\|_{k_1+a, k_2+a, k_1},$$

for some constant  $C = C(k_1, k_2)$ .

iii) *For every  $\delta > 0$  and every integer  $p$ , there exists an integer  $k_0$  such that for every  $\psi$  in  $E_0$ , there exists  $\varphi$  in  $E_0$  such that for every  $k_1$  and  $k_2$  larger than  $k_0$ ,*

$$\begin{cases} \|\varphi\|_{k_1, k_2} \leq C \|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \\ \|(I - T_0)\varphi - \psi\|_{k_1, k_2} \leq C \Delta^p \|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \end{cases}$$

for some constant  $C = C(\delta, p, k_1, k_2)$ .

*Proof.* We will use the decomposition of the space  $E_0$  on two subspaces  $E_0 = E_{00} \oplus F$  where

$$E_{00} = \left\{ \varphi \in C^\infty(\mathbb{T}^2) \mid \forall x \in \mathbb{T}, \int_{\mathbb{T}} \varphi(x, y) dy = 0 \right\}$$

and  $F$  is the subspace of functions of  $E_0$  which depend only on the first variable. Notice that the spaces  $E_{00}$  and  $F$  are stable by  $T_0$ , and by the derivations  $\partial_1$  and  $\partial_2$ . It is easy to check that  $I - T_0$  is a bijection of  $F$  onto itself satisfying estimates i), ii) and iii) (it satisfies in fact stronger estimates), since for  $\varphi$  in  $F$ ,  $T_0\varphi(x) = \mathbb{E}[\varphi(x + \alpha)]$ , so that the cohomological equation  $\varphi - T_0\varphi = \psi$  can be solved using Fourier series expansion (see Lemma 1 of Chapter 5). Thus, we will mainly study the restriction of  $I - T_0$  on  $E_{00}$ .

Up to replacing  $R$  by  $URU^{-1}$  with  $U(x, y) = (x, y + u(x))$ ,  $u$  being defined by the relation  $u(x) - \mathbb{E}[u(x + \alpha)] = \mathbb{E}[\beta(x) - \hat{\beta}]$  (see Lemma 1 of previous chapter), we will from now on assume that

$$\forall x \in \mathbb{T}, \mathbb{E}[\beta(x) - \hat{\beta}] = 0.$$

If  $\tilde{R}$  is an independant copy of  $R$ , then

$$(R \circ \tilde{R} - \tilde{R} \circ R)(x, y) = (0, (\beta(x + \tilde{\alpha}) + \tilde{\beta}(x)) - (\tilde{\beta}(x + \alpha) + \beta(x))).$$

Since  $\mathbb{E}[\tilde{\beta}] = \text{constant}$ ,

$$\mathbb{E}[R \circ \tilde{R} - \tilde{R} \circ R | \alpha, \beta] = (0, T_0\beta - \beta),$$

and in consequence

$$\| \|T_0\beta - \beta\| \|_{L^2(\mathbb{T})} \leq \Delta.$$

Thus, still by Lemma 1 of the previous chapter, we have for every integer  $k$ ,

$$\| \|\beta - \hat{\beta}\| \|_k \leq C \| \|T_0\beta - \beta\| \|_{k+2\sigma}$$

for some  $C = C(k)$ , and for  $K$  large enough, we have by convexity inequality

$$\|T_0\beta - \beta\|_k \leq C \sqrt{\| \|T_0\beta - \beta\| \|_{L^2(\mathbb{T}^2)} \| \|T_0\beta - \beta\| \|_K} \leq C \sqrt{2\|\beta\|_K} \sqrt{\| \|T_0\beta - \beta\| \|_K}$$

so that we finally get the estimate

$$\| \|\beta - \hat{\beta}\| \|_{L^2(\mathbb{T}^2)} \leq C \sqrt{\Delta}$$

for some  $C = C(k, M_0)$ .

When  $\Delta = 0$ , this implies that  $R$  is almost surely a translation on  $\mathbb{T}^2$ : in this case we do not need to use the previous decomposition of  $E_0$ , Proposition 3 can be proved by Fourier arguments:

**Lemma 1.** *If  $R$  is almost surely a translation, then  $(I - T_0)$  is a bijection from  $E_0$  onto itself, and for every  $\delta > 0$  there exists an integer  $k_0$  such that for every integers  $k_1$  and  $k_2$  larger than  $k_0$  and for every  $\psi$  in  $E_0$ ,*

$$\| (I - T_0)^{-1} \psi \|_{k_1, k_2} \leq C \| \psi \|_{(1+\delta)k_1, (1+\delta)k_2}$$

for some  $C = C(\delta, k_1, k_2)$ .

*Proof.* The equation  $\varphi - T_0\varphi = \psi$  is equivalent to

$$\forall p = (p_1, p_2) \in \mathbb{Z}^2, \hat{\varphi}(p)(1 - \mathbb{E}[e^{2i\pi(p_1\alpha + p_2\beta)}]) = \hat{\psi}(p)$$

Thus defining  $\varphi$  by

$$\varphi(x) = \sum_{p \in \mathbb{Z}^2 - \{0\}} \frac{\hat{\psi}(p)}{1 - \mathbb{E}[e^{2i\pi(p_1\alpha + p_2\beta)}]} e^{2i\pi(p_1x + p_2y)},$$



we have  $\varphi - T_0\varphi = \psi$  and using that

$$|1 - \mathbb{E}[e^{2i\pi(p_1\alpha + p_2\beta)}]| \geq \frac{A^2}{|p_1|^{2\sigma} + |p_2|^{2\sigma}}$$

we deduce that  $\psi$  is  $C^\infty$  and that there exists an integer  $a$  such that

$$\|\varphi\|_{L^2(\mathbb{T}^2)} \leq C(\|\partial_1^a \psi\|_{L^2(\mathbb{T}^2)} + \|\partial_2^a \psi\|_{L^2(\mathbb{T}^2)}).$$

If  $k_1$  and  $k_2$  are larger than  $\frac{a}{\delta}$  and if  $i_1$  and  $i_2$  are integers satisfying  $\frac{i_1}{k_1} + \frac{i_2}{k_2} \leq 1$ , then since  $(I - T_0)\partial_1^{i_1}\partial_2^{i_2}\varphi = \partial_1^{i_1}\partial_2^{i_2}\psi$  we get by the previous inequality

$$\|\partial_1^{i_1}\partial_2^{i_2}\varphi\|_{L^2(\mathbb{T}^2)} \leq C(\|\partial_1^{i_1+a}\partial_2^{i_2}\psi\|_{L^2(\mathbb{T}^2)} + \|\partial_1^{i_1}\partial_2^{i_2+a}\psi\|_{L^2(\mathbb{T}^2)}).$$

with

$$\frac{i_1 + a}{(1 + \delta)k_1} + \frac{i_2}{(1 + \delta)k_2} = \frac{1}{1 + \delta} \left[ \left( \frac{i_1}{k_1} + \frac{i_2}{k_2} \right) + \frac{a}{k_1} \right] \leq 1$$

so that  $\|\partial_1^{i_1+a}\partial_2^{i_2}\psi\|_{L^2(\mathbb{T}^2)} \leq \|\psi\|_{(1+\delta)k_1, (1+\delta)k_2}$ , and  $\|\partial_1^{i_1}\partial_2^{i_2+a}\psi\|_{L^2(\mathbb{T}^2)} \leq \|\psi\|_{(1+\delta)k_1, (1+\delta)k_2}$  in the same way. The result follows.  $\square$

In the sequel, we will assume that  $\Delta \neq 0$ . Now we will prove the point iii) of Proposition 3

**Lemma 2.** *For every integer  $p$  and every  $\delta > 0$ , there exists  $k_0$  such that for every  $k_1, k_2$  larger than  $k_0$  and every  $\psi$  in  $E_0$ , there exists  $\varphi$  in  $C^\infty(\mathbb{T}^2)$  such that*

$$\begin{cases} \|\varphi\|_{k_1, k_2} \leq C\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \\ \|(I - T_0)\varphi - \psi\|_{k_1, k_2} \leq C\Delta^p\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2}. \end{cases}$$

*Proof.* We set  $\tilde{T}_0\varphi(x, y) = \mathbb{E}[\varphi(x + \alpha, y + \beta)]$ . This operator  $\tilde{T}_0$  is close to  $T_0$  when  $\Delta$  is small, in the sense that for every  $\varphi$  in  $E_0$ ,

$$\|(T_0 - \tilde{T}_0)\varphi\|_{k_1, k_2} \leq C\|\beta - \hat{\beta}\|_{k_1}\|\varphi\|_{k_1, k_2+1} \leq C\sqrt{\Delta}\|\varphi\|_{k_1, k_2+1}$$

Moreover, we know by Lemma 1 that  $I - \tilde{T}_0$  is invertible on  $E_0$ . For  $\psi$  in  $E_0$ , we define functions  $\varphi_p$  and  $\psi_p$  by induction by  $\varphi_0 = 0$ ,  $\psi_p = (I - T_0)\varphi_p$  and  $\varphi_{p+1} = \varphi_p + (I - \tilde{T}_0)^{-1}(\psi - \psi_p)$ . Now let  $p$  in  $\mathbb{N}$ , let us assume that for every  $\delta > 0$  if  $k_1$  and  $k_2$  are large enough we have

$$\begin{cases} \|\varphi_p\|_{k_1, k_2} \leq C\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \\ \|\psi_p - \psi\|_{k_1, k_2} \leq C\Delta^{\frac{p}{2}}\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2}. \end{cases}$$

for some constant  $C = C(\delta, k_1, k_2)$ . Then

$$\|\varphi_{p+1} - \varphi_p\|_{k_1, k_2} = \|(I - \tilde{T}_0)^{-1}(\psi - \psi_p)\|_{k_1, k_2} \leq C\|\psi - \psi_p\|_{(1+\delta)k_1, (1+\delta)k_2} \leq C\Delta^{\frac{p}{2}}\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2},$$

hence

$$\begin{cases} \|\varphi_{p+1}\|_{k_1, k_2} \leq \|\varphi_p\|_{k_1, k_2} + C\Delta^{\frac{p}{2}}\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \leq C\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \\ \|\psi_{p+1} - \psi\|_{k_1, k_2} = \|(T_0 - \tilde{T}_0)(\varphi_{p+1} - \varphi_p)\|_{k_1, k_2} \leq C\|\varphi_{p+1} - \varphi_p\|_{k_1, k_2+1} \leq C\Delta^{\frac{p+1}{2}}\|\psi\|_{(1+\delta)k_1, (1+\delta)k_2} \end{cases}.$$

We deduce by induction that  $\varphi = \varphi_{2p}$  satisfies the claimed estimates.  $\square$

**Lemma 3.** *There exists an integer  $k_0$  such that for every  $\varphi$  in  $E_{00}$ ,*

$$\|\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^8} \|\partial_2^{k_0}(\varphi - T_0\varphi)\|_{L^2(\mathbb{T}^2)}$$

*Proof.* Let  $\varepsilon = \mathbb{E}[\|\partial_2^k(\varphi \circ R - \varphi)\|_{L^2(\mathbb{T}^2)}^2]^{1/2} = \mathbb{E}[\|\partial_2^k(\varphi - \varphi \circ R^{-1})\|_{L^2(\mathbb{T}^2)}^2]^{1/2}$  for some integer  $k$  we will chose later. We have

$$\begin{aligned} \varepsilon^2 &= 2 \int_{\mathbb{T}^2} \partial_2^k \varphi(x, y)^2 dx dy - 2 \int_{\mathbb{T}^2} \partial_2^k \varphi(x, y) T_0(\partial_2^k \varphi)(x, y) dx dy \\ &= \int_{\mathbb{T}^2} \varphi \partial_2^{2k}(\varphi - T_0\varphi) dx dy \\ &\leq 2\|\varphi\|_{L^2(\mathbb{T}^2)} \|\partial_2^{2k}(\varphi - T_0\varphi)\|_{L^2(\mathbb{T}^2)} \end{aligned} \quad (6.2)$$

Notice that if  $S$  is a composition of  $n$  any copies (not necessarily independant) of  $R$  or  $R^{-1}$ , then

$$\mathbb{E}[\|\partial_2^k(\varphi \circ S - \varphi)\|_{L^2(\mathbb{T}^2)}^2]^{1/2} \leq n\varepsilon.$$

Letting  $\tilde{R} : (x, y) \mapsto (x + \tilde{\alpha}, y + \tilde{\beta})$  be an independant copy of  $R$ , we set

$$S_0 = R^{-1} \circ \tilde{R}^{-1} \circ R \circ \tilde{R},$$

which is on the form

$$S_0(x, y) = (x, y + b(x))$$

where  $b(x) = \tilde{\beta}(x) + \beta(x + \tilde{\alpha}) - \tilde{\beta}(x + \alpha) - \beta(x)$ .

Next, let  $R_i : (x, y) \mapsto (x + \alpha_i, y + \beta_i(x))$  ( $i = 1, \dots, n$ ) be  $n$  copies of  $R$  for some  $n$  we will chose later such that  $R, \tilde{R}, R_1, \dots, R_n$  are independant. We set

$$S = (R_n \circ \dots \circ R_1)^{-1} \circ S_0 \circ (R_n \circ \dots \circ R_1),$$

which is on the form

$$S(x, y) = (x, y + b(x + s_n))$$

where  $s_n = \alpha_1 + \dots + \alpha_n$ .

By the previous remark, we have

$$\mathbb{E}[\|\partial_2^k(\varphi \circ S - \varphi)\|_{L^2(\mathbb{T}^2)}^2]^{1/2} \leq (2n + 4)\varepsilon.$$

Writing

$$\varphi(x, y) = \sum_{p \in \mathbb{Z}} \varphi_p(x) e^{2ip\pi y}, \quad (6.3)$$

we have

$$\begin{aligned} \varphi \circ S(x, y) - \varphi(x, y) &= \varphi(x, y + b(x + s_n)) - \varphi(x, y) \\ &= \sum_{p \in \mathbb{Z}} \varphi_p(x) (e^{2ip\pi b(x + s_n)} - 1) e^{2ip\pi y}. \end{aligned}$$

Seeing  $\varphi_p(x)(e^{2ip\pi b(x+s_n)} - 1)$  as a Fourier coefficient of  $y \mapsto \varphi \circ S(x, y) - \varphi(x, y)$  leads to

$$|\varphi_p(x)(e^{2ip\pi b(x+s_n)} - 1)| \leq \frac{C}{|p|^k} \|\partial_2^k(\varphi \circ S(x, \cdot) - \varphi(x, \cdot))\|_{L^2(\mathbb{T})}$$

Denoting  $u = \mathbb{E}[e^{2ip\pi b}]$ , we have  $\mathbb{E}[e^{2ip\pi b(x+s_n)}] = T_0^n u(x)$  (we still denote  $T_0$  the restriction of  $T_0$  on functions depending only of the first variable), and so

$$\|\varphi_p(T_0^n u - 1)\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{|p|^k} (2n + 4)\varepsilon,$$

hence

$$\left\| \varphi_p \left( \frac{1}{n} \sum_{j=0}^{n-1} T_0^j u - 1 \right) \right\|_{L^2(\mathbb{T})} \leq \frac{C}{|p|^k} (2n + 4)\varepsilon. \quad (6.4)$$

Setting  $m_p = \mathbb{E} \int_{\mathbb{T}} e^{2ip\pi b(t)} dt$ , writing  $u = v - T_0 v + m_p$  (by Lemma 1 of previous chapter) leads to

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} T_0^j u - m_p \right| \leq \frac{2\|v\|_0}{n} \leq \frac{C\|u\|_{l_0}}{n} \leq \frac{C|p|^{l_0}}{n}$$

for some constant  $C$  and integer  $l_0$ . Next, writing that for every integer  $j$ ,  $\|u\|_{j+1} \leq 1 + \|2\pi p b' u\|_j \leq 1 + C|p| \|u\|_j$  for some  $C = C(j)$ , we deduce by induction on  $k$  that  $\|u\|_j \leq C|p|^j$  for some  $C = C(j)$ . In particular,  $\|u\|_{l_0} \leq C|p|^{l_0}$  for some  $C$  and hence

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} T_0^j u - 1 \right| \geq |1 - m_p| - \frac{C_0|p|^{l_0}}{n}$$

for some constant  $C_0$ . Choosing  $n = \frac{2C|p|^{l_0}}{|1 - m_p|}$ , we obtain

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} T_0^j u - 1 \right| \geq \frac{|1 - m_p|}{2}$$

and so (6.4) becomes

$$\|\varphi_p\|_{L^2(\mathbb{T})} \leq \frac{Cn\varepsilon}{|p|^k|1 - m_p|} \leq \frac{C\varepsilon}{|p|^{k-l_0}|1 - m_p|^2} \quad (6.5)$$

for some constant  $C$ .

Let us look for a lower bound for  $|1 - m_p|$ . We have

$$|1 - m_p| \geq \mathbb{E} \int_{\mathbb{T}} (1 - \cos(2\pi p b(t))) dt \geq 2\pi^2 \mathbb{E} \int_{\mathbb{T}} \{pb(t)\}^2 dt,$$

where  $\{t\}$  is the distance of  $t$  to  $\mathbb{Z}$ . Next, using that  $\hat{b}(0) = 0$ , we have the alternative:

-either there exists  $t_0$  and a realization of  $b$  such that  $|pb(t_0)| = \frac{1}{2}$ , and then for  $|t - t_0| \leq \frac{1}{4p\|b'\|_0}$  we have

$$|pb(t) - pb(t_0)| \leq p\|b'\|_0|t - t_0| \leq \frac{1}{4}$$

hence  $\{pb(t)\} \geq \frac{1}{4}$ , so

$$\int_{\mathbb{T}} \{pb(t)\}^2 dt \geq \int_{\left[t_0 - \frac{1}{4p\|b'\|_0}, t_0 + \frac{1}{4p\|b'\|_0}\right]} \frac{dt}{16} = \frac{1}{64p\|b'\|_0},$$

and since the probability space is finite, this implies that

$$\mathbb{E} \int_{\mathbb{T}} \{pb(t)\}^2 dt \geq \frac{1}{Cp\|b'\|_0}$$

for some constant  $C$ .

-or  $|pb(t)| < \frac{1}{2}$  for every  $t$  in  $\mathbb{T}$  and every realization of  $b$  and then

$$\mathbb{E} \int_{\mathbb{T}} \{pb(t)\}^2 dt = \mathbb{E} \int_{\mathbb{T}} (pb(t))^2 dt = p^2 \Delta^2.$$

Thus in the two cases,

$$\mathbb{E} \int_{\mathbb{T}} \{pb(t)\}^2 dt \geq \frac{\Delta^2}{C|p|}$$

for some  $C$ , and

$$|1 - m_p| \geq \frac{\Delta^2}{C|p|}.$$

Finally, using these lower bounds of  $|1 - m_p|$ , (6.5) becomes

$$\|\varphi_p\|_{L^2(\mathbb{T})} \leq \frac{C\varepsilon}{|p|^{k-l_0-2}\Delta^4}, \quad (6.6)$$

and, choosing  $k = l_0 + 4$ , we have by (6.3), (6.6) and (6.2):

$$\|\varphi\|_{L^2(\mathbb{T}^2)} \leq \sum_{p \in \mathbb{Z}} \|\varphi_p\|_{L^2(\mathbb{T})} \leq \frac{C\varepsilon}{\Delta^4} \leq \frac{C}{\Delta^4} \|\varphi\|_{L^2(\mathbb{T}^2)}^{\frac{1}{2}} \|\partial_2^{2k}(\varphi - T_0\varphi)\|_{L^2(\mathbb{T}^2)}^{\frac{1}{2}}.$$

The result follows.  $\square$

**Lemma 4.** *There exists an integer  $a$  such that for every  $\varphi$  in  $E_{00}$  and every integers  $k_1$  and  $k_2$  larger than 1,*

$$\|\varphi\|_{k_1, k_2} \leq \frac{C}{\Delta^{8k_1}} \|\varphi - T_0\varphi\|_{k_1+a, k_2+ak_1}$$

for some  $C = C(k_1, k_2)$ .

*Proof.* The relation  $(I - T_0)\partial_1\varphi = \partial_1(\varphi - T_0\varphi) - \mathbb{E}[\beta' \partial_2\varphi \circ R]$  leads to the inequality

$$\|\partial_1^{k_1}(I - T_0)\partial_1\varphi\|_{L^2(\mathbb{T}^2)} \leq \|\partial_1^{k_1+1}(I - T_0)\varphi\|_{L^2(\mathbb{T}^2)} + C\|\partial_1^{k_1}\partial_2\varphi\|_{L^2(\mathbb{T}^2)}$$

Thus, for every positive integer  $j$  less than  $k_1$ ,

$$\|\partial_1^{k_1-j}(I - T_0)\partial_1^j\varphi\|_{L^2(\mathbb{T}^2)} \leq \|\partial_1^{k_1-(j-1)}(I - T_0)\partial_1^{j-1}\varphi\|_{L^2(\mathbb{T}^2)} + C\|\partial_1^{k_1-1}\partial_2\varphi\|_{L^2(\mathbb{T}^2)},$$

and we deduce by successive use of this inequality that

$$\|(I - T_0)\partial_1^{k_1}\varphi\|_{L^2(\mathbb{T}^2)} \leq \|\partial_1^{k_1}(I - T_0)\varphi\|_{L^2(\mathbb{T}^2)} + C\|\partial_1^{k_1-1}\partial_2\varphi\|_{L^2(\mathbb{T}^2)}$$

for some constant  $C$ . Next, Using Lemma 3 and the fact that  $\partial_2$  and  $T_0$  commute, we get

$$\|\partial_1^{k_1}\partial_2^{k_2}\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^8} \|(I - T_0)\partial_1^{k_1}\partial_2^{k_2+k_0}\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^8} (\|\partial_1^{k_1}\partial_2^{k_2+k_0}(I - T_0)\varphi\|_{L^2(\mathbb{T}^2)} + \|\partial_1^{k_1-1}\partial_2^{k_2+k_0+1}\varphi\|_{L^2(\mathbb{T}^2)})$$

and by iteration of this inequality,

$$\|\partial_1^{k_1}\partial_2^{k_2}\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^{8k_1}} \sum_{i=0}^{k_1} \|\partial_1^{k_1-i}\partial_2^{k_2+ai+k_0}(I - T_0)\varphi\|_{L^2(\mathbb{T}^2)}$$

with  $a = k_0 + 1$ . Now, If  $i_1, i_2$  are integers satisfying  $\frac{i_1}{k_1} + \frac{i_2}{k_2} \leq 1$ , then

$$\|\partial_1^{i_1}\partial_2^{i_2}\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^{8k_1}} \sum_{i=0}^{i_1} \|\partial_1^{i_1-i}\partial_2^{i_2+ai+k_0}(I - T_0)\varphi\|_{L^2(\mathbb{T}^2)}$$

with

$$\begin{aligned} \frac{i_1 - i}{k_1 + k_0} + \frac{i_2 + ai + k_0}{k_2 + ak_1} &= \frac{i_1}{k_1 + k_0} + \frac{i_2}{k_2 + ak_1} + \left( -\frac{1}{k_1 + k_0} + \frac{a}{k_2 + ak_1} \right) i + \frac{k_0}{k_2 + ak_1} \\ &\leq \frac{i_1}{k_1} \left( \frac{k_1}{k_1 + k_0} \right) + \frac{i_2}{k_2} \left( \frac{k_2}{k_2 + ak_1} \right) + \frac{k_0}{k_2 + ak_1} \\ &\leq \frac{i_1}{k_1} \left( \frac{k_1 + k_2}{k_1 + k_2 + k_0} \right) + \frac{i_2}{k_2} \left( \frac{k_1 + k_2}{k_1 + k_2 + k_0} \right) + \frac{k_0}{k_1 + k_2 + k_0} = 1 \end{aligned}$$

hence

$$\|\partial_1^{i_1}\partial_2^{i_2}\varphi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^{8k_1}} \|(I - T_0)\varphi\|_{k_1+k_0, k_2+ak_1}$$

and hence

$$\|\varphi\|_{k_1, k_2} \leq \frac{C}{\Delta^{8k_1}} \|(I - T_0)\varphi\|_{k_1+k_0, k_2+ak_1}.$$

□

**Lemma 5.** For every  $\psi$  in  $E_{00}$ , there exists  $\varphi$  in  $E_{00}$  such that  $\varphi - T_0\varphi = \psi$

*Proof.* We set  $S_n = \sum_{k=0}^{n-1} T_0^k$  and we fix  $\psi$  in  $E_{00}$ . Using Lemma 3, we have for every  $n$  the inequality

$$\|S_n\psi\|_{L^2(\mathbb{T}^2)} \leq \frac{C}{\Delta^8} \|(I - T_0)S_n\psi\|_{0,k_0} = \frac{C}{\Delta^8} \|(I - T_0^n)\psi\|_{0,k_0} \leq \frac{C}{\Delta^8} \|\psi\|_{0,k_0},$$

hence  $(S_n\psi)$  is bounded in  $L^2(\mathbb{T}^2)$ . Setting

$$\varphi_N = \frac{1}{N} \sum_{n=0}^{N-1} S_n\psi,$$

we have

$$(I - T_0)\varphi_N = \psi - \frac{1}{N}S_N\psi,$$

hence  $(I - T_0)\varphi_N$  converges to  $\psi$  in  $L^2(\mathbb{T}^2)$  as  $N$  tends to  $+\infty$ . Since  $(S_n\psi)$  is bounded in  $L^2(\mathbb{T}^2)$ , so does  $(\varphi_N)$  and there exists a subsequence of  $(\varphi_N)$  weakly converging to some function  $\varphi$  in  $L^2(\mathbb{T}^2)$ , which satisfies  $(I - T_0)\varphi = \psi$ .

We will prove that  $\varphi$  is  $C^\infty$  on  $\mathbb{T}^2$ . Denoting

$$\Delta_h\varphi(x, y) = \frac{\varphi(x, y+h) - \varphi(x, y)}{h},$$

we have  $(I - T_0)\Delta_h\varphi = \Delta_h\psi$ . Since  $(\|\Delta_h\psi\|_{0,k_0})_{h \neq 0}$  is bounded, by Lemma 3,  $(\Delta_h\varphi)_{h \neq 0}$  is bounded in  $L^2(\mathbb{T}^2)$ , hence  $\partial_2\varphi$  belongs to  $L^2(\mathbb{T}^2)$  and  $(I - T_0)\partial_2\varphi = \partial_2\psi$ . By induction,  $\partial_2^k\varphi$  belongs to  $L^2(\mathbb{T}^2)$  for any integer  $k$  and  $(I - T_0)\partial_2^k\varphi = \partial_2^k\psi$ .

Now, let  $k_1$  be in  $\mathbb{N}$  and let us assume that  $\partial_1^{k_1}\partial_2^{k_2}\varphi$  belongs to  $L^2(\mathbb{T}^2)$  for every integer  $k_2$ . Using that

$$\partial_1(I - T_0)\varphi = (I - T_0)\varphi + \mathbb{E}[\beta'\partial_2\varphi \circ R],$$

we can write

$$\partial_1^{k_1}\partial_2^{k_2}\psi = (I - T_0)\partial_1^{k_1}\partial_2^{k_2}\varphi + r$$

where  $\partial_1\partial_2^{k_2}r$  belongs to  $L^2(\mathbb{T}^2)$  for every  $k_2$ . Next, denoting

$$\widetilde{\Delta}_h\varphi(x, y) = \frac{\varphi(x+h, y) - \varphi(x, y)}{h},$$

we have

$$\widetilde{\Delta}_h\partial_1^{k_1}\partial_2^{k_2}\psi = (I - T_0)\widetilde{\Delta}_h\partial_1^{k_1}\partial_2^{k_2}\varphi + \mathbb{E}[\partial_1^{k_1}\partial_2^{k_2}\varphi(x+\alpha, y+\beta(x+h)) - \partial_1^{k_1}\partial_2^{k_2}\varphi(x+\alpha, y+\beta(x))] + \widetilde{\Delta}_h r$$

and we deduce that  $((I - T_0)\widetilde{\Delta}_h\partial_1^{k_1}\partial_2^{k_2}\varphi)_{h \neq 0}$  is bounded in  $L^2(\mathbb{T}^2)$  for every  $k_2$ , hence that  $(\|(I - T_0)(\widetilde{\Delta}_h\partial_1^{k_1}\partial_2^{k_2}\varphi)\|_{0,k_0})_{h \neq 0}$  is bounded for every  $k_2$ , and next by Lemma 3 that  $(\widetilde{\Delta}_h\partial_1^{k_1}\partial_2^{k_2}\varphi)_{h \neq 0}$  is bounded in  $L^2(\mathbb{T}^2)$ , and so that  $\partial_1^{k_1+1}\partial_2^{k_2}\varphi$  belongs to  $L^2(\mathbb{T}^2)$ .  $\square$

Now, we can finish the proof of Proposition 3. At this point we have proved that  $I - T_0$  is a bijection of  $E_0$  satisfying the estimates ii) and iii). Let us now prove i). Let  $\psi$  in  $E_0$ , and  $\varphi = (I - T_0)^{-1}\psi$ . By Lemma 2, we can find  $\tilde{\varphi}$  such that for  $k_1, k_2$  large enough

$$\begin{cases} \|\tilde{\varphi}\|_{k_1, k_2} \leq C\|\psi\|_{2k_1, 2k_2} \\ \|(I - T_0)(\tilde{\varphi} - \varphi)\|_{k_1, k_2} = \|(I - T_0)\tilde{\varphi} - \psi\|_{k_1, k_2} \leq C\Delta^p\|\psi\|_{2k_1, 2k_2}. \end{cases}$$

Using Lemma 4, the second inequality implies that

$$\|\varphi - \tilde{\varphi}\|_{k_1, k_2} \leq \frac{C}{\Delta^{8k_1}} \|(I - T_0)(\varphi - \tilde{\varphi})\|_{k_1 + a, k_2 + ak_1} \leq \Delta^{p-8k_1} \|\psi\|_{k_1, \bar{k}_2}$$

with  $\bar{k}_1 = 2(k_1 + a)$ ,  $\bar{k}_2 = 2(k_2 + ak_1)$ . Choosing  $p = 8k_1$ , we deduce

$$\|\varphi\|_{k_1, k_2} \leq \|\tilde{\varphi}\|_{k_1, k_2} + \|\varphi - \tilde{\varphi}\|_{k_1, k_2} \leq C\|\psi\|_{\bar{k}_1, \bar{k}_2}.$$

□

### 6.3 Estimates on the stationary measure and the sum of Lyapunov exponents, first conjugation

We fix a diophantine random Anzai cocycle  $R$  as in the previous section,  $f = R + \zeta$  in  $\widehat{Diff}_0^\infty(\mathbb{T}^2)$  a perturbation of  $R$ , and  $\mu$  a stationary measure of  $f$ . For  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ , we will denote

$$U\varphi = (I - T_0)^{-1}(\varphi - \varphi(0)).$$

If  $\mu$  and  $\nu$  are probability measures on  $\mathbb{T}^2$ , we denote

$$d_k^*(\mu, \nu) = \sup \left\{ \int_{\mathbb{T}^2} \varphi d\mu - \int_{\mathbb{T}^2} \varphi d\nu \mid \varphi \in C^k(\mathbb{T}^d), \|\varphi\|_k \leq 1 \right\}.$$

**Proposition 4.** *There exists an integer  $K$  such that for every  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ , we have*

$$\int_{\mathbb{T}^2} \varphi d\mu = \int_{\mathbb{T}^2} \varphi dz + \int_{\mathbb{T}^2} (\vec{\nabla} U\varphi) \cdot \bar{\zeta} dz + O(\varepsilon^2 \|\varphi\|_K),$$

where  $\bar{\zeta} = \mathbb{E}[\zeta \circ R^{-1}]$ ,  $\varepsilon = \|\zeta\|_K$ . In particular, if  $\bar{\zeta} = \text{constant} + O(\varepsilon^2)$ , then  $d_K^*(\mu, dz) = O(\varepsilon^2)$ .

*Proof.* For  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ , setting  $T\varphi = \mathbb{E}[\varphi \circ f]$  and  $T_0\varphi = \mathbb{E}[\varphi \circ R]$ , we have  $|T\varphi - T_0\varphi| \leq \varepsilon\|\varphi\|_1$ . Using the  $T$ -invariance of  $\mu$ , we get

$$\left| \int_{\mathbb{T}^2} (\varphi - T_0\varphi) d\mu \right| \leq \varepsilon\|\varphi\|_1.$$

We now use Proposition 3. For  $\psi$  in  $C^\infty(\mathbb{T}^2)$ , applying the previous inequality to  $\varphi = U\psi$  gives for some  $C$  and  $K$

$$\left| \int_{\mathbb{T}^2} \psi d\mu - \int_{\mathbb{T}^2} \psi dz \right| \leq C\epsilon \|U\psi\|_1 \leq C\epsilon \|\psi\|_K.$$

Thus  $\mu$  is equal to Lebesgue measure in first approximation. To get a more precise approximation, we now write  $T\varphi - T_0\varphi = \mathbb{E}[(\vec{\nabla}\varphi \circ R) \cdot \zeta] + O(\epsilon^2 \|\varphi\|_2)$ , which gives by integration and our previous estimate on  $\mu$ :

$$\begin{aligned} \int_{\mathbb{T}^2} (\varphi - T_0\varphi) d\mu &= \int_{\mathbb{T}^2} \mathbb{E}[(\vec{\nabla}\varphi \circ R) \cdot \zeta] d\mu + O(\epsilon^2 \|\varphi\|_2) \\ &= \int_{\mathbb{T}^2} \mathbb{E}[(\vec{\nabla}\varphi \circ R) \cdot \zeta] dz + O(\epsilon^2 \|\varphi\|_K) \\ &= \int_{\mathbb{T}^2} \vec{\nabla}\varphi \cdot \bar{\zeta} dz + O(\epsilon^2 \|\varphi\|_K). \end{aligned}$$

Applying as previously this formula to  $\varphi = U\psi$  this gives for some  $K_1$ :

$$\int_{\mathbb{T}^2} \psi d\mu - \int_{\mathbb{T}^2} \psi dz = \int_{\mathbb{T}^2} (\vec{\nabla}U\psi) \cdot \bar{\zeta} dz + O(\epsilon^2 \|\psi\|_{K_1}).$$

□

**Proposition 5.** *i) There exists a diffeomorphism  $G = Id + g$  such that, denoting  $\tilde{f} = G^{-1}fG = R + \tilde{\zeta}$ :*

- For any  $k_1, k_2$  larger than 1,

$$\|g\|_{k_1, k_2} \leq \frac{C}{\Delta^{16k_1}} \|\zeta\|_{k_1 + a, k_2 + ak_1}.$$

for some  $C = C(k_1, k_2)$ .

- For any  $k$ , there exists  $K$  such that

$$\|g\|_k \leq C \|\zeta\|_K$$

for some  $C = C(k)$ .

- For any integer  $k$ , there exists an integer  $K$  such that

$$\left\| \mathbb{E} \left[ \tilde{\zeta} \circ R^{-1} - \int_{\mathbb{T}^2} \tilde{\zeta}(z) dz \right] \right\|_k \leq C \|\zeta\|_K^2$$

for some  $C = C(k)$ .

*ii) For any  $\delta > 0$  and integer  $p$ , there exists a diffeomorphism  $G = Id + g$  such that, denoting  $\tilde{f} = G^{-1}fG = R + \tilde{\zeta}$ :*

- For any  $k_1, k_2$  large enough,

$$\|g\|_{k_1, k_2} \leq C \|\zeta\|_{(1+\delta)k_1, (1+\delta)k_2}.$$

for some  $C = C(k_1, k_2)$ .



6.3. ESTIMATES ON THE STATIONARY MEASURE AND THE SUM OF LYAPUNOV EXPONENTS, FIRST COM

- For any integer  $k$ , there exists an integer  $K$  such that

$$\left\| \mathbb{E} \left[ \tilde{\zeta} \circ R^{-1} - \int_{\mathbb{T}^2} \tilde{\zeta}(z) dz \right] \right\|_k \leq C(\Delta^p \|\zeta\|_K + \|\zeta\|_K^2)$$

*Proof.* Let  $G = Id + g$  in  $Dif f_0^\infty(\mathbb{T}^2)$  with  $\|g\|_1 = O(\|\zeta\|_{k_0})$  for some integer  $k_0$ , and let  $\tilde{f} = G^{-1}fG = R + \tilde{\zeta}$ . Setting  $\zeta = (\zeta_1, \zeta_2)$ ,  $\tilde{\zeta} = (\tilde{\zeta}_1, \tilde{\zeta}_2)$  and  $G = (G_1, G_2) = Id + (g_1, g_2)$ , we have

$$\begin{cases} \tilde{\zeta}_1 = \zeta_1 + (g_1 - g_1 \circ R) + O(\|\zeta\|_{k_0}^2) \\ \tilde{\zeta}_2 = \zeta_2 + (\beta \circ G_1 - \beta) + (g_2 - g_2 \circ R) + O(\|\zeta\|_{k_0}^2), \end{cases}$$

hence, denoting  $\tilde{T}_0\varphi = \mathbb{E}[\varphi \circ R^{-1}]$ ,  $\eta_1 = \mathbb{E}[\zeta_1 \circ R^{-1}]$  and  $\eta_2 = \mathbb{E}[(\zeta_2 + \beta \circ G_1 - \beta) \circ R^{-1}]$ , we have

$$\begin{cases} \tilde{\zeta}_1 = \eta_1 - (g_1 - \tilde{T}_0 g_1) + O(\|\zeta\|_{k_0}^2) \\ \tilde{\zeta}_2 = \eta_2 - (g_2 - \tilde{T}_0 g_2) + O(\|\zeta\|_{k_0}^2) \end{cases}.$$

Thus, using Proposition 3 with  $\tilde{T}_0$  instead of  $T_0$  and denoting  $\tilde{U}$  the associated inverse operator, setting  $g_1 = \tilde{U}(\tilde{T}_0 \zeta_1)$  and next  $g_2 = \tilde{U}(\tilde{T}_0(\zeta_2 \circ R^{-1} + (\beta \circ G_1 - \beta)))$  we obtain the point i). If instead we define  $g_1$  and  $g_2$  using the point iii) of Proposition 3, we obtain the point ii).  $\square$

**Proposition 6.** *If  $f = R + \tilde{\zeta}$  satisfies for some  $\varepsilon > 0$  and some integer  $k_0$  large enough the inequalities*

$$\begin{cases} \|\zeta\|_{k_0} \leq \varepsilon \\ \left\| \mathbb{E} \left[ \zeta \circ R^{-1} - \int_{\mathbb{T}^2} \zeta(z) dz \right] \right\|_k \leq \varepsilon^2 \end{cases},$$

and if moreover  $\gamma_1(f, \mu) + \gamma_2(f, \mu) \geq 0$ , or  $f_*\mu = \mu$  a.s., then we have,

$$\|\det(d_z f) - 1\|_{L^2(\mathbb{T}^2)} = O(\varepsilon^{\frac{3}{2}}).$$

where  $d_z f$  is the differential of  $f$  at the point  $z$ .

*Proof.* By Proposition 4,  $d_k^*(\mu, dz) = O(\varepsilon^2)$ . In consequence, we have

$$\begin{aligned} \gamma_1(f, \mu) + \gamma_2(f, \mu) &= \mathbb{E} \int_{\mathbb{T}^2} \ln \det(d_z f) d\mu(z) \\ &= \mathbb{E} \int_{\mathbb{T}^2} \ln \det(d_z f) dz + O(\varepsilon^3) \\ &= \mathbb{E} \int_{\mathbb{T}^2} (\det(d_z f) - 1) dz - \frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2} (\det(d_z f) - 1)^2 dz + O(\varepsilon^3). \end{aligned}$$

Since  $f$  is a diffeomorphism of  $\mathbb{T}^2$ ,  $\int_{\mathbb{T}^2} \det(d_z f) dz = 1$  hence

$$\gamma_1(f, \mu) + \gamma_2(f, \mu) = -\frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2} (\det(d_z f) - 1)^2 dz + O(\varepsilon^3).$$

This gives the expected estimate when  $\gamma_1(f, \mu) + \gamma_2(f, \mu) = 0$ .

To obtain this same estimate when  $f_*\mu = \mu$  a.s., we write that for  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ ,

$$\int_{\mathbb{T}^2} \varphi d\mu = \int_{\mathbb{T}^2} \varphi(z) dz + O(\varepsilon^2 \|\varphi\|_k)$$

and

$$\begin{aligned} \int_{\mathbb{T}^2} \varphi \circ f^{-1} d\mu &= \int_{\mathbb{T}^2} \varphi(f^{-1}(z)) dz + O(\varepsilon^2 \|\varphi \circ f^{-1}\|_k) \\ &= \int_{\mathbb{T}^2} \varphi(z) \det(d_z f) dz + O(\varepsilon^2 \|\varphi\|_k) \end{aligned}$$

hence

$$\int_{\mathbb{T}^2} \varphi(z) (\det(d_z f) - 1) dz = O(\varepsilon^2 \|\varphi\|_k)$$

Setting  $\varphi = \det(d_z f) - 1$  and taking the expectation of the last formula, we get

$$\mathbb{E} \int_{\mathbb{T}^2} (\det(d_z f) - 1)^2 dz = O(\varepsilon^3)$$

□

Thus, using the last two propositions, under assumption (I) or (II) of Theorem 1 we can conjugate  $f$  to some  $\tilde{f}$  which is close to be Lebesgue invariant. In the next section, we will use more finely these assumptions in studying the action of the differential of  $\tilde{f}$  on the projective space  $\mathcal{P}(\mathbb{R}^2)$ , and we will obtain that in fact  $\tilde{f}$  is close to an Anzai cocycle, up to an algebraic conjugation.

## 6.4 Projective action

The aim of this section is to prove the following proposition:

**Proposition 7.** *There exist a positive real number  $\delta_0$  (depending only on  $A, \sigma$  and  $M_0$ ) such that for every  $\kappa > 0$  there exists a constant  $C$  and an integer  $k$  such that if  $f = R + \zeta$  satisfies assumption (I) or (II) of Theorem 1, and that in addition  $\mathbb{E}[\beta(x - \alpha)] = \hat{\beta}$ ,  $\|f - R\|_k \leq \varepsilon$  and  $\|\mathbb{E}[\zeta \circ R^{-1} - \hat{\zeta}(0)]\|_k \leq \varepsilon^2$ , then there exists a matrix  $P$  in  $Sl_2(\mathbb{Z})$  and an Anzai cocycle  $\tilde{R} : (x, y) \rightarrow (x + \tilde{\alpha}, y + \tilde{\beta}(x))$  such that  $\|P^{-1} f P - \tilde{R}\|_0 \leq C\varepsilon^{1+\delta_0}$ ,  $\|P\| \leq C\varepsilon^{-\kappa}$ , and  $P = Id$  if  $\|\beta'\|_0 \geq 2\varepsilon$ .*

*Proof.* We assume that  $R$  satisfies the condition  $\mathbb{E}[\beta(x - \alpha)] = \hat{\beta}$ . We fix  $\varepsilon > 0$  and an integer  $k_0$  large enough and we assume that  $f$  satisfies the estimates  $\|f - R\|_k \leq \varepsilon$  and  $\|\mathbb{E}[\zeta \circ R^{-1} - \hat{\zeta}(0)]\|_k \leq \varepsilon^2$ .

By Proposition 6, we already know that

$$\det(d_z f) = 1 + O(\varepsilon^{\frac{3}{2}}) \tag{6.7}$$

To prove the more precise fact that  $f$  is close, up to an algebraic conjugation, to a random Anzai cocycle, we will study the action of the random cocycle

$(z, h) \rightarrow (f(z), d_z f(h))$  on  $\mathbb{T}^2 \times \mathcal{P}(\mathbb{R}^2)$ , where  $\mathcal{P}(\mathbb{R}^2)$  is the projective plane. Notice that in the case where  $f$  is close to a translation of  $\mathbb{T}^2$ ,  $d_z f$  is close to  $Id$ . And in the general case where  $f$  is close to an Anzai cocycle, we can conjugate the cocycle by  $(z, h) \rightarrow (z, D(h))$  where  $D = \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$  to obtain for an appropriate choice of  $\lambda$  a cocycle whose second component is close to  $Id$ .

Let us state some notations:

- Let  $f_1$  and  $f_2$  be the coordinates of  $f$  and for  $z$  in  $\mathbb{T}^2$ , and let  $\mathcal{J}_z f = \begin{pmatrix} \partial_1 f_1 & \partial_2 f_1 \\ \partial_1 f_2 & \partial_2 f_2 \end{pmatrix}$  be the Jacobian matrix of  $f$  at point  $z$ .
- Let  $\lambda \leq 1$  a positive real number we will chose later,  $D = \begin{pmatrix} \lambda & 0 \\ 0 & \frac{1}{\lambda} \end{pmatrix}$  and  $M_z = D(\mathcal{J}_z f)D^{-1} = Id + N_z$ . For  $z = (x, y)$  in  $\mathbb{T}^2$ , we have

$$N_z = \begin{pmatrix} \partial_1 \zeta_1(z) & \frac{\partial_2 \zeta_1(z)}{\lambda} \\ \lambda(\partial_1 \zeta_2(z) + \beta'(x)) & \partial_2 \zeta_2(z) \end{pmatrix}.$$

We denote  $m = \|\beta'\|_0$  and  $\varepsilon_1 = \lambda m + \frac{\varepsilon}{\lambda}$  so that  $\|N_z\| \leq \varepsilon_1$ .

- Let

$$u_z = Id + \eta_z$$

the diffeomorphism of  $\mathbb{T}$  induced by the projective action of  $M_z$ , that is such that  $\frac{M_z(e^{i\pi\theta})}{|M_z(e^{i\pi\theta})|} = e^{i\pi u_z(\theta)}$ .

- Let  $U$  the random cocycle  $(z, \theta) \mapsto (f(z), u_z(\theta))$  on  $\mathbb{T}^2 \times \mathbb{T}$ ,  $\hat{\mu}$  a stationary measure of  $U$ , and  $\mu, \nu$  the projections of  $\hat{\mu}$  on the first and second factors.

Let us begin by proving an elementary estimate of  $\eta_z$ :  
Writing that

$$e^{2i\pi u_z(\theta)} = \frac{M_z e^{i\pi\theta}}{M_z e^{i\pi\theta}} = \frac{e^{i\pi\theta} + N_z e^{i\pi\theta}}{e^{-i\pi\theta} + N_z e^{i\pi\theta}} = e^{2i\pi\theta} \left( 1 + 2Im(N_z e^{i\pi\theta} \cdot e^{-i\pi\theta}) + O(\varepsilon_1^2) \right)$$

leads to the estimate

$$\begin{aligned} \eta_z(\theta) &= \frac{1}{2\pi} Im(N_z e^{i\pi\theta} e^{-i\pi\theta}) + O(\varepsilon_1^2) \\ &= \lambda(\partial_1 \zeta_2 + \beta') \cos^2(\pi\theta) + (\partial_2 \zeta_2 - \partial_1 \zeta_1) \cos(\pi\theta) \sin(\pi\theta) \\ &\quad - \frac{\partial_2 \zeta_1}{\lambda} \sin^2(\pi\theta) + O(\varepsilon_1^2) \end{aligned} \quad (6.8)$$

Proof of Proposition 7 will be based on the study of the measure  $\hat{\mu}$ . By Proposition 6, we already know that  $\mu$  is close to Lebesgue measure. In the next lemma we will prove that  $\hat{\mu}$  is close to product measure.

**Lemma 6.** *Under assumptions of Proposition 7 and previous notations, there exists an integer  $k$  such that for every  $\varphi$  in  $C^\infty(\mathbb{T}^2 \times \mathbb{T})$ ,*

$$\iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) d\hat{\mu}(z, \theta) = \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) dz d\nu(\theta) + O(\varepsilon_1^2 \|\varphi\|_k)$$

*Proof.* We will follow the same scheme as Proposition 4. As in previous sections, for  $\varphi$  in  $C^\infty(\mathbb{T}^2)$  we denote  $T_0\varphi = \mathbb{E}[\varphi \circ R^{-1}]$  and  $U = (I - T_0)^{-1}(\varphi - \hat{\varphi}(0))$ . We know by Proposition 3 that for every  $k$ , there exists  $K$  and  $C$  such that for every  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ ,  $\|U\varphi\|_k \leq C\|\varphi\|_K$ .

For  $\varphi$  in  $C^\infty(\mathbb{T}^2 \times \mathbb{T})$ ,

$$\begin{aligned} \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) d\hat{\mu}(z, \theta) &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(f(z), u_z(\theta)) d\hat{\mu}(z, \theta) \\ &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(R(z), \theta) d\hat{\mu}(z, \theta) + O(\varepsilon_1 \|\varphi\|_1) \end{aligned}$$

and so, for  $\psi \in C^\infty(\mathbb{T}^2 \times \mathbb{T})$ , choosing  $\varphi(\cdot, \theta) = U(\psi(\cdot, \theta))$  we obtain for some integer  $k$ :

$$\iint_{\mathbb{T}^2} \psi(z, \theta) d\hat{\mu}(z, \theta) = \iint_{\mathbb{T}^2 \times \mathbb{T}} \psi(z, \theta) dz d\nu(\theta) + O(\varepsilon_1 \|\psi\|_k). \quad (6.9)$$

Next, for  $\varphi$  in  $C^\infty(\mathbb{T}^2 \times \mathbb{T})$ ,

$$\mathbb{E}[\varphi(f(z), u_z(\theta))] = \mathbb{E}[\varphi(R(z), \theta)] + \mathbb{E}[\vec{\nabla}_z \varphi(R(z), \theta) \cdot \zeta(z)] + \mathbb{E}[\partial_\theta \varphi(R(z), \theta) \eta_z(\theta)] + O(\varepsilon_1^2 \|\varphi\|_2)$$

hence by (6.9)

$$\begin{aligned} \iint_{\mathbb{T}^2 \times \mathbb{T}} (\varphi(z, \theta) - \mathbb{E}[\varphi(R(z), \theta)]) d\hat{\mu}(z, \theta) &= \iint_{\mathbb{T}^2 \times \mathbb{T}} \vec{\nabla}_z \varphi(z, \theta) \mathbb{E}[\zeta(R^{-1}(z))] dz d\nu(\theta) \\ &\quad + \iint_{\mathbb{T}^2 \times \mathbb{T}} \partial_\theta \varphi(z, \theta) \mathbb{E}[\eta_{R^{-1}(z)}(\theta)] dz d\nu(\theta) \\ &\quad + O(\varepsilon_1^2 \|\varphi\|_{k+1}). \end{aligned}$$

Since  $\mathbb{E}[\zeta_i \circ R^{-1} - \hat{\zeta}_i(0)] \leq \varepsilon^2$  by assumption and  $\varepsilon \geq \varepsilon_1$ , the integral

$$\iint_{\mathbb{T}^2 \times \mathbb{T}} \vec{\nabla}_z \varphi(z, \theta) \mathbb{E}[\zeta(R^{-1}(z))] dz d\nu(\theta)$$

is bounded by above by  $\varepsilon_1^2 \|\varphi\|_1$  for some. We estimate the integral

$$\iint_{\mathbb{T}^2 \times \mathbb{T}} \partial_\theta \varphi(z, \theta) \mathbb{E}[\eta_{R^{-1}(z)}(\theta)] dz d\nu(\theta)$$

thanks to (6.8). We have  $|\mathbb{E}[\vec{\nabla} \zeta_i \circ R^{-1}]| \leq \varepsilon^2 \leq \varepsilon_1^2$  and:

$$\begin{aligned} \partial_2 (\mathbb{E}[\zeta_i \circ R^{-1}]) &= \mathbb{E}[\partial_2 \zeta_i \circ R^{-1}] \\ \partial_1 (\mathbb{E}[\zeta_i \circ R^{-1}]) &= \mathbb{E}[\partial_1 \zeta_i \circ R^{-1}] + \mathbb{E}[\beta' \partial_2 \zeta_i \circ R^{-1}] = \mathbb{E}[\partial_1 \zeta_i \circ R^{-1}] + O(\varepsilon_1^2) \\ \mathbb{E}[\beta'(x - \alpha)] &= 0, \end{aligned}$$

hence  $\mathbb{E}[\eta_{R^{-1}(z)}(\theta)] = O(\varepsilon_1^2)$ , and the integral  $\iint_{\mathbb{T}^2 \times \mathbb{T}} \partial_\theta \varphi(z, \theta) \mathbb{E}[\eta_{R^{-1}(z)}(\theta)] dz d\nu(\theta)$  is bounded by above by  $C\varepsilon_1^2 \|\varphi\|_1$ .

Consequently,

$$\iint_{\mathbb{T}^2 \times \mathbb{T}} (\varphi(z, \theta) - \mathbb{E}[\varphi(R(z), \theta)]) d\hat{\mu}(z, \theta) = O(\varepsilon_1^2 \|\varphi\|_k),$$

and choosing  $\varphi(\cdot, \theta) = U(\psi(\cdot, \theta))$  for  $\psi$  in  $C^\infty(\mathbb{T}^2 \times \mathbb{T})$  we obtain

$$\iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) d\hat{\mu}(z, \theta) = \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) dz d\nu(\theta) + O(\varepsilon_1^2 \|\varphi\|_k)$$

for some integer  $K$ . □

Now, we will study  $\nu$  the second factor of  $\hat{\mu}$ . We will use that  $\nu$  is almost invariant by some elliptic operator, and use the following proposition

**Proposition 8.** *Let  $d$  and  $k$  be two integers. Let  $S$  be an operator on  $C^\infty(\mathbb{T}^2)$  defined by*

$$S\varphi = \varphi + (a\varphi' + b\varphi'')$$

where  $a, b$  belong to  $C^\infty(\mathbb{T})$ ,  $b > 0$  on  $\mathbb{T}^2$  and  $a$  vanishes at most  $d$  times on  $\mathbb{T}$ . We set  $M = \max(\|a\|_k, \|b\|_k)$ ,  $m = \inf_{\mathbb{T}} b$  and  $q = 1 + \frac{M}{m}$ . Then for any probability  $\nu$  on  $\mathbb{T}$ , denoting  $\varepsilon = d_k^*(S^* \nu, \nu)$  there exists a smooth density  $h(\theta) d\theta$  satisfying  $\|h\|_k \leq Cq^K$ ,  $\|\ln h\|_k \leq Cq^K$  and

$$d_k^*(\nu, h(\theta) d\theta) \leq Cq^K \left( \frac{\varepsilon}{M} \right)^{\frac{1}{3}}$$

where  $C$  and  $K$  are positive constants depending only on  $k$  and  $d$ .

Before proving this proposition let us see how it applies in our context.

**Corollary 1.** *There exists a constant  $C$  and an integer  $K$  such that for every  $\delta > 0$ , we have the following alternative:*

- either there exists  $\theta_0$  such that

$$\|\eta(\cdot, \theta_0)\|_{L^2(\mathbb{T}^2)} \leq \varepsilon_1^{1+\delta}$$

- or there exists a smooth density  $g(\theta) d\theta$  satisfying  $\|g\|_k \leq C\varepsilon_1^{-K\delta}$ ,  $\|\ln g\|_k \leq C\varepsilon_1^{-K\delta}$  and  $d_k^*(\nu, g(\theta) d\theta) \leq C\varepsilon_1^{\frac{1}{3}-K\delta}$ .

*Proof.* By Lemma 6, there exists an integer  $k$  and a constant  $C$  such that  $d_k^*(\hat{\mu}, dz \times \nu) \leq C\varepsilon_1^2$ . We set

$$\begin{cases} a(\theta) = \int_{\mathbb{T}^2} \eta_z(\theta) dz \\ b(\theta) = \frac{1}{2} \mathbb{E} \int_{\mathbb{T}} \eta_z(\theta)^2 dz = \frac{1}{2} \|\eta(\cdot, \theta)\|_{L^2(\mathbb{T}^2)}^2 \end{cases}, \quad (6.10)$$

By (6.8) we have the estimates  $\|\eta_z\|_k = O(\varepsilon_1)$ ,  $\|b\|_k = O(\varepsilon_1^2)$ , and also  $\|a\|_k = O(\varepsilon_1^2)$  using that  $\int_{\mathbb{T}^2} N_z dz = 0$ . We set  $S\varphi = \varphi + a\varphi' + b\varphi''$ . If  $\varphi$  is a  $C^\infty$  function on  $\mathbb{T}$ , then

$$\begin{aligned} \int_{\mathbb{T}} \varphi(\theta) d\nu(\theta) &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \varphi(u_z(\theta)) d\hat{\mu}(z, \theta) \\ &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \left( \varphi(\theta) + \varphi'(\theta)\eta_z(\theta) + \frac{1}{2}\varphi''(\theta)\eta_z(\theta)^2 \right) d\hat{\mu}(z, \theta) + O(\varepsilon_1^3 \|\varphi\|_2) \\ &= \int_{\mathbb{T}} \varphi(\theta) d\nu(\theta) + \int_{\mathbb{T}} (a(\theta)\varphi'(\theta) + b(\theta)\varphi''(\theta)) d\nu(\theta) + O(\varepsilon_1^3 \|\varphi\|_k) \\ &= \int_{\mathbb{T}} S\varphi d\nu + O(\varepsilon_1^3 \|\varphi\|_k). \end{aligned}$$

Thus,  $d_k^*(S^*v, v) \leq C\varepsilon_1^3$  for some constant  $C$ . In order to use Proposition 8, we need a bound on the number of zeros of  $a$ . This is not obvious that such a bound exists, and we will rather consider a polynomial approximation of  $a$ . From the relation

$$e^{2i\pi u_z(\theta)} = \frac{M_z e^{i\pi\theta}}{M_z e^{i\pi\theta}} = \frac{e^{i\pi\theta} + N_z e^{i\pi\theta}}{e^{-i\pi\theta} + N_z e^{i\pi\theta}} = e^{2i\pi\theta} \left( \frac{1 + N_z e^{i\pi\theta} e^{-i\pi\theta}}{e^{-i\pi\theta} + N_z e^{i\pi\theta} e^{-i\pi\theta}} \right)$$

we obtain

$$\eta_z(\theta) = \frac{1}{2i\pi} \ln \left( \frac{1 + N_z e^{i\pi\theta} e^{-i\pi\theta}}{e^{-i\pi\theta} + N_z e^{i\pi\theta} e^{-i\pi\theta}} \right)$$

and by Taylor expansion we deduce the existence of a trigonometric polynomial  $\tilde{\eta}_z$  of degree 2 such that  $\eta_z = \tilde{\eta}_z + O(\varepsilon_1^3)$ . Thus  $\tilde{a} = \int_{\mathbb{T}^2} \tilde{\eta}_z dz$  is trigonometric polynomial of degree 2 satisfying  $\tilde{a} = a + O(\varepsilon_1^3)$ . Since  $\tilde{a}$  has at most 4 zeros on  $\mathbb{T}$ , we can apply Proposition 8 on the operator  $\tilde{S}\varphi = \varphi + \tilde{a}\varphi' + b\varphi''$ .

We have for some constant  $C$  the estimates  $d_k(\tilde{S}^*v, v) \leq C\varepsilon_1^3$  and  $\max(\|a\|_k, \|b\|_k) \leq C\varepsilon_1^2$ . If moreover we assume that  $b \geq \varepsilon^{2+\delta}$  on  $\mathbb{T}^2$ , then we obtain by Proposition 8 a density probability  $h(t)dt$  such that  $\|h\|_k \leq C\varepsilon_1^{-K\delta}$ ,  $\|\ln h\|_k \leq C\varepsilon_1^{-K\delta}$  and  $d_k^*(v, h(\theta)d\theta) \leq C\varepsilon_1^{\frac{1}{3}-K\delta}$  for some constants  $C$  and  $K$ .  $\square$

Now we will prove Proposition 8:

*Proof.* We begin by a lemma studying solutions of some elliptic differential equations:

**Lemma 7.** *Let  $a, b$  be in  $C^\infty(\mathbb{T})$  and  $k$  be an integer. and let  $M = \max(\|a\|_k, \|b\|_k)$ ,  $m = \inf_{\mathbb{T}} b$  and  $q = 1 + \frac{M}{m}$ . Then there exists an unique smooth periodic function  $g$  such that  $\int_{\mathbb{T}} g(t)dt = 1$  and  $ag + bg'$  is constant. Moreover,  $g$  is positive,  $\|g\|_k \leq Cq^K$  and  $\|\ln g\|_k \leq Cq^K$  where  $C$  and  $K$  are positive constants depending only on  $k$ .*

*Proof.* Let  $Q(\theta) = -\int_0^\theta \frac{a(t)}{b(t)} dt$ . For  $c_0$  in  $\mathbb{R}$ ,  $g : \mathbb{R} \rightarrow \mathbb{R}$  is solution of the equation  $ag + bg' = c_0$  if and only if  $(ge^Q)' = \frac{c_0 e^Q}{b}$  hence if and only if there exists  $d_0$  in  $\mathbb{R}$  such that

$$\forall \theta \in \mathbb{R}, g(\theta) = d_0 e^{-Q(\theta)} + c_0 R(\theta),$$

where  $R(\theta) = \int_0^\theta \frac{e^{Q(t)-Q(\theta)}}{b(t)} dt$ . By Cauchy-Lipschitz theorem,  $g$  is 1-periodic if and only if  $g(1) = g(0)$  since  $\theta \mapsto g(1 + \theta)$  is solution of the same equation, and so if and only if  $d_0(e^{-Q(1)} - 1) + c_0 R(1) = 0$ . Since  $R(1)$  is non null, we deduce that the space of 1-periodic functions  $g$  such that  $ag + bg'$  is constant has dimension 1.

If  $g$  in  $C^\infty(\mathbb{T})$  is such that  $ag + bg'$  is constant equal to  $c_0$ , then:  
 -If  $c_0 = 0$ ,  $g$  cannot vanish by Cauchy-Lipschitz theorem, unless it is identically null.  
 -If  $c_0 \neq 0$ , at each point  $\theta$  where  $g$  vanishes we must have  $b(\theta)g'(\theta) = c_0$ , so  $g'(\theta)$  has the same strict sign as  $c_0$ . But if  $\theta_1$  and  $\theta_2$  are two consecutive zeros of  $g$ ,  $g'(\theta_1)$  and  $g'(\theta_2)$  cannot have same sign. Consequently,  $g$  cannot vanish on  $\mathbb{R}$ .

Thus, unless  $g$  is identically null,  $g$  does not vanish on  $\mathbb{T}$ , and in particular,  $\int_{\mathbb{T}} g(t) dt \neq 0$ .

In any case, there exists a unique function  $g$  such that  $ag + bg'$  is constant and  $\int_{\mathbb{T}} g(t) dt = 1$ . Moreover, since  $g$  does not vanish, it is necessarily positive.

Let us now prove that  $g$  satisfies the claimed estimates. Let  $\theta_0$  be a point where  $g$  is minimal. Then  $g'(\theta_0) = 0$  hence  $|c_0| = |a(\theta_0)g(\theta_0)| \leq M \inf(g)$ , and hence

$$\left| \frac{g'}{g} \right| \leq \left| \frac{c_0}{bg} \right| + \left| \frac{a}{b} \right| \leq 2q.$$

For  $k \geq 2$ , derivating  $k - 1$  times the relation  $ag + bg' = c_0$  gives

$$bg^{(k)} + \sum_{j=0}^{k-1} \alpha_j g^{(j)} = 0,$$

where  $q_j$  is some linear combination of derivatives of  $a$  and  $b$  of order less than  $k$ . In particular we have  $|\alpha_j| \leq CM$  for some  $C = C(k)$  and we deduce that

$$\left| \frac{g^{(k)}}{g} \right| \leq \frac{1}{|b|} \sum_{j=0}^{k-1} |\alpha_j| \left| \frac{g^{(j)}}{g} \right| \leq Cq \sup_{0 \leq j \leq k-1} \left| \frac{g^{(j)}}{g} \right|$$

for some  $C = C(k)$ , and so by induction,

$$\left| \frac{g^{(k)}}{g} \right| \leq Cq^k \tag{6.11}$$

for some  $C = C(k)$ .

Next, for every  $\theta$  in  $\mathbb{T}$ ,

$$\left| g(\theta) - \int_{\mathbb{T}} g(t) dt \right| \leq \int_{\mathbb{T}} |g'(t)| dt \leq \left( \sup_{\mathbb{T}} \left| \frac{g'}{g} \right| \right) \int_{\mathbb{T}} g(t) dt \leq 2q$$

so  $|g| \leq 2q + 1 \leq 3q$ , and by (6.11),  $|g^{(k)}| \leq Cq^{k+1}$  for some  $C = C(k)$ .

Finally, by induction on  $k$  we have

$$(\ln g)^{(k)} = \sum_{0 \leq i_1 + \dots + i_k \leq k} a_{i_1, \dots, i_k} \left( \frac{g'}{g} \right)^{i_1} \cdots \left( \frac{g^{(k)}}{g} \right)^{i_k}$$

where the  $a_i$ 's depend only on  $k$ . It follows by (6.11) that  $\|\ln g\|_k \leq Cq^{k^2}$  for some  $C = C(k)$ .  $\square$

We will now prove Lemma 1. Setting  $a_1 = b' - a$ , we fix  $g(\theta)d\theta$  be the density probability given by Lemma 7 with  $b$  defined in (6.10) so that

$$a_1 g + b g' = c_0 \tag{6.12}$$

for some constant  $c_0$ . We will distinguish the cases where  $c_0$  is small and large.

First we will prove that when  $|c_0|$  is not too small,  $g(\theta)d\theta$  is a good approximation of  $\nu$ :

**Lemma 8.** *There exists constants  $C$  and  $K$  depending only on  $k$  such that*

$$d_k^*(\nu, g(\theta)d\theta) \leq C \frac{q^K \varepsilon}{|c_0|^2}$$

*Proof.* By assumptions, for  $\varphi$  in  $C^\infty(\mathbb{T})$  we have

$$\int_{\mathbb{T}} (a\varphi' + b\varphi'') dv = O(\varepsilon \|\varphi\|_k). \tag{6.13}$$

Noticing that  $g(\theta)d\theta$  is null on the functions of the form  $a\varphi' + b\varphi''$ , if  $\psi$  in  $C^\infty(\mathbb{T}^2)$  satisfies  $\int_{\mathbb{T}} \psi(t)g(t)dt = 0$  then there exists  $\varphi$  such that  $\psi = a\varphi' + b\varphi''$ , and we can assume  $\int_{\mathbb{T}} \varphi(t)dt = 0$  up to adding a constant to  $\varphi$ . In order to use (6.13) and conclude that  $\int_{\mathbb{T}} \psi dv$  is small, we need to estimate  $\|\varphi\|_k$ . Using that  $(b' - a)g + b g' = c_0$  or equivalently  $ag = (bg)' - c_0$ , we write

$$\psi g = ag\varphi' + bg\varphi'' = ((bg)' - c_0)\varphi' + bg\varphi'' = (bg\varphi')' - c_0\varphi'.$$

If  $|bg\varphi'|$  is maximum at  $\theta_0$ ,  $(bg\varphi')'(\theta_0) = 0$ , hence

$$|c_0\varphi'(\theta_0)| = |\psi(\theta_0)g(\theta_0)| \leq Cq^K \|\psi\|_0,$$

and

$$|c_0(bg\varphi')(\theta_0)| = |c_0\varphi'(\theta_0)| b(\theta_0)g(\theta_0) \leq CMq^{2K} \|\psi\|_0$$



. Thus,

$$\|c_0 b g \varphi'\|_0 \leq C M q^{2K} \|\psi\|_0.$$

We have  $b \geq m$  by definition, and if  $g$  is minimal at point  $\theta_1$ , we have by (6.12)

$$|c_0| = |a_1(\theta_1)g(\theta_1)| \leq C M |g(\theta_1)|,$$

so that  $g \geq \frac{1}{C M} |c_0| \geq \frac{1}{C} \varepsilon_1^{\frac{1}{3}}$ . These lower bounds leads to

$$\|\varphi'\|_0 \leq \frac{q^{2K}}{c_0^2} \|\psi\|_0.$$

Next, derivating  $k$  times the equality  $a\varphi' + b\varphi'' = \psi$ , we have

$$\varphi^{(k)} = \frac{1}{b} \left( \psi^{(k-2)} + \sum_{j=0}^{k-1} \alpha_j \varphi^{(j)} \right)$$

where  $\alpha_j$  is a linear combination of derivatives of  $a^{(i)}$ ,  $b^{(i)}$  ( $0 \leq i \leq k$ ) and we deduce by induction on  $k$  that  $\|\varphi\|_k \leq C \frac{q^k}{m} \|\psi\|_k$  for some  $C = C(k)$ , and thus by (6.13),

$$\int_{\mathbb{T}} \psi d\nu = \int_{\mathbb{T}} (a\varphi' + b\varphi'') d\nu = O(\varepsilon \|\varphi\|_k) = O\left(\frac{q^K \varepsilon}{c_0^2} \|\psi\|_k\right),$$

hence  $d_k^*(d\nu(\theta), g(\theta)d\theta) \leq \frac{C q^K}{c_0^2} \varepsilon$ .

□

Now we will treat the case where  $c_0$  is small. The density  $g(\theta)d\theta$  is not necessarily a good approximation in this case, but we still can find an approximating density:

**Lemma 9.** *There exists a smooth density  $h(\theta)d\theta$  such that  $d_k^*(\nu, h(\theta)d\theta) \leq C M q^K (\varepsilon + |c_0|)$ ,  $\|h\|_k \leq C q^K$  and  $\|\ln h\|_k \leq C q^K$ .*

*Proof.* For  $\varphi$  in  $C^\infty(\mathbb{T})$ , applying (6.13) to  $\varphi_1(\theta) = \int_0^\theta (\varphi(t) - \hat{\varphi}(0)) dt$  gives

$$\int_{\mathbb{T}} (a\varphi + b\varphi') d\nu = \tilde{c}_0 \int_{\mathbb{T}} \varphi(\theta) d\theta + O(\varepsilon \|\varphi\|_k)$$

where  $\tilde{c}_0 = \int_{\mathbb{T}} a(\theta) d\nu(\theta)$ , and we also have

$$\int_{\mathbb{T}} (a\varphi + b\varphi') g(\theta) d\theta = \int_{\mathbb{T}} (ag - (bg)') \varphi d\theta = -c_0 \int_{\mathbb{T}} \varphi(\theta) d\theta.$$

Thus, using Lemma 7 we choose  $\varphi$  such that  $a\varphi' + b\varphi$  is constant with  $\int_{\mathbb{T}} \varphi(\theta) d\theta = 1$  and  $\|\varphi\|_k \leq C q^K$  and we deduce that  $\tilde{c}_0 = -c_0 + O(q^K \varepsilon)$ . In consequence, denoting  $\alpha = |c_0| + q^K \varepsilon$ , for every  $\varphi$  in  $C^\infty(\mathbb{T})$ ,

$$\int_{\mathbb{T}} (a\varphi + b\varphi') d\nu = O(\alpha \|\varphi\|_k). \quad (6.14)$$

We set  $A(\theta) = \int_0^\theta \frac{a(t)}{b(t)} dt$ .

First, let  $I = [\theta_0, \theta_1]$  an interval be an interval where  $A$  does not have local minima and let  $\psi$  in  $C^\infty(\mathbb{T})$  supported in  $] \theta_0 + q^{-1}, \theta_1 - q^{-1}[$ , and satisfying  $\int_I \psi(t) \frac{e^{A(t)}}{b(t)} dt = 0$ . Then the function  $\varphi(\theta) = e^{-A(\theta)} \int_0^\theta \frac{\psi(t)}{b(t)} e^{A(t)} dt$  is 1-periodic, supported in  $I$ , satisfies  $a\varphi + b\varphi' = \psi$ , and:

-If  $\theta \in [\theta_0, \theta_0 + q^{-1}]$ , then for  $t$  in  $[\theta_0, \theta]$ ,  $|A(t) - A(\theta)| \leq \frac{\|A'\|_0}{q} \leq 1$  and so

$$|\varphi(\theta)| \leq e \int_0^\theta \frac{|\psi(t)|}{b(t)} dt \leq \frac{3}{m} \|\psi\|_0.$$

-If  $\theta \in [\theta_1, \theta_1 + q^{-1}]$ , then using that  $\varphi(\theta) = -e^{-A(\theta)} \int_\theta^{\theta_1} \frac{\psi(t)}{b(t)} e^{A(t)} dt$ , we get in the same way  $|\varphi(\theta)| \leq \frac{3}{m} \|\psi\|_0$ .

-If  $\theta \in [\theta_0, \theta_1]$ , then either  $A$  is increasing on  $[\theta_0, \theta]$  and

$$|\varphi(\theta)| \leq e^{-A(\theta)} \left( e^{A(\theta_0)} |\varphi(\theta_0)| + \int_{\theta_0}^\theta \frac{|\psi(t)|}{b(t)} e^{A(t)} dt \right) \leq |\varphi(\theta_0)| + \int_{\mathbb{T}} \frac{|\psi(t)|}{b(t)} dt \leq \frac{2}{m} \|\psi\|_0,$$

or  $A$  is decreasing on  $[\theta, \theta_1]$  and in the same way,

$$|\varphi(\theta)| \leq e^{-A(\theta)} \left( e^{A(\theta_1)} |\varphi(\theta_1)| + \int_\theta^{\theta_1} \frac{|\psi(t)|}{b(t)} e^{A(t)} dt \right) \leq |\varphi(\theta_1)| + \int_{\mathbb{T}} \frac{|\psi(t)|}{b(t)} dt \leq \frac{2}{m} \|\psi\|_0.$$

Thus we conclude that  $\|\varphi\|_0 \leq \frac{3}{m} \|\psi\|_0$ . Then, derivating  $k$  times the equality  $a\varphi + b\varphi' = \psi$ , we have

$$\varphi^{(k)} = \frac{1}{b} \left( \psi^{(k)} + \sum_{j=0}^{k-1} \alpha_j \varphi^{(j)} \right)$$

where  $\alpha_j$  is a linear combination of derivatives of  $a^{(i)}$ ,  $b^{(i)}$  ( $0 \leq i \leq k$ ) and we deduce by induction on  $k$  that  $\|\varphi\|_k \leq C \frac{q^k}{m} \|\psi\|_k$  for some  $C = C(k)$ , and so by (6.14),

$$\int_{\mathbb{T}} \psi d\nu = O\left(\frac{q^k \alpha}{m} \|\psi\|_k\right).$$

Thus, we deduce that there exists  $\lambda \geq 0$  such that for every  $C^\infty$  function  $\psi$  supported in  $I$ ,

$$\int_{\mathbb{T}} \psi d\nu = \lambda \int_{\mathbb{T}} \psi h_0 dt + O\left(\frac{q^k \alpha}{m} \|\psi\|_k\right)$$

Now we divide the torus in intervals  $I_1, \dots, I_m$  such that  $A$  has not local minima on each  $I_n$ . We assume these intervals ordered so that  $I_n = ]\theta_n, \theta_{n+1}[$  for some sequence  $\theta_1, \dots, \theta_m, \theta_{m+1} = \theta_1 \bmod 1$ . We denote  $\tilde{I}_n = ]\theta_n - q^{-1}, \theta_{n+1} + q^{-1}[$  and  $J_n = ]\theta_n - q^{-1}, \theta_n + q^{-1}[$ . We have proved that for each  $n$  there exists  $\lambda_n > 0$

such that  $\nu$  can be approximated by  $h_n = \lambda_n \frac{e^A}{b}$  on  $\tilde{I}_n$  in the sense that for every function  $\varphi \in C^\infty$  on  $\mathbb{T}$  supported in  $\tilde{I}_n$ ,

$$\int_{\mathbb{T}} \varphi d\nu = \int_{\mathbb{T}} \varphi h_n d\nu + O\left(\frac{q^k \alpha}{m} \|\varphi\|_k\right)$$

Now we fix  $u$  a  $C^\infty$  function on  $[-1, 1]$  into  $[0, 1]$  such that  $u(-1) = 0$ ,  $u(1) = 1$  and  $u^{(j)}(-1) = u^{(j)}(1) = 0$  for every  $j \geq 1$ , and we define  $h$  by  $h = h_n$  on  $\tilde{I}_n$  and  $h(\theta) = h_n(\theta)^{1-u(q(\theta-\theta_n))} h_{n+1}(\theta)^{u(q(\theta-\theta_n))}$ .

Let us prove that  $h(\theta)d\theta$  is an approximation of  $\nu$ . We fix a  $C^\infty$  function  $v$  with compact support from  $]-1, 1[$  into  $[0, 1]$  such that  $v = 1$  on a neighborhood of 0 and we denote  $\chi(\theta) = \sum_n v(q(\theta - \theta_n))$ , and for  $\varphi$  in  $C^\infty(\mathbb{T})$  we write  $\varphi = (1 - \chi)\varphi + \chi\varphi$ . The function  $(1 - \chi)\varphi$  is null at the neighborhood of the  $\theta_n$ 's hence for every  $n$ ,

$$\int_{I_n} (1 - \chi)\varphi d\nu = \int_{I_n} (1 - \chi)\varphi h_n dt + O\left(\frac{q^k \alpha}{m} \|\varphi\|_k\right) = \int_{I_n} (1 - \chi)\varphi h dt + O\left(\frac{q^{2k} \alpha}{m} \|\varphi\|_k\right).$$

The function  $\chi\varphi$  is supported in  $\cup_n J_n$ . Since  $J_n \subset \tilde{I}_n \cap \tilde{I}_{n+1}$ , we have

$$\int_{J_n} \chi\varphi d\nu = \int_{J_n} \chi\varphi h_n dt + O\left(\frac{q^k \alpha}{m} \|\chi\varphi\|_k\right) = \int_{J_n} \chi\varphi h_{n+1} dt + O\left(\frac{q^k \alpha}{m} \|\chi\varphi\|_k\right),$$

and if for exemple  $\lambda_n \leq \lambda_{n+1}$  then  $h_n \leq h \leq h_{n+1}$  on  $J_n$ , hence

$$\int_{J_n} \chi\varphi h_n dt \leq \int_{J_n} \chi\varphi h dt \leq \int_{J_n} \chi\varphi h_{n+1} dt$$

and hence

$$\int_{J_n} \chi\varphi h dt = \int_{J_n} \chi\varphi d\nu + O\left(\frac{q^k \alpha}{m} \|\chi\varphi\|_k\right) = \int_{J_n} \chi\varphi d\nu + O\left(\frac{q^{2k} \alpha}{m} \|\varphi\|_k\right).$$

Thus,

$$\int_{\mathbb{T}} \varphi d\nu = \int_{\mathbb{T}} \varphi h dt = \int_{\mathbb{T}} \varphi h dt + O\left(\frac{q^{2k} \alpha}{m} \|\varphi\|_k\right).$$

Finally, up to replace  $h$  by  $h/\hat{h}(0)$ , we can assume that  $h(\theta)d\theta$  is a probability density. Next, since for each  $n$ ,  $\ln(h_n) = \ln(\lambda_n) + A - \ln b$ , we obtain that for  $1 \leq j \leq k$ ,

$$\|(\ln h_n)^{(j)}\|_0 \leq C \|A^{(j)}\|_0 + \|(\ln b)^{(j)}\|_0 \leq Cq$$

for some  $C$  and  $K$  depending of  $k$ , and hence that  $\|(\ln h)^{(j)}\|_0 \leq Cq^K$  for some  $C$  and  $K$  depending on  $k$  and  $d$  since  $\ln h$  is locally either equal to some  $\ln h_n$  or a smooth combination of  $\ln h_n$  and  $\ln h_{n+1}$ . Next, as in the proof of Lemma 7 we deduce from  $|\frac{h'}{h}| \leq Cq^K$  that  $\int_{\mathbb{T}} |h'| dt \leq Cq^K$  and hence that  $\|h\|_0 \leq Cq^K$  and next by successive derivations of the relation  $h' = (\ln h)'h$  that  $\|h\|_k \leq Cq^K$  for some  $C, K$ .  $\square$

Thus, using the two previous lemmas, we can find a density probability  $h(\theta)d\theta$  such that  $\|\ln h\|_k \leq Cq^K$ ,  $\|h\|_k \leq Cq^K$  and

$$d_k^*(\nu, h(\theta)d\theta) \leq Cq^K \inf\left(\frac{\varepsilon}{|c_0|^2}, \frac{|c_0| + \varepsilon}{M}\right)$$

If  $|c_0| \geq (M\varepsilon)^{\frac{1}{3}}$ , then

$$\frac{\varepsilon}{|c_0|^2} \leq \left(\frac{\varepsilon}{M}\right)^{\frac{1}{3}}$$

and if  $|c_0| \geq (M\varepsilon)^{\frac{1}{3}}$  then

$$\frac{|c_0| + \varepsilon}{M} \leq \left(\frac{\varepsilon}{M}\right)^{\frac{1}{3}} + \frac{\varepsilon}{M} \leq C\left(\frac{\varepsilon}{M}\right)^{\frac{1}{3}}$$

(using the obvious inequality  $\varepsilon \leq 2M$ ). The proposition 8 is proved □

At this point, we have proved that either there exist a direction close to be invariant by the matrices  $M_z$ , or that the stationary measure  $\hat{\mu}$  can be approximated by a controlled density. Now we will see that in the second case, the fact that  $f$  satisfies assumptions (I) or (II) imply that the cocycle  $U$  is close to be linearizable:

**Lemma 10.** *If  $f$  satisfies the assumption (I) or (II) of Theorem 1, we have for some  $C = C(k, \delta)$  and  $\tilde{\delta} = \tilde{\delta}(k, \delta)$  the alternative:*

- either there exists  $\theta_0$  in  $\mathbb{T}$  such that

$$\|\eta(\cdot, \theta_0)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{1+\delta},$$

- or there exists a diffeomorphism  $G$  in  $\text{Diff}_0(\mathbb{T})$  satisfying  $\|G - \text{Id}\|_k \leq C\varepsilon^{-\delta}$  and a family  $(r_z = \text{Id} + \omega(z))_{z \in \mathbb{T}^2}$  of random translations on  $\mathbb{T}$  such that:

$$\|G u \cdot G^{-1} - r_\cdot\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{\frac{7}{6} - K\delta}.$$

*Proof.* We can assume that the first alternative does not hold, so that we can apply Lemma 1 to approximate  $\nu$  by some smooth density  $g$ . We set  $G(\theta) = \int_0^\theta g(t)dt$ .

We first treat the case where  $\gamma_2(f, \mu) \geq 0$ . For any real numbers  $\theta_1$  and  $\theta_2$ , and any  $z$  in  $\mathbb{T}^2$ , the area of the parallelogram generated by the vectors  $M_z(e^{i\pi\theta_1})$  and  $M_z(e^{i\pi\theta_2})$  is equal to  $|\det(M_z)| \cdot |\sin(\pi(\theta_1 - \theta_2))|$ , and also to  $|M_z(e^{i\pi\theta_1})| \cdot |M_z(e^{i\pi\theta_2})| \cdot |\sin(u_z(\theta_1) - u_z(\theta_2))|$ . Letting  $\theta_2 - \theta_1$  tend to 0 and using (6.7) leads to the equality

$$|M_z(e^{i\pi\theta})|^2 \cdot |u'_z(\theta)| = |\det(M_z)| = |\det(d_z f)|,$$

and hence

$$2 \iint_{\mathbb{T}^2} \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln |M_z(e^{i\pi\theta})| d\hat{\mu}(z, \theta) + \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln u'_z(\theta) d\hat{\mu}(z, \theta) = O(\varepsilon_1^2).$$

Moreover,  $\gamma_1(f, \mu) \geq \gamma_2(f, \mu) \geq 0$  and

$$\gamma_1(f, \mu) + \gamma_2(f, \mu) = \mathbb{E} \int_{\mathbb{T}^2} \ln \det d_z f d\mu(z) = O(\varepsilon_1^2)$$

hence we have  $\gamma_i(f, \mu) = O(\varepsilon_1^2)$  for  $i = 1, 2$ . This implies that

$$\mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln |M_z(e^{i\pi\theta})| d\hat{\mu}(z, \theta) = O(\varepsilon_1^2),$$

and hence

$$\mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln u'_z(\theta) d\hat{\mu}(z, \theta) = O(\varepsilon_1^2).$$

Thus, the Lyapunov exponent of the cocycle  $U$  associated to the stationary measure  $\hat{\mu}$  is close to 0. The idea is now to use that since  $\hat{\mu}$  is close to  $g(\theta) dz d\theta$ , the conjugation of  $U$  by  $(z, \theta) \mapsto (z, G(\theta))$  should be almost  $dz d\theta$ -stationary, which would allow to easily estimate the associated Lyapunov exponent. But we will not exactly proceed like this because we do not have a good control on derivatives of  $G^{-1}$ . We will rather copy the proof of the invariance of Lyapunov exponent by conjugation, which leads to the following computation:

By  $\hat{\mu}$  invariance we have

$$\begin{aligned} \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln u'_z(\theta) d\hat{\mu}(z, \theta) &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} (\ln u'_z(\theta) + \ln g(u_z(\theta)) - \ln g(\theta)) d\hat{\mu}(z, \theta) \\ &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln \left( \frac{(G \circ u_z)'(\theta)}{G'(\theta)} \right) d\hat{\mu}(z, \theta) \\ &= \mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} \ln(1 + w) d\hat{\mu} \end{aligned}$$

where we have set

$$w(z, \theta) = \frac{(G \circ u_z)'(\theta) - G'(\theta)}{G'(\theta)}$$

We estimate  $w$  using the relation

$$\ln(1 + w) = \ln(1 + \eta') + \ln g \circ u - \ln g.$$

By Taylor formulas at order 1, we have

$$\|\ln(1 + \eta') - \eta'\|_{C^k(\mathbb{T}^2 \times \mathbb{T})} \leq C \|\eta'\|_{C^{k+1}(\mathbb{T}^2 \times \mathbb{T})}^2 \leq C\varepsilon_1^2$$

for some  $C = C(k)$ , and

$$\|\ln g \circ u - \ln g - (\ln g)' \eta\|_{C^k(\mathbb{T}^2 \times \mathbb{T})} \leq C \|\ln g\|_{k+1} \|\eta\|_{C^k(\mathbb{T}^2 \times \mathbb{T})}^2 \leq C\varepsilon_1^{2-K\delta}$$

for some  $C = C(k)$ .

Thus,

$$\|w - (\eta' + (\ln g)' \eta)\|_{C^k(\mathbb{T}^2 \times \mathbb{T})} \leq C \varepsilon_1^{2-K\delta}$$

and hence

$$\|\ln(1+w) - (\eta' + (\ln g)' \eta)\|_{C^k(\mathbb{T}^2 \times \mathbb{T})} \leq C \varepsilon_1^{2-K\delta}.$$

In particular,  $\|w\|_{C^k(\mathbb{T}^2 \times \mathbb{T})} \leq C \varepsilon_1^{1-K\delta}$ , and  $\bar{w}(\theta) = \int_{\mathbb{T}^2} w(z, \theta) dz$  satisfies  $\|\bar{w}\|_k \leq C \varepsilon_1^{2-K\delta}$ .

Now, using that  $d_k^*(\hat{\mu}, dz \times \nu) \leq C \varepsilon_1^2$  and  $d_k(\nu, d(\theta)d\theta)$ , we obtain

$$\begin{aligned} \int_{\mathbb{T}^2 \times \mathbb{T}} \ln(1+w) d\hat{\mu} &= \int_{\mathbb{T}^2 \times \mathbb{T}} \ln(1+w) dz d\nu(\theta) + O(\varepsilon_1^{3-K\delta}) \\ &= \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} \left( w - \frac{w^2}{2} \right) dz d\nu(\theta) + O(\varepsilon_1^{3-K\delta}) \\ &= \mathbb{E} \int_{\mathbb{T}} \bar{w}(\theta) d\nu(\theta) - \frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} w(z, \theta)^2 dz d\nu(\theta) + O(\varepsilon_1^{3-K\delta}) \\ &= \mathbb{E} \int_{\mathbb{T}} \bar{w}(\theta) g(\theta) dz d\theta - \frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} w(z, \theta)^2 g(\theta) dz d\theta + O(\varepsilon_1^{\frac{7}{3}-2K\delta}) \\ &= \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} \left( \frac{(G \circ u_z - G)'(\theta)}{g(\theta)} \right) g(\theta) dz d\theta \\ &\quad - \frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} \left( \frac{(G \circ u_z - G)'(\theta)}{g(\theta)} \right)^2 g(\theta) dz d\theta + O(\varepsilon_1^{\frac{7}{3}-2K\delta}) \\ &= -\frac{1}{2} \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} (G \circ u_z - G)'(\theta)^2 \frac{dz d\theta}{g(\theta)} + O(\varepsilon_1^{\frac{7}{3}-2K\delta}) \end{aligned}$$

We deduce that

$$\mathbb{E} \iint_{\mathbb{T}^2 \times \mathbb{T}} (G \circ u_z - G)'(\theta)^2 d\theta dz \leq \|g\|_0 \mathbb{E} \int_{\mathbb{T}^2 \times \mathbb{T}} (G \circ u_z - G)'(\theta)^2 \frac{dz d\theta}{g(\theta)} \leq C \varepsilon_1^{\frac{7}{3}-3K\delta}$$

for some  $C = C(k)$ . Thus, setting  $\omega(z) = G(u_z(0))$ ,

$$\mathbb{E} \int_{\mathbb{T}^2} \|G \circ u_z - G - \omega(z)\|_0^2 dz \leq C \varepsilon_1^{\frac{7}{3}-3K\delta}.$$

If  $f$  satisfies assertion (II), hence belongs almost surely to some determinist solvable group  $G$ , then the cocycle  $(z, h) \mapsto (f(z), d_z f(h))$  almost surely belongs to the solvable group  $\hat{G} = \{(z, h) \mapsto (g(z), d_z g(h)), g \in G \cap \text{Diff}^1(\mathbb{T}^2)\}$ , and then that this is also the case for  $U$ . Hence, by Proposition 1, there exists a probability  $\hat{\mu}$  almost surely invariant by  $U$ , which can be approximated by a density  $g(\theta)d\theta dz$ . Thus, for  $\varphi$  in  $C^\infty(\mathbb{T}^2 \times \mathbb{T})$ , the integral of  $\varphi \circ U^{-1} - \varphi$  along  $\hat{\mu}$  is null, and differs from its integral along  $g(\theta)d\theta dz$  by a remainder

$$O\left(\varepsilon_1^{\frac{7}{3}-K\delta} (\|\varphi \circ U^{-1}\|_k + \|\varphi\|_k)\right) = O(\varepsilon_1^{\frac{7}{3}-K\delta} \|\varphi\|_k).$$

By a change of variable,

$$\int_{\mathbb{T} \times \mathbb{T}^2} \varphi \circ U^{-1}(z, \theta) g(\theta) d\theta dz = \int_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) g(u_z(\theta)) u'_z(\theta) \det(d_z f) dz d\theta$$

and by Proposition 5,  $\det d_z f = 1 + O(\varepsilon^{\frac{3}{2}})$ , hence

$$\int_{\mathbb{T}^2 \times \mathbb{T}} \varphi(z, \theta) (G \circ u_z - G)' dz d\theta = O(\varepsilon_1^{\frac{7}{3} - K\delta} \|\varphi\|_k)$$

Choosing  $\varphi(z, \theta) = (G \circ u_z - G)'(\theta)$ , we deduce as in the first case that

$$\int_{\mathbb{T}^2} \|G \circ u_z - G - \omega(z)\|_0^2 dz = O(\varepsilon_1^{\frac{7}{3} - \delta}).$$

□

**Lemma 11.** *There exists  $\delta_0 > 0$  such that for every  $\kappa > 0$ , if  $f$  satisfies assumption (I) or (II) of Theorem 1, there exists a vector  $p$  in  $\mathbb{Z}^2$ , a map  $k : \mathbb{T}^2 \rightarrow \mathbb{R}$  and some positive constants  $C$  and  $\delta_0$  such that*

$$\|d_z f(p) - k(z)p\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon^{1+\delta_0}$$

with  $|p| \leq \varepsilon^{-\kappa}$ , and  $p = (0, 1)$  if  $m = \|\beta'\|_0 \geq 2\varepsilon$ .

*Proof.* Now, we will fix the value  $\lambda$  in the definition of  $M_z$  and  $\varepsilon_1$  of the beginning of the section by:

$$\begin{cases} \lambda = \sqrt{\frac{\varepsilon}{m}} \text{ if } m > 2\varepsilon, \text{ in which case } \varepsilon_1 = 2\sqrt{m\varepsilon}. \\ \lambda = 1 \text{ if } m \leq 2\varepsilon, \text{ in which case } \varepsilon_1 = \varepsilon + m \leq 3\varepsilon. \end{cases}$$

Let  $\delta > 0$ . We will first prove that if  $\delta$  is small enough, this is necessarily the first alternative of Lemma 10 which occurs. Thus, let us assume that there exists  $G$  and  $\omega$  such that for every  $\theta$  in  $\mathbb{T}$ ,

$$\|G \circ u_z(\theta) - G(\theta) - \omega(\cdot)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{\frac{7}{6} - K\delta}.$$

By Taylor expansion, denoting  $g = G'$ ,

$$|G \circ u_z(\theta) - G(\theta) - g(\theta)\eta_z(\theta)| \leq C\|g\|_1 \|\eta_z\|_0 \leq C\varepsilon_1^{2-K\delta},$$

hence for every  $\theta$  in  $\mathbb{T}$ ,

$$\|g(\theta)\eta_z(\theta) - \omega(\cdot)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{\frac{7}{6} - K\delta}. \quad (6.15)$$

We first treat the case where  $\inf g \leq \varepsilon_1^\delta$ : by (6.15), for every  $\theta_0$  and  $\theta_1$  in  $\mathbb{T}$  we have

$$\|g(\theta_0)\eta_z(\theta_0) - g(\theta_1)\eta_z(\theta_1)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{\frac{7}{6} - K\delta}$$

, and hence

$$g(\theta_0)\eta_z(\theta_0) \leq g(\theta_1)\varepsilon_1^2 + \varepsilon_1^{\frac{7}{6}-K\delta}$$

Choosing  $\theta_0$  such that

$$g(\theta_0) = \sup_{\mathbb{T}} g \geq \int_{\mathbb{T}} g(t) dt = 1$$

and  $\theta_1$  such that

$$g(\theta_1) = \inf_{\mathbb{T}} g \geq \varepsilon_1^\delta$$

we obtain that

$$\|\eta \cdot (\theta_0)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{1+\delta}.$$

Thus we can assume that  $\inf g \geq \varepsilon_1^\delta$ . If  $W$  is a complex measure on  $\mathbb{T}^2$ , we set  $N_W = \int_{\mathbb{T}^2} N_z dW(z)$  and  $\eta_W = \int_{\mathbb{T}^2} \eta_z dW(z)$ . The interest of these notations is the following: If  $W$  is a complex measure on  $\mathbb{T}^2$  such that  $|W|(\mathbb{T}^2) \leq 1$ , then by (6.15) we have for every  $\theta$  and  $\theta'$  in  $\mathbb{T}$

$$|g(\theta_0)\eta_W(\theta_0) - g(\theta_1)\eta_W(\theta_1)| \leq C\varepsilon_1^{\frac{7}{6}-K\delta}$$

hence

$$|\eta_W(\theta)| \leq \frac{1}{g(\theta)} \left( g(\theta')|\eta_W(\theta')| + C\varepsilon_1^{\frac{7}{6}-K\delta} \right) \leq \varepsilon_1^{-(K+1)\delta} |\eta_W(\theta)| + \varepsilon_1^{\frac{7}{6}-(K+1)\delta},$$

and in particular, if there exists  $\theta_0$  in  $\mathbb{T}$  such that  $N_W(e^{i\pi\theta_0}) = 0$ , what implies  $\eta_W(\theta_0) = 2\text{Im}(N_W(e^{i\pi\theta_0})e^{-i\pi\theta_0}) + O(\varepsilon_1^2) = O(\varepsilon_1^2)$ , then  $\|\eta_W\|_0 = O(\varepsilon_1^{\frac{7}{6}-2K\delta})$  and hence  $N_W = O(\varepsilon_1^{\frac{7}{6}-2K\delta})$ .

Using the previous remark:

-If  $m > 2\varepsilon$ , and hence

$$N_z = \begin{pmatrix} \partial_1 \zeta_1(z) & \frac{\partial_2 \bar{\zeta}_1(z)}{\lambda} \\ \lambda(\partial_1 \zeta_2(z) + \beta'(x)) & \partial_2 \zeta_2(z) \end{pmatrix},$$

we choose  $dW = \delta_x \times dy$ , so that  $N_W = \begin{pmatrix} \times & 0 \\ \times & 0 \end{pmatrix}$  vanishes at the vector  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . By the above remark,  $N_W = O(\varepsilon_1^{\frac{7}{6}-K\delta})$ , which implies, setting  $V(x) = \beta'(x) + \int_{\mathbb{T}} \partial_1 \zeta_2(x, y) dy$ , that  $\lambda V(x) = O(\varepsilon_1^{\frac{7}{6}-\delta})$ , and using that  $\|V\|_0 \geq m - \varepsilon \geq \frac{m}{2}$  we get  $\varepsilon_1 = \lambda m = O(\varepsilon_1^{\frac{7}{6}-\delta})$ , which is absurd if  $\varepsilon_1$  and  $\delta$  are small enough.



If  $m \leq 2\varepsilon$ , then

$$N_z = J\Phi(z) = \begin{pmatrix} \partial_1 \zeta_1(z) & \partial_2 \zeta_1(z) \\ \partial_1 \zeta_2(z) + \beta'(x) & \partial_2 \zeta_2(z) \end{pmatrix}$$

where  $\Phi = f - Id$ . For  $p = (p_1, p_2)$  in  $\mathbb{Z}^2$ , we choose  $dW = e^{2i\pi p \cdot z} dz$ , so that  $N_W = (p_j \hat{\Phi}_i(p))_{1 \leq i, j \leq 2}$  is null at vector  $\begin{pmatrix} p_2 \\ -p_1 \end{pmatrix}$ , hence  $N_W = O(\varepsilon_1^{\frac{7}{6}-\delta})$ . Since this estimate is valid for any choice of  $p$ , this implies that for  $i, j = 1, 2$  and  $k_1$  large enough

$$\begin{aligned} \|\partial_i \Phi_j\|_0 &\leq \sum_{p \in \mathbb{Z}^2} |p_j \hat{\Phi}_i(p)| \\ &\leq C \sum_{|p| \leq \varepsilon_1^{-\delta}} \varepsilon_1^{\frac{7}{6}-2K\delta} + \sum_{|p| \geq \varepsilon_1^{-\delta}} \frac{\|\Phi\|_{k_1}}{|p|^{k_1}} \\ &\leq C \varepsilon_1^{\frac{7}{6}-(2K+1)\delta}. \end{aligned}$$

We deduce that  $N_z = O(\varepsilon_1^{\frac{7}{6}-(2K+1)\delta})$  and so that  $\eta_z = O(\varepsilon_1^{\frac{7}{6}-(2K+1)\delta})$ .

Thus, if  $\delta_0$  is small enough, there necessarily exists  $\theta_0$  such that  $\|\eta_1(\theta_0)\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{1+\delta_0}$ . This implies that

$$\|N_z e^{i\pi\theta_0} - k(\cdot) e^{i\pi\theta_0}\|_{L^2(\mathbb{T}^2)} \leq C\varepsilon_1^{1+\delta_0}$$

where  $k(z) = \operatorname{Re}(N_z e^{i\pi\theta_0} e^{-i\pi\theta_0})$ .

If  $m > 2\varepsilon$ , for  $W = \delta_x \times dy$ ,  $N_W$  and  $V(x)$  defined as previously and  $k_W = \int_{\mathbb{T}^2} k(z) dW(z)$ , we have  $N_W e^{i\pi\theta_0} = k_W e^{i\pi\theta_0} + O(\varepsilon_1^{1+\delta_0})$  with  $N_W = \begin{pmatrix} 0 & 0 \\ \lambda V(x) & 0 \end{pmatrix} + O(\varepsilon_1^2)$ , hence

$$\begin{cases} \lambda V(x) \cos(\pi\theta_0) = k_W \sin(\pi\theta_0) + O(\varepsilon_1^{1+\delta_0}) \\ k_W \cos(\pi\theta_0) = O(\varepsilon_1^{1+\delta_0}) \end{cases}.$$

Combining these two equalities gives  $\lambda V(x) \cos^2(\pi\theta_0) = O(\varepsilon_1^{1+\delta_0})$  with  $\|V\|_0 \geq \frac{m}{2}$ , hence  $\cos^2(\pi\theta_0) = O(\varepsilon_1^{1+\delta_0})$ . Thus,  $e^{i\pi\theta_0} = \pm i + O(\varepsilon_1^{\delta_0})$ , hence, denoting  $e_2$  the second canonical vector of  $\mathbb{R}^2$ , we successively obtain  $N_z e_2 = k(z) e_2 + O(\varepsilon_1^{1+\delta_0})$ ,  $M_z e_2 = (1 + k(z)) e_2 + O(\varepsilon_1^{1+\delta_0})$  and  $d_z f(e_2) = (1 + k(z)) e_2 + O(\lambda \varepsilon_1^{1+\delta_0})$ , with  $\lambda \varepsilon_1^{1+\delta_0} = m^{\frac{\delta_0}{2}} \varepsilon_1^{1+\frac{\delta_0}{2}} = O(\varepsilon_1^{1+\frac{\delta_0}{2}})$ .

If  $m \leq 2\varepsilon$ , we set  $dW = e^{2i\pi p \cdot z} dz$ , where  $p = (p_1, p_2) \in \mathbb{Z}^2$ . On one hand, we have  $N_W e^{i\pi\theta_0} = k_W e^{i\pi\theta_0} + O(\varepsilon_1^{1+\delta_0})$ , and on another hand,  $N_W$  vanishes at vector  $\begin{pmatrix} p_2 \\ -p_1 \end{pmatrix}$ , hence writing  $N_W$  in the basis  $\begin{pmatrix} p_2 \\ -p_1 \end{pmatrix}, \begin{pmatrix} \cos(\pi\theta_0) \\ \sin(\pi\theta_0) \end{pmatrix}$  gives the relation

$$\operatorname{Tr}(N_W) = k_W + O(\varepsilon_1^{1+\delta_0}).$$

But  $\text{Tr}(N_z) = \det(M_z) - 1 + O(\varepsilon^2) = O(\varepsilon^{1+\delta_0})$  so  $k_W = O(\varepsilon^{1+\delta_0})$  and so  $N_W(e^{i\pi\theta_0}) = O(\varepsilon^{1+\delta_0})$ , that is

$$(p_1 \cos(\pi\theta_0) + p_2 \sin(\pi\theta_0))\hat{\zeta}_i(p_1, p_2) = O(\varepsilon^{1+\delta_0})$$

for  $i = 1, 2$ . Now we distinguish two cases:

1) If  $|\hat{\zeta}_i(p)| \leq \varepsilon^{1+\frac{\delta_0}{2}}$  ( $i = 1, 2$ ) for every  $p \in \mathbb{Z}^2 - \{(0, 0)\}$  such that  $|p| \leq \varepsilon^{-\kappa}$ , then, for  $k_1$  large enough,

$$\|\zeta_i - \hat{\zeta}_i(0)\|_0 \leq \sum_{p \in \mathbb{Z}^2 - \{(0,0)\}} |\hat{\zeta}_i(p)| \leq \sum_{|p| \leq \varepsilon^{-\kappa}} \varepsilon^{1+\frac{\delta_0}{2}} + \sum_{|p| \geq \varepsilon^{-\kappa}} \frac{\|\zeta_i\|_{k_1}}{|p|^{k_1}} = O(\varepsilon^{1+\frac{\delta_0}{3}}).$$

If  $|\hat{\zeta}_i(p)| \geq \varepsilon^{1+\frac{\delta_0}{2}}$  for some  $i$  in  $\{1, 2\}$  and some  $p = (p_1, p_2)$  in  $\mathbb{Z}^2 - \{(0, 0)\}$  such that  $|p| \leq \varepsilon^{-\kappa}$ , then  $|p_1 \cos(\pi\theta_0) + p_2 \sin(\pi\theta_0)| \leq \varepsilon^{\frac{\delta_0}{2}}$ . Thus,  $e^{i\pi\theta_0} = \pm \frac{\tilde{p}}{|p|} + O(\varepsilon^{\frac{\delta_0}{2}})$ , where  $\tilde{p} = (-p_2, p_1)$ . Consequently,  $N_z(\tilde{p}) = O(\varepsilon^{1+\frac{\delta_0}{3}})$ .  $\square$

We can complete the proof of Proposition 7. Let  $p = (p_1, p_2)$  and  $k(z)$  given by Lemma 11. Up to dividing  $p_1$  and  $p_2$  by some common divisor, we can assume  $\gcd(p_1, p_2) = 1$ . According to Bezout Lemma, there exists  $P$  in  $Sl_2(\mathbb{Z})$  such that  $Pe_2 = \tilde{p}$  and  $\|P\| \leq |p| \leq \varepsilon^{-\delta_1}$ , and we can choose  $P = Id$  if  $p = (0, 1)$ , hence if  $m \geq 2\varepsilon$ . The diffeomorphism  $\tilde{f} = P^{-1}fP = (\tilde{f}_1, \tilde{f}_2)$  satisfies

$$d_z \tilde{f}(e_2) = e_2 + P^{-1}N_z P e_2 = e_2 + O(\|P^{-1}\| \varepsilon^{1+\delta_0}) = e_2 + O(\varepsilon^{1+\frac{\delta_0}{2}}),$$

that is

$$\partial_2 \tilde{f}_1 = O(\varepsilon^{1+\frac{\delta_0}{2}}). \quad (6.16)$$

Thus,  $\tilde{f}_1$  depends only on  $x$ , up to a remainder  $O(\varepsilon^{1+\delta_0})$ . Next, we have by (6.7)

$$\det(d_z \tilde{f}) = \det(d_z f) = 1 + O(\varepsilon^{\frac{3}{2}}),$$

which leads to

$$\partial_1 \tilde{f}_1 + \partial_2 \tilde{f}_2 = 2 + O(\varepsilon^{\frac{3}{2}}).$$

Using that  $\int_{\mathbb{T}} \partial_2 \tilde{f}_2(x, y) dy = 1$  and that  $\int_{\mathbb{T}} \partial_1 \tilde{f}_1(x, y) dy = \partial_1 \tilde{f}_1(x, \cdot) + O(\varepsilon^{1+\frac{\delta_0}{2}})$ , we obtain

$$\begin{cases} \partial_1 \tilde{f}_1 = 1 + O(\varepsilon^{1+\frac{\delta_0}{2}}) \\ \partial_2 \tilde{f}_2 = 1 + O(\varepsilon^{1+\frac{\delta_0}{2}}). \end{cases} \quad (6.17)$$

Estimates (6.16) and (6.17) imply that  $\tilde{f} = \tilde{R} + O(\varepsilon^{1+\frac{\delta_0}{2}})$  for some Anzai cocycle  $\tilde{R}$ .  $\square$

## 6.5 KAM iteration

Let  $f = R + \zeta$  a random diffeomorphism of  $\mathbb{T}^2$  having a finite number of realizations, where  $R : (x, y) \mapsto (x + \alpha, y + \beta(x))$  is a diophantine random Anzai cocycle satisfying  $\mathbb{E}[\beta(x - \alpha)] = 0$ . We will denote  $\varepsilon_{k_1, k_2} = \|\zeta\|_{k_1, k_2}$  and  $\Delta = \|\|R \circ \tilde{R} - \tilde{R} \circ R\|\|$ . As a corollary of the two previous sections (more precisely of Proposition 5 and 7), we have the following conjugation result:

**Lemma 12.** *There exists  $a$  and  $k_0$  in  $\mathbb{N}$  and  $\delta_0, \varepsilon > 0$  such that if  $f$  satisfies assumption (I) or (II) of Theorem 1, and if  $\varepsilon_{k_0, k_0} \leq \varepsilon$  then:*

i) *There exists a diffeomorphism  $G = P + g$  where  $P$  is the algebraic part of  $G$ , and an Anzai cocycle  $R_1$ , such that*

$$\|\|GfG^{-1} - R_1\|\|_0 \leq C\varepsilon_{k_0, k_0}^{1+\delta_0}$$

and for every  $k_1, k_2$

$$\|g\|_{k_1, k_2} \leq \frac{C}{\Delta^{16k_1}} \varepsilon_{k_1+a, k_2+ak_1+a}.$$

for some  $C = C(k_1, k_2)$ .

ii) *For every integer  $p$ , there exists  $G = P + g$  where  $P$  is an algebraic automorphism, and an Anzai cocycle  $R_1$ , such that*

$$\|\|GfG^{-1} - R_1\|\|_0 \leq C(\varepsilon_{k_0, k_0} + \sqrt{\Delta^p \varepsilon_{k_0, k_0}})^{1+\delta_0}$$

and for every  $\delta > 0$  and every integers  $k_1, k_2$  large enough,

$$\|g\|_{k_1, k_2} \leq C\varepsilon^{(1+\delta)k_1, (1+\delta)k_2}.$$

Moreover,  $\|P\| \leq C\varepsilon_{k_0, k_0}^{-\kappa}$  and  $P = Id$  if  $\|\|\beta'\|\|_0 \geq 2\|\zeta\|_{k_0}$ .

In order to iterate this conjugation procedure, we will need some estimates on large derivatives of the conjugated diffeomorphism. In this view, we will replace the conjugation  $G$  by some approximation by a trigonometric polynomial.

If  $\varphi$  belongs to  $C^\infty(\mathbb{T}^2)$  and  $\lambda_1, \lambda_2$  are real numbers, we denote

$$P_{\lambda_1, \lambda_2} \varphi(x, y) = \sum_{|p_1| \leq \lambda_1, |p_2| \leq \lambda_2} \hat{\varphi}(p) e^{2i\pi(p_1 x + p_2 y)}.$$

The operator  $P_{\lambda_1, \lambda_2}$  satisfies the estimates:

**Lemma 13.** *For every  $\delta > 0$  and every integers  $k_1, k_2$ , and  $\bar{k}_1, \bar{k}_2$  large enough, there exists a constant  $C$  such that for every  $\varphi$  in  $C^\infty(\mathbb{T}^2)$ :*

$$\left\{ \begin{array}{l} \|P_{\lambda_1, \lambda_2} \varphi\|_{\bar{k}_1, \bar{k}_2} \leq C(\lambda_1^{\bar{k}_1 - k_1} + \lambda_2^{\bar{k}_2 - k_2}) \|\varphi\|_{k_1, k_2} \\ \|\varphi - P_{\lambda_1, \lambda_2} \varphi\|_{k_1, k_2} \leq C \left( \frac{1}{\lambda_1^{(k_1 - k_1) - \delta \bar{k}_1}} + \frac{1}{\lambda_2^{(k_2 - k_2) - \delta \bar{k}_2}} \right) \|\varphi\|_{\bar{k}_1, \bar{k}_2} \end{array} \right.$$

*Proof.* If  $\frac{i_1}{k_1} + \frac{i_2}{k_2} \leq 1$ , then

$$\begin{aligned} \|\partial_1^{i_1} \partial_2^{i_2} P_{\lambda_1, \lambda_2} \varphi\|_{L^2(\mathbb{T}^2)}^2 &\leq C \sum_{|p_1| \leq \lambda_1, |p_2| \leq \lambda_2} |p_1|^{2i_1} |p_2|^{2i_2} |\hat{\varphi}(p_1, p_2)|^2 \\ &\leq C \sum_{|p_1| \leq \lambda_1, |p_2| \leq \lambda_2} (|p_1|^{2\bar{k}_1} + |p_2|^{2\bar{k}_2}) |\hat{\varphi}(p_1, p_2)|^2 \\ &\leq C (\lambda_1^{2(\bar{k}_1 - k_1)} \|\partial_1^{k_1} \varphi\|_{L^2(\mathbb{T}^2)}^2 + \lambda_2^{2(\bar{k}_2 - k_2)} \|\partial_2^{k_2} \varphi\|_{L^2(\mathbb{T}^2)}^2) \\ &\leq C (\lambda_1^{2(\bar{k}_1 - k_1)} + \lambda_2^{2(\bar{k}_2 - k_2)}) \|\varphi\|_{k_1, k_2}^2 \end{aligned}$$

and if  $\frac{i_1}{k_1} + \frac{i_2}{k_2} \leq 1$ ,

$$\begin{aligned} \|\partial_1^{i_1} \partial_2^{i_2} (\varphi - P_{\lambda_1, \lambda_2} \varphi)\|_{L^2(\mathbb{T}^2)} &\leq C \left( \sum_{|p_1| > \lambda_1, p_2 \in \mathbb{Z}} + \sum_{p_1 \in \mathbb{Z}, |p_2| > \lambda_2} \right) |p_1|^{i_1} |p_2|^{i_2} |\hat{\varphi}(p_1, p_2)| \\ &\leq C \left( \sum_{|p_1| > \lambda_1, p_2 \in \mathbb{Z}} + \sum_{p_1 \in \mathbb{Z}, |p_2| > \lambda_2} \right) \left( \frac{|p_1|^{k_1} + |p_2|^{k_2}}{|p_1|^{\bar{k}_1} + |p_2|^{\bar{k}_2}} \right) \|\varphi\|_{\bar{k}_1, \bar{k}_2} \\ &\leq C \left( \sum_{|p_1| > \lambda_1, p_2 \in \mathbb{Z}} \frac{|p_1|^{k_1}}{|p_1|^{(1-\delta)\bar{k}_1} |p_2|^{\delta\bar{k}_2}} + \sum_{p_1 \in \mathbb{Z}, |p_2| > \lambda_1} \frac{|p_2|^{k_2}}{|p_1|^{\delta\bar{k}_1} |p_2|^{(1-\delta)\bar{k}_2}} \right) \|\varphi\|_{\bar{k}_1, \bar{k}_2} \\ &\leq C \left( \frac{1}{\lambda_1^{(\bar{k}_1 - k_1) - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{(\bar{k}_2 - k_2) - \delta\bar{k}_2}} \right) \|\varphi\|_{\bar{k}_1, \bar{k}_2} \end{aligned}$$

□

**Lemma 14.** Let  $\delta$  a positive real number. We set  $\varepsilon_k = \varepsilon_{\delta k, k}$ .

There exists  $a$  and  $k_0$  in  $\mathbb{N}$  and  $\delta_0, \varepsilon > 0$  such that if  $f$  satisfies assumption (I) or (II) of Theorem 1, and if  $\varepsilon_{k_0} \leq \varepsilon$  then:

i) For every  $\lambda > 1$ , there exist a diffeomorphism  $G_0 = P + g_0$ , an Anzai cocycle  $R_1 : (x, y) \mapsto (x + \alpha_1, y + \beta_1(x))$  and a family of constants  $(C_k)_{k \in \mathbb{N}}$  depending on  $\delta$  and  $\lambda$  such that, denoting  $f_1 = G_0 f G_0^{-1} = R_1 + \zeta_1$  and  $\varepsilon_k^{(1)} = \|\|\zeta_1\|\|_{\delta k, k}$ , we have for every integer  $k, \bar{k}$ :

$$\varepsilon_k^{(1)} \leq \frac{C_{\bar{k}}}{\Delta^{16\bar{k}}} \left( \lambda^{k+a} \varepsilon_{k_0}^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}}{\lambda^{(\bar{k}-k) - \delta\bar{k}-a}} \right) ..$$

ii) For every integer  $p$  and every  $\lambda > 1$ , there exists a diffeomorphism  $G_0 = P + g_0$ , an Anzai cocycle  $R_1 : (x, y) \mapsto (x + \alpha_1, y + \beta_1(x))$  and a family of constants  $(C_k)_{k \in \mathbb{N}}$  depending on  $\delta, \lambda$  and  $p$  such that, denoting  $f_1 = G_0 f G_0^{-1} = R_1 + \zeta_1$  and  $\varepsilon_k^{(1)} = \|\|\zeta_1\|\|_{\delta k, k}$ , we have

$$\varepsilon_k^{(1)} \leq C_k \left( \lambda^{k+a} (\varepsilon_{k_0} + \sqrt{\Delta^p \varepsilon_{k_0}})^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}}{\lambda^{(\bar{k}-k) - \delta\bar{k}-a}} \right)$$

Moreover in the two cases, we have  $\mathbb{E}[\beta'_1(x - \alpha_1)] = 0$ ,  $\|P\| \leq C \varepsilon_{k_0}^{-\kappa}$  and  $P = Id$  if  $\|\|\beta'\|\|_0 \geq 2\varepsilon_{k_0}$  or  $\|\|\tilde{\beta}'\|\|_0 \leq 2\varepsilon_{k_0}^{(1)}$ .

*Proof.* Let  $G = P + g$  and  $R_1$  given by point i) of Lemma 12,  $G_0 = P + g_{\lambda_1, \lambda_2}$  and  $f_\lambda = G_0 f G_0^{-1} = R_1 + \zeta_1$  where  $g_{\lambda_1, \lambda_2} = P_{\lambda_1, \lambda_2} g$ .

We know that

$$\|g\|_{k_1-a, k_2-ak_1} \leq \frac{C}{\Delta^{16k_1}} \varepsilon_{k_1, k_2},$$

hence

$$\|g_{\lambda_1, \lambda_2}\|_{k_1, k_2} \leq \frac{C}{\Delta^{16k_1}} (\lambda_1^a + \lambda_2^{ak_1}) \varepsilon_{k_1, k_2}$$

and

$$\|g - g_{\lambda_1, \lambda_2}\|_{k_1, k_2} \leq \frac{C}{\Delta^{16\bar{k}_1}} \left( \frac{1}{\lambda_1^{(\bar{k}_1 - k_1) - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{(\bar{k}_2 - k_2) - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2}.$$

Next, we have for some constant  $C$  the following estimates:

$$\begin{aligned} \|GfG^{-1} - R_1\|_0 &\leq C\varepsilon_{k_0, k_0}^{1+\delta_0} \\ \|G_0fG_0^{-1} - GfG^{-1}\|_0 &\leq \|G_0fG_0^{-1} - GfG_0^{-1}\|_0 + \|GfG_0^{-1} - GfG^{-1}\|_0 \\ &\leq \|G_0 - G\|_0 + \|Gf - GfG^{-1}G_0\|_0 \\ &\leq (1 + \|(GfG^{-1})'\|_0) \|G - G_0\|_0 \\ &\leq C \left( \frac{1}{\lambda_1^{\bar{k}_1 - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{\bar{k}_2 - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2} \end{aligned}$$

Combining the two last inequalities, we obtain:

$$\varepsilon_{0,0}^{(1)} \leq \|G_0fG_0^{-1} - R_1\|_0 \leq C \left[ \varepsilon_{k_0, k_0}^{1+\delta_0} + \left( \frac{1}{\lambda_1^{\bar{k}_1 - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{\bar{k}_2 - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2} \right].$$

Moreover, we also have

$$\varepsilon_{\bar{k}_1, \bar{k}_2}^{(1)} \leq C(\varepsilon_{\bar{k}_1, \bar{k}_2} + \|g\|_{\bar{k}_1, \bar{k}_2}) \leq \frac{C}{\Delta^{16\bar{k}_1}} (\lambda_1^a + \lambda_2^{a\bar{k}_1}) \varepsilon_{\bar{k}_1, \bar{k}_2}$$

and consequently, using the decomposition  $\zeta = P_{\lambda_1, \lambda_2} \zeta + (\zeta - P_{\lambda_1, \lambda_2} \zeta)$ , we deduce

$$\begin{aligned} \varepsilon_{k_1, k_2}^{(1)} &\leq C \left[ (\lambda_1^{k_1+a} + \lambda_2^{k_2+a}) \varepsilon_0^{(1)} + \left( \frac{1}{\lambda_1^{(\bar{k}_1 - k_1) - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{(\bar{k}_2 - k_2) - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2}^{(1)} \right] \\ &\leq \frac{C}{\Delta^{16\bar{k}_1}} \left[ (\lambda_1^{k_1+a} + \lambda_2^{k_2+a}) \varepsilon_{k_0, k_0}^{1+\delta_0} + (\lambda_1^{k_1+a} + \lambda_2^{k_2+a}) \left( \frac{1}{\lambda_1^{\bar{k}_1 - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{\bar{k}_2 - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2} \right] \\ &\quad + (\lambda_1^a + \lambda_2^{ak_1}) \left( \frac{1}{\lambda_1^{(\bar{k}_1 - k_1) - \delta\bar{k}_1}} + \frac{1}{\lambda_2^{(\bar{k}_2 - k_2) - \delta\bar{k}_2}} \right) \varepsilon_{\bar{k}_1, \bar{k}_2} \end{aligned}$$

Now we set  $\lambda_1 = \lambda$ ,  $\lambda_2 = \lambda^\delta$ ,  $k_1 = k$ ,  $k_2 = \delta^{-1}k$ ,  $\bar{k}_1 = \bar{k}$ ,  $\bar{k}_2 = \delta^{-1}\bar{k}$ . The previous inequality becomes

$$\varepsilon_k^{(1)} \leq \frac{C}{\Delta^{16k}} \left( \lambda^{k+a} \varepsilon_{k_0}^{1+\delta_0} + \frac{1}{\lambda^{(\bar{k}-k)-(a+1)\delta\bar{k}-a}} \varepsilon_{\bar{k}} \right).$$

The proof of ii) is analog.

Finally, we can assume that  $\mathbb{E}[\beta'_1(x - \alpha)] = 0$  up to conjugate by some diffeomorphism  $U : (x, y) \mapsto (x, y + u(x))$ , and if  $\|\beta'_1\|_0 \leq 2\varepsilon_{k_0}$ , then  $f_1$  is close to a translation and hence  $P^{-1}fP$  is still close to a translation so that we can assume  $P = Id$ .

□

The convergence of the procedure will be assured by the following combinatorial Lemma

**Lemma 15.** *Let  $a$  and  $k_0$  in  $\mathbb{N}$  and  $\delta_0 > 0$ . There exists  $k_1$  and  $s$  in  $\mathbb{N}$  and  $\delta > 0$  such that, setting  $Q = 1 + \frac{\delta_0}{2}$  and  $\lambda_n = e^{Q^n}$ , if  $(\varepsilon_k^{(n)})_{k,n \in \mathbb{N}}$  and  $(C_k)_{k \in \mathbb{N}}$  are families of positive real numbers satisfying for every integers  $n, k, \bar{k}$  with  $\bar{k} \geq k \geq k_0$*

$$\varepsilon_k^{(n+1)} \leq C_{\bar{k}} \left( \lambda_n^{k+a} (\varepsilon_{k_0}^{(n)})^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(\bar{k}-k)-\delta\bar{k}-a}} \right) \quad (6.18)$$

and  $\varepsilon_{k_1}^{(0)} \leq C_{k_1}^{-s}$ , then for every  $k, p$  in  $\mathbb{N}$ ,  $\varepsilon_k^{(n)} = O(\lambda_n^{-p})$  as  $n$  goes to  $+\infty$ .

*Proof.* Up to replacing  $\varepsilon_k^n$  by  $\varepsilon_{k+k_0}^{(n)}$ , we can assume  $k_0 = 0$ . We fix  $\delta > 0$  small enough. Let  $\bar{k}$  and  $p$  be large enough and satisfying the following conditions:

$$\begin{cases} p(1 + \delta_0) > \bar{k}(1 + \delta) + 2a \\ (1 - \delta \frac{Q}{Q-1})\bar{k} - a \frac{Q}{Q-1} \geq (1 + \delta)p \\ p(1 + \delta_0) - a \geq (1 + \delta)p \end{cases} \quad (6.19)$$

(when  $\delta = 0$ , these inequalities become  $p(1 + \delta_0) > \bar{k} + \text{constant}$ ,  $\bar{k} \geq p + \text{constant}$  and  $p \geq \text{constant}$ , hence are compatible, and still are when  $\delta$  is small enough)

We fix an integer  $N$  and we set  $M = \sup_{n \leq N} \varepsilon_0^{(n)} \lambda_n^p$ .

Using inequality (6.18) with  $k = \bar{k}$  we have

$$\varepsilon_{\bar{k}}^{(n+1)} \leq C_{\bar{k}} \left( \frac{M^{1+\delta_0}}{\lambda_n^{p(1+\delta_0)-\bar{k}-a}} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{-\delta\bar{k}-a}} \right) \leq C_{\bar{k}} (M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(n)}) \lambda_n^{\delta\bar{k}+a},$$

and we deduce

$$M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(n)} \leq 2C_{\bar{k}} \lambda_{n-1}^{\delta\bar{k}+a} (M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(n-1)}) \leq \dots \leq (2C_{\bar{k}})^n (M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(0)}) (\lambda_0 \dots \lambda_{n-1})^{\delta\bar{k}+a} \leq (2C_{\bar{k}})^n \lambda_n^{\frac{\delta\bar{k}+a}{Q-1}}$$

and hence

$$\varepsilon_{\bar{k}}^{(n)} \leq C_{\bar{k}}^n (M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(0)}) \lambda_n^{\frac{\delta\bar{k}+a}{Q-1}} \quad (6.20)$$

(up to replacing  $C_{\bar{k}}$  by  $3C_{\bar{k}}$ ). Next, inequality (6.18) with  $k = 0$  gives

$$\begin{aligned} \varepsilon_0^{(n+1)} &\leq C_{\bar{k}} \left( \lambda_n^a (\varepsilon_0^{(n)})^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(1-\delta)\bar{k}-a}} \right) \\ &\leq C_{\bar{k}} \left( \frac{M^{1+\delta_0}}{\lambda_n^{p(1+\delta_0)-a}} + \frac{C_k^n (M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(0)})}{\lambda_n^{(1-\delta\frac{Q}{Q-1})\bar{k}-a\frac{Q}{Q-1}}} \right) \\ &\leq \frac{\widetilde{M}}{\lambda_{n+1}^{(1+\delta)p}} \end{aligned}$$

where  $\widetilde{M} = C_{\bar{k}}^{n+1} (2M^{1+\delta_0} + \varepsilon_{\bar{k}}^{(0)})$ . A first consequence is that the sequence  $M_n = \sup_{k \leq n} \varepsilon_0^{(k)} \lambda_k^p$  satisfies the inequality

$$M_{n+1} \leq C_{\bar{k}}^{n+1} (M_n^{1+\delta_0} + \varepsilon_{\bar{k}}^{(0)}),$$

which implies that  $(M_n)$  is bounded if

$$\varepsilon_{\bar{k}}^{(0)} \leq C_{\bar{k}}^{-\frac{1}{\delta_0}}.$$

A second consequence is that once we have proved that  $\varepsilon_0^{(n)} = O(\lambda_n^{-p})$  for some  $p$ , we have  $\varepsilon_0^{(n)} = O(\lambda_n^{-(1+\delta)p})$ , and by induction we deduce that  $\varepsilon_0^{(n)} = O(\lambda_n^{-p})$  for every  $p > 0$ .

Finally, we deduce by (6.20) that for every  $\bar{k}$ ,

$$\varepsilon_{\bar{k}}^{(n)} = O(\lambda_n^{\frac{\delta\bar{k}}{Q-1}+a})$$

and next by inequality (6.18) that for every  $p, k, \bar{k}$ ,

$$\varepsilon_k^{(n+1)} = O \left( \frac{1}{\lambda_n^{p(1+\delta_0)-k-a}} + \frac{1}{\lambda_n^{(1-\delta\frac{Q}{Q-1})\bar{k}-k-a\frac{Q}{Q-1}}} \right)$$

which gives the result choosing  $\bar{k}$  and  $p$  large enough.  $\square$

By successive uses of the Lemma 14, we will now define a sequence of random diffeomorphisms  $f_n$  conjugated to  $f$  and converging to a random Anzai cocycle, defined by  $f_0 = f$  and the induction relation  $f_{n+1} = G_n f_n G_n^{-1}$  where  $G_n$  is constructed using point i) or ii) of Lemma 14.

Thus, we assume now that  $f = R + \zeta$  satisfies assumptions of theorem 1. If  $(\alpha_0, \hat{\beta}_0)$  is a  $(A, \sigma)$ -diophantine rotation vector of  $f$ , we can assume that  $\int_{\mathbb{T}^2} (R - Id) dz = (\alpha_0, \hat{\beta}_0)$ . Indeed, we have  $\int_{\mathbb{T}^2} (R - Id) d\mu = (\alpha_0, \hat{\beta}_0)$  for some stationary measure  $\mu$  for  $f$ . The first coordinate of  $R - Id$  is constant, hence

close to  $\alpha_0$ . In consequence  $\mu$  is small on maps on the form  $\mathbb{E}[\varphi(x + \alpha_0)] - \varphi(x)$ , and using that  $\alpha_0$  is diophantine, we deduce that  $\mu$  is close to Lebesgue on maps depending only on  $x$ , and so  $\int_{\mathbb{T}^2} (R - Id) d\mu$  is close to  $\int_{\mathbb{T}^2} (R - Id) dz$ .

Now we fix  $\delta > 0$ , we set  $\lambda_n = e^{(1+\delta_0/2)^n}$  as in Lemma 15, we define  $f_0 = f$ ,  $R_0 = R$ ,  $\zeta_0 = \zeta$ , and once we have constructed  $f_n = R_n + \zeta_n$ , if Lemma 14 applies we define  $f_{n+1} = G_n f_n G_n^{-1} = R_{n+1} + \zeta_{n+1}$  where  $G_n = P_n + g_n$  is a conjugation constructed using point i) or ii) of Lemma 14 with  $\lambda = \lambda_n$  (construction ii) will be used with some fixed integer  $p$  we will chose later). We set  $H_n = G_{n-1} \circ \dots \circ G_0$ ,  $\varepsilon_k^{(n)} = \|\zeta_n\|_{\delta k, k}$  and  $\Delta_n = \|R_n \circ \tilde{R}_n - \tilde{R}_n \circ R_n\|_{L^2(\mathbb{T}^2)}$ , where  $\tilde{R}_n$  is an independant copy of  $R_n$ . More precisely, we will use construction ii) of Lemma 14 as long as  $\Delta_n$  is small behind  $\zeta_n$ , and construction i) once  $\Delta_n$  is large behind  $\zeta_n$ .

Using that  $(\alpha_0, \hat{\beta}_0)$  is a diophantine random rotation vector of  $f$ , we can assume that  $R_n$  is diophantine as long as  $f_n$  is defined: indeed, by Proposition 4, the stationary measure  $\mu_n = (H_n)_* \mu$  for  $f_n$  is close to the Lebesgue measure (up to a remainder  $O\left(\left(\varepsilon_{k_0}^{(n-1)}\right)^2\right)$ ), hence

$$(\alpha_0, \hat{\beta}_0) = \int_{\mathbb{T}^2} (H_n \circ f_n - H_n) d\mu_n \approx \int_{\mathbb{T}^2} (H_n \circ R_n - H_n) dz = P(\alpha_n, \hat{\beta}_n)$$

and hence we can assume  $R_n$  is diophantine up to adding it a constant.

There will exist  $n_0$  such that the diffeomorphism  $G_n$  will be defined using construction ii) for  $n < n_0$  (with some fixed  $p$  we will chose later) and construction i) for  $n \geq n_0$ .

As long as  $f_n$  is defined,  $\varepsilon_k^{(n)}$  satisfies for  $n \leq n_0$

$$\varepsilon_k^{(n+1)} \leq C_{\bar{k}} \left( \lambda_n^{k+a} \left( \varepsilon_{k_0}^{(n)} + \sqrt{\Delta_n^p \varepsilon_{k_0}^{(n)}} \right)^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(\bar{k}-k)-ap}} \right)$$

and for  $n \geq n_0$

$$\varepsilon_k^{(n+1)} \leq \frac{C_{\bar{k}}}{\Delta_n^{16\bar{k}}} \left( \lambda_n^{k+a} \left( \varepsilon_{k_0}^{(n)} \right)^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(\bar{k}-k)-\delta\bar{k}-a}} \right).$$

If we control derivatives of  $H_n$ , then the numbers  $\Delta_n$  are comparable. If we assume that  $\Delta_n \leq 2\Delta_m$  for every  $n, m$ , then for  $n \geq n_0$  we have

$$\varepsilon_k^{(n+1)} \leq \frac{C_{\bar{k}}}{\Delta_{n_0}^{16\bar{k}}} \left( \lambda_n^{k+a} \left( \varepsilon_{k_0}^{(n)} \right)^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(\bar{k}-k)-\delta\bar{k}-a}} \right),$$

and so by Lemma 15, there exists integers  $s$  and  $k_1$  such that if  $\varepsilon_{k_1}^{(m)} \leq \left( \frac{\Delta_m}{C_{k_1}} \right)^s$ ,



then for every  $n \geq m$  and every  $k, q > 0$

$$\varepsilon_k^{(n)} \leq \frac{C}{\lambda_n^q}$$

for some constant  $C = C(k, q)$ .

Thus, we use the following scheme: As long as  $\Delta_n^s \leq C_{k_1}^s \varepsilon_{k_1}^{(n)}$  we use construction ii) with  $p = s$  to define  $f_{n+1}$ . If  $\Delta_{n_0}^s \geq C_{k_1}^s \varepsilon_{k_1}^{(n_0)}$  for some  $n_0$ , then we use construction i) to define  $f_k$  for  $k > n$ .

Hence, for  $n < n_0$  we have

$$\varepsilon_k^{(n+1)} \leq C_k C_{k_1} \left( \lambda_n^{k+a} (\varepsilon_{k_1}^{(n)})^{1+\delta_0} + \frac{\varepsilon_{\bar{k}}^{(n)}}{\lambda_n^{(\bar{k}-k)-\delta\bar{k}-a}} \right)$$

Thus, there exists  $k_2$  in  $\mathbb{N}$  such that if  $\varepsilon_{k_2}^{(0)}$  is small enough then we have the estimate  $\varepsilon_k^{(n)} = O(\lambda_n^{-q})$  for any  $k, q$  as long as  $f_n$  is defined and  $\|H_n'\|_0$  is controlled. Let us prove that this estimate of  $\varepsilon_k^{(n)}$  implies an estimate on  $H_n$ .

Noticing that there is at most one integer  $n_0$  such that  $P_{n_0} = P \neq Id$ , we have  $H_n = P_n + h_n$  with

$$h_n = \sum_{j=0}^{n-1} P_j (g_j \circ H_{j-1}).$$

For every  $l$ , we have for some  $C$  and  $\bar{k}$

$$\begin{aligned} \|h_n\|_k &\leq \sum_{j=0}^{n-1} \|P_j (g_j \circ H_j)\|_k \\ &\leq C \sum_{j=0}^{n-1} \|P_j\| \|g_j\|_k (1 + \|h_j\|_k) \\ &\leq C \varepsilon_{\bar{k}}^{(0)} \sup_{j < n} (1 + \|h_j\|_k). \end{aligned}$$

Using an induction we deduce that  $\|h_n\|_k = O(\varepsilon_{\bar{k}}^{(0)})$ , and next that

$$\sum_{j=0}^{n-1} \|P_j (g_j \circ H_m)\|_l = O(\varepsilon_{\bar{k}}^{(0)}).$$

In consequence, we deduce by induction that  $f_n, G_n$  and  $H_n$  are defined for every  $n$ , and that  $h_n$  normally converges in  $C^l(\mathbb{T}^2)$  to some limit  $h$  satisfying  $\|h\|_k = O(\varepsilon_{\bar{k}}^{(0)})$ , and  $H = P + h$  is thus  $C^\infty$ , close to  $P$  and invertible if  $f$  is close enough to  $R$ , and  $HfH^{-1}$  is almost surely a limit of Anzai cocycles, hence is almost surely an Anzai cocycle.



# Bibliographie

- [1] Hirotada Anzai. Ergodic skew product transformations on the torus. *Osaka Math. J.*, 3 :83–99, 1951.
- [2] VI Arnold. Small divisors I : On mappings of the circle onto itself. *Amer. Math. Soc. Transl., Ser, 2(46)* :213–284, 1965.
- [3] A. Avila and R. Krikorian. Reducibility or nonuniform hyperbolicity for quasiperiodic schrödinger cocycles. *Annals of Mathematics*, 164 :911–940, 2006.
- [4] A. Avila and M. Viana. Extremal Lyapunov exponents of smooth cocycles. *preprint*.
- [5] L. Barreira and Y.B. Pesin. *Nonuniform hyperbolicity : Dynamics of systems with nonzero Lyapunov exponents*, volume 115. Cambridge Univ Pr, 2007.
- [6] D. Dolgopyat and R. Krikorian. On simultaneous linearization of diffeomorphisms of the sphere. *Duke Mathematical Journal*, 136(3) :475–506, 2007.
- [7] B. Fayad and K. Khanin. Smooth linearization of commuting circle diffeomorphisms. *Annals of Mathematics*, 170 :101–101, 2009.
- [8] H. Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(3) :377–428, 1963.
- [9] M.R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Publications Mathématiques de l’IHÉS*, 49(1) :5–233, 1979.
- [10] Y. Kifer. Ergodic theory of random transformations. 1986.
- [11] Y. Le jan. Equilibre statistique pour les produits de difféomorphismes aléatoires indépendants. In *Annales de l’IHP Probabilités et statistiques*, volume 23, pages 111–120. Elsevier, 1987.
- [12] F. Ledrappier. Quelques propriétés des exposants caractéristiques, école d’été de probabilités de saint-flour, xii 1982, 305 396. *Lecture Notes in Math*, 1097.
- [13] F. Ledrappier. Positivity of the exponent for stationary sequences of matrices. *Lyapunov Exponents*, pages 56–73, 1986.

- [14] J. Moser. On commuting circle mappings and simultaneous Diophantine approximations. *Mathematische Zeitschrift*, 205(1) :105–121, 1990.
- [15] VI Oseledec. A multiplicative ergodic theorem, lyapunov characteristic numbers for dynamical systems, trudy mosk. *Mat. Obsc. 19 (1968)*, 179, 210.
- [16] L. Pastur and A. Figotin. Spectra of random and almost-periodic operators. *Grundlehren der mathematischen Wissenschaften*, 297.
- [17] G. Royer. Croissance exponentielle de produits markoviens de matrices aléatoires. *Ann. Inst. Henri Poincaré*, 16 :49–62, 1980.
- [18] AD Virtser. On products of random matrices and operators. *Theory of Probability and its Applications*, 24 :367, 1979.
- [19] J.C. Yoccoz. Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne. *Ann. Sci. Ecole Norm. Sup.(4)*, 17(3) :333–359, 1984.