Instituto Nacional de Matemática Pura e Aplicada

# Continuidade dos expoentes de Lyapunov para 2D-matrizes aleatórias

**Carlos Bocker Neto**

Tese apresentada para obtenção do grau de
Doutor em Ciências

Rio de Janeiro
Dezembro/2009

INSTITUTO NACIONAL DE MATEMÁTICA PURA E APLICADA – IMPA

ORIENTADOR: MARCELO MIRANDA VIANA DA SILVA

COMISSÃO EXAMINADORA

MARCELO MIRANDA VIANA DA SILVA

ARTUR AVILA CORDEIRO DE MELO

ENRIQUE RAMIRO PUJALS

ALEXANDER EDUARDO ARBIETO MENDOZA

JAIRO DA SILVA BOCHI

AUTOR:

CARLOS BOCKER NETO

# Dedicatória

À minha filha Isabele

# Agradecimentos

Ao Instituto de Matemática Pura e Aplicada - IMPA.

Ao Marcelo Viana pela confiança, apoio e especialmente pela *excelente* orientação em todos os momentos que estive no IMPA que muitas vezes foram além do ensino.

Ao Jiagang Yang pela ajuda e conversas principalmente no início deste trabalho. Também agradeço a Alexandre Simas pelos diversos auxílios e a Eduardo Teles pela confecção das figuras para a apresentação desta tese.

A todos amigos e colegas do IMPA. Especialmente a Artem Raibekas, Cristina Lizana, Ítalo Raony, José Régis, Michel Cambrainha, Mohammad Fanaee, Samuel e a todos aqueles que por algum momento tiveram a paciência de ouvir e conversar sobre este trabalho.

Aos membros da banca e ao professor Carlos Gustavo Moreira (Gugu).

A todos os funcionários do IMPA pela contribuição para que o IMPA seja este excelente ambiente de estudo e produtividade científica.

Ao CNPq e a Faperj pelo auxílio financeiro.

Aos professores Marcio Lima do Nascimento, Francisco Julio Sobreira de A. Corrêa e Jacob Palis Jr. pelo apoio para que eu chegasse até este instituto.

Aos amigos da república Yestin, Ricardo Javier, Eduardo, Mitchael, John Beiro, Sérgio, Alien, Edilberto, Anita.

À minha filha Isabele pelo carinho e pelo sorriso maravilhoso.

À minha esposa Bernadete pelo companheirismo e dedicação.

A meus pais Álvaro e Terezinha pelo apoio e carinho incondicionais.

Enfim, agradeço a todos que, direta ou indiretamente, contribuiram para que este trabalho fosse realizado.

# Resumo de Tese de Doutorado

Continuidade dos expoentes de Lyapunov para 2D-matrizes aleatórias

Carlos Bocker Neto
*Orientador*: Marcelo Viana
*Data da defesa*: 14 de dezembro de 2009
*Palavras chaves*: continuidade dos expoentes de Lyapunov, *su*-estados, medidas estacionárias.

*Resumo*: A variação dos expoentes de Lyapunov é o tema central da tese. Como resultado principal nós mostramos que os expoentes de Lyapunov de cociclos 2-dimensionais dependendo somente de uma coordenada sobre shifts de Bernoulli dependem continuamente do cociclo e da probabilidade invariante.

# Contents

# Abstract

The Lyapunov exponents of zero range $\mathrm{GL}(2,\mathbb{C})$-cocycles over Bernoulli shifts depend continuously on the cocycle and on the invariant probability.

# Resumo

Os expoentes de Lyapunov de cociclos $\mathrm{GL}(2,\mathbb{C})$ dependendo somente de uma coordenada sobre shifts de Bernoulli dependem continuamente do cociclo e da probabilidade invariante.

# Chapter 1

# Introduction

Let $A_1, \ldots, A_m$ be invertible $2 \times 2$ matrices and let $p_1, \ldots, p_m$ be (strictly) positive numbers with $p_1 + \cdots + p_m = 1$. Consider

$$L^n = L_{n-1} \cdots L_1 L_0, \quad n \geq 1,$$

where the $L_j$ are independent random variables with identical probability distributions, given by

$$\text{probability}(\{L_j = A_i\}) = p_k \quad \text{for all } j \geq 0 \text{ and } i = 1, \ldots, m.$$

It is a classical fact, going back to Furstenberg, Kesten [15], that there exist numbers $\lambda_+$ and $\lambda_-$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|L^n\| = \lambda_+ \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} \log \|(L^n)^{-1}\|^{-1} = \lambda_- \tag{1.1}$$

almost surely. The results in this paper imply that these *extremal Lyapunov exponents* always vary continuously with the choice of the matrices and the probability weights:

**Theorem A.** *The extremal Lyapunov exponents $\lambda_+$ and $\lambda_-$ depend continuously on $(A_1, \ldots, A_m, p_1, \ldots, p_m)$ at all points.*

This conclusion holds in much more generality. Indeed, we may take the probability distribution of the random variables $L_j$ to be any probability measure $\nu$ on $\mathrm{GL}(2, \mathbb{C})$ with compact support. Let $\lambda_+(\nu)$ and $\lambda_-(\nu)$, respectively, denote the values of the (almost certain) limits in (1.1). Then we have:

**Theorem B.** *For every $\varepsilon > 0$ there exists $\delta > 0$ and a weak* neighborhood $V$ of $\nu$ in the space of probability measures on $\mathrm{GL}(2, \mathbb{C})$ such that $|\lambda_\pm(\nu) - \lambda_\pm(\nu')| < \varepsilon$ for every probability measure $\nu' \in V$ whose support is contained in the $\delta$-neighborhood of the support of $\nu$.*

The situation in Theorem A corresponds to the special case when the measures have finite supports:

$$\nu = p_1 \delta_{A_1} + \cdots + p_m \delta_{A_m} \quad \text{and} \quad \nu' = p'_1 \delta_{A'_1} + \cdots + p'_m \delta_{A'_m}.$$

Clearly, the support of $\nu'$ is Hausdorff close to the support of $\nu$ if $A'_i$, $p'_i$ are close to $A_i$, $p_i$ for all $i$. In this regard, recall that we assume that all $p_i > 0$: the conclusion of Theorem A may fail if this condition is removed, as we will see in Remark 7.1.5.

# Chapter 2

# Continuity of Lyapunov exponents

In this chapter we put the previous results in a broader context and give a convenient translation of Theorem B to the theory of linear cocycles.

## 2.1 Linear cocycles

Let $\pi : \mathcal{V} \to M$ be a finite-dimensional (real or complex) vector bundle and $F : \mathcal{V} \to \mathcal{V}$ be a *linear cocycle* over some measurable transformation $f : M \to M$. By this we mean that $\pi \circ F = f \circ \pi$ and the actions $F_x : \mathcal{V}_x \to \mathcal{V}_{f(x)}$ on the fibers are linear isomorphisms. Take $\mathcal{V}$ to be endowed with some measurable Riemannian metric, that is, an Hermitian product on each fiber depending measurably on the base point. Let $\mu$ be an $f$-invariant probability measure on $M$ such that

$$\log \|(F_x)^{\pm 1}\| \in L^1(\mu).$$

Then it follows from the sub-additive ergodic theorem (Kingman [24]) that the numbers

$$\lambda_+(F,x) = \lim_{n\to\infty} \frac{1}{n} \log \|F_x^n\| \quad \text{and} \quad \lambda_-(F,x) = \lim_{n\to\infty} \frac{1}{n} \log \|(F_x^n)^{-1}\|^{-1}$$

are well-defined $\mu$-almost everywhere.

The theorem of Oseledets [30] provides a more detailed statement. Namely, at $\mu$-almost every point $x \in M$, there exist numbers

$$\hat\lambda_1(F,x) > \cdots > \hat\lambda_{k(x)}(F,x)$$

and a filtration

$$\mathcal{V}_x = \hat E_x^1 > \hat E_x^2 > \cdots > \hat E_x^{k(x)} > \{0\} = \hat E_x^{k(x)+1} \tag{2.1}$$

such that $F_x(\hat E_x^j) = \hat E_{f(x)}^j$ and

$$\lim_{n\to\infty} \frac{1}{n} \log \|F_x^n(v)\| = \hat\lambda_j(F,x) \quad \text{for all } v \in \hat E_x^j \setminus \hat E_x^{j+1}.$$

When $f$ is invertible one can say more: there exists a splitting

$$\mathcal{V}_x = E_x^1 \oplus E_x^2 \oplus \cdots \oplus E_x^{k(x)} \tag{2.2}$$

such that $F_x(E_x^j) = E_{f(x)}^j$ and

$$\lim_{n\to\pm\infty} \frac{1}{n} \log \|F_x^n(v)\| = \hat{\lambda}_j(F,x) \quad \text{for all } v \in E_x^j \setminus \{0\}.$$

The number $k(x) \geq 1$ and the *Lyapunov exponents* $\hat{\lambda}_j(F,\cdot)$ are measurable functions of the point $x$, with

$$\hat{\lambda}_1(F,x) = \lambda_+(F,x) \quad \text{and} \quad \hat{\lambda}_{k(x)}(F,x) = \lambda_-(F,x),$$

and they are constant on the orbits of $f$. In particular, they are constant $\mu$-almost everywhere if $\mu$ is ergodic.

## 2.2 Continuity problem

Next, let $\lambda_1(F,x) \geq \cdots \geq \lambda_d(F,x)$ be the list of all Lyapunov exponents, where each is counted according to its multiplicity $m_j(x) = \dim \hat{E}_x^j - \dim \hat{E}_x^{j+1}$ ($= \dim E_x^j$ in the invertible case). Of course, $d =$ dimension of $\mathcal{V}$. The *average Lyapunov exponents* of $F$ are defined by

$$\lambda_i(F,\mu) = \int \lambda_i(F,\cdot)\, d\mu, \quad \text{for } i = 1,\ldots,d.$$

The results in this paper are motivated by the following basic question:

**Problem 2.2.1.** What are the continuity points of

$$(F,\mu) \mapsto (\lambda_1(F,\mu),\ldots,\lambda_d(F,\mu)) \;?$$

It is well known that the sum of the $k$ largest Lyapunov exponents (any $1 \leq k < d$)

$$F \mapsto \lambda_1(F,\mu) + \cdots + \lambda_k(F,\mu)$$

is always upper semi-continuous relative to the $L^\infty$-norm in the space of cocycles. Indeed, this is an easy consequence of the identity

$$\lambda_1(F,\mu) + \cdots + \lambda_k(F,\mu) = \inf_{n\geq 1} \frac{1}{n} \int \log \|\Lambda^k(F_x^n)\|\, d\mu(x) \tag{2.3}$$

where $\Lambda^k$ denotes the $k$th exterior power. Similarly, the sum of the $k$ smallest Lyapunov exponents is always lower semi-continuous. However, Lyapunov exponents are, usually, *discontinuous* functions of the data. A number of results, both positive and negative, will be recalled in a while.

## 2.3 Continuity theorem

Let $\mathcal{X}$ be a polish space, that is, a separable completely metrizable topological space. Let $p$ be a probability measure on $\mathcal{X}$ and $A : \mathcal{X} \to \mathrm{GL}(2, \mathbb{C})$ be a measurable function such that

$$\log \|A^{\pm 1}\| \quad \text{are bounded.} \tag{2.4}$$

Let $f : M \to M$ be the shift map on $M$ and let $\mu = p^{\mathbb{Z}}$. Consider the linear cocycle

$$F : M \times \mathbb{C}^2 \to M \times \mathbb{C}^2, \quad F(\mathbf{x}, v) = (f(\mathbf{x}), A_{x_0}(v)),$$

where $x_0 \in \mathcal{X}$ denotes the zeroth coordinate of $\mathbf{x} \in M$. In the spaces of cocycles and probability measures on $\mathcal{X}$ we consider the distances defined by, respectively,

$$d(A, B) = \sup_{x \in \mathcal{X}} \|A_x - B_x\| \quad d(p, q) = \sup_{|\phi| \leq 1} |\int \phi \, d(p - q)| \tag{2.5}$$

where the second sup is over all measurable functions $\phi : \mathcal{X} \to \mathbb{R}$ with $\sup |\phi| \leq 1$. In the space of pairs $(A, p)$ we consider the topology determined by the bases of neighborhoods

$$V(A, p, \varepsilon, \mathcal{Z}) = \{(B, q) : d(A, B) < \varepsilon, \; q(\mathcal{Z}) = 1, \; d(p, q) < \varepsilon\} \tag{2.6}$$

where $\varepsilon > 0$ and $\mathcal{Z}$ is any measurable subset of $\mathcal{X}$ with $p(\mathcal{Z}) = 1$.

**Theorem C.** *The extremal Lyapunov exponents $\lambda_\pm(A, p) = \lambda_\pm(F, \mu)$ depend continuously on $(A, p)$ at all points.*

It is easy to deduce Theorem C from Theorem B: if $d(A, B)$ and $d(p, q)$ are small then $\nu' = B_* q$ is weak* close to $\nu = A_* p$ and the support of $\nu'$ is contained in a small neighborhood of the support of $\nu$; moreover, $\lambda_\pm(A, p) = \lambda_\pm(\nu)$ and $\lambda_\pm(B, q) = \lambda_\pm(\nu')$. In this way one even gets a more general version of Theorem C, where $\mathcal{X}$ can be any measurable space. In fact, our presentation goes the other way around: we prove Theorem C directly, in Chapters 3 and 4, and then we deduce Theorem B from it, in Chapter 6.1.

We also get that the Oseledets decomposition depends continuously on the cocycle *in measure*. Given $B : \mathcal{X} \to \mathrm{GL}(2, \mathbb{C})$, let $E^s_{B,\mathbf{x}}$ and $E^u_{B,\mathbf{x}}$ be the Oseledets subspaces of the corresponding cocycle at a point $\mathbf{x} \in M$ (when they exist).

**Theorem D.** *Suppose $\lambda_-(A, p) < \lambda_+(A, p)$. Then for any sequence $A^k : \mathcal{X} \to \mathrm{GL}(2, \mathbb{C})$ such that $d(A^k, A) \to 0$, and for any $\varepsilon > 0$,*

$$\mu\big(\{x \in M : \angle(E^u_{A,x}, E^u_{A^k,x}) < \epsilon \text{ and } \angle(E^s_{A,x}, E^s_{A^k,x}) < \epsilon\}\big) \to 1.$$

A few words are in order on our choice of the topology (2.6). As we are going to see, the proof of Theorem C splits into two cases, depending on whether the cocycle is almost irreducible (Chapter 3.1) or diagonal (Chapter 3.2). In the irreducible case, continuity of the Lyapunov exponents was known before ([16], see also [3]) and only requires the weak* topology. In a nutshell, this is because in the irreducible case

$$\lambda_+(A,p) = \int \log \frac{\|A(\mathbf{x})(v)\|}{\|v\|}\, d\mu(\mathbf{x})\, d\eta(v) \tag{2.7}$$

for *every* stationary measure $\eta$. Then one only has to note that the set of stationary measures varies semi-continuously with the data. The main point in Theorem C is to handle the diagonal case, where (2.7) breaks down, and that is where we need the full strength of (2.6).

Restricted to the space of pairs $(A,p)$ where $A$ is continuous (and bounded), it suffices to consider the neater bases of neighborhoods

$$V(A,p,\varepsilon) = \{(B,q) : d(A,B) < \varepsilon,\ \operatorname{supp} q \subset \operatorname{supp} p,\ d(p,q) < \varepsilon\}. \tag{2.8}$$

However, this will not be used in the present paper.

## 2.4 Some previous results

The issue of dependence of Lyapunov exponents on the linear cocycle or the base dynamics has been addressed by several authors. In a pioneer work, Ruelle [35] proved real-analytic dependence of the largest exponent on the cocycle, for linear cocycles admitting an invariant convex cone field. Short afterwards, Furstenberg, Kifer [16, 23] and Hennion [20] studied the dependence of the largest exponent of i.i.d. random matrices on the probability distribution, proving continuity with respect to the weak* topology in the essentially irreducible case. Kifer [23] also observed that discontinuities may occur when the probability vector degenerates (cf. Remark 7.1.5 below). Moreover, Johnson [21] found examples of discontinuous dependence of the exponent on the energy $E$, for Schrödinger cocycles over quasi-periodic flows.

For i.i.d. random matrices satisfying strong irreducibility and the contraction property, Le Page [31, 32] proved local Hölder continuous and even smooth dependence of the largest exponent on the cocycle; the assumptions ensure that the largest exponent is simple (multiplicity 1), by work of Guivarc'h, Raugi [19] and Gol'dsheid, Margulis [17]. Le Page's result can not be improved: a construction of Halperin quoted by Simon, Taylor [36] shows that for every $\alpha > 0$ one can find *random Schrödinger cocycles*

$$\begin{pmatrix} E - V_n & -1 \\ 1 & 0 \end{pmatrix}$$

(the $V_n$ are i.i.d. random variables) near which the exponents fail to be $\alpha$-Hölder continuous. For i.i.d. random matrices with finitely many values and, more generally, for locally constant cocycles over Markov shifts, Peres [33] showed that simple exponents are locally real-analytic functions of the transition data.

Recently, Bourgain, Jitomirskaya [12, 13] proved continuous dependence of the exponents on the energy $E$, for *quasi-periodic* Schrödinger cocycles that is, with $V_n = V(f^n(\theta))$ where

$f$ is a quasi-periodic translation on a torus. Furthermore, Avila, Viana [3] studied the continuity of the Lyapunov exponents in the very broad context of *smooth* cocycles. The continuity criterium in [3, Section 5] is the starting point for the proof of our Theorem C.

### Organization

In Chapter 3 we reduce Theorem C to a key result on stationary measures of nearby cocycles. This key statement is proved in Chapter 4. In Chapter 6.1 we deduce Theorems B and D. Finally, in Chapter 7 we describe an example of discontinuity of Lyapunov exponents for Hölder cocycles, and we close with a short list of open problems and conjectures.


### Acknowledgements

We are grateful to Artur Avila, Jairo Bochi, and Jiagang Yang for very useful conversations. Lemma 6.1.1 is due to Artur Avila.


## 2.5 Continuity versus cocycle regularity

Going back to linear cocycles, the answer to the Continuity Problem is bound to depend on the class of cocycles under consideration, including its topology. Knill [25, 26] considered $L^\infty$ cocycles with values in $\mathrm{SL}(2,\mathbb{R})$ and proved that, as long as the base dynamics is aperiodic, discontinuities always exist: the set of cocycles with non-zero exponents is never open. This was refined to the $C^0$ case by Bochi [5, 6], partly inspired also by Mañé [29]: an $\mathrm{SL}(2,\mathbb{R})$-cocycle is a continuity point in the $C^0$ topology if and only if it is uniformly hyperbolic or else the exponents vanish. Most striking, the theorem of Mañé-Bochi [6, 29] remains true restricted to the subset of $C^0$ derivative cocycles, that is, of the form $F = Df$ for some $C^1$ area preserving surface diffeomorphism $f$. These results have been extended to arbitrary dimension by Bochi, Viana [7, 8]. Let us also note that $\mathrm{GL}(d,\mathbb{R})$-cocycles whose exponents are all equal form an $L^p$-residual subset, for any $p \in [1,\infty)$, by Arnold, Cong [2], Arbieto, Bochi [1]. Consequently, they are precisely the continuity points for the Lyapunov exponents relative to the $L^p$ topology.

These results show that discontinuity of Lyapunov exponents is quite common among cocycles with low regularity. Locally constant cocycles, as we deal with here, sit at the opposite end of the regularity spectrum, and the results in the present paper show that in this context continuity does hold at every point. For cocycles with intermediate regularities the Continuity Problem is very much open. However, our construction in Chapter 7.1 shows that for any $r \in (0,\infty)$ there exist locally constant cocycles over Bernoulli shifts that are points of discontinuity for the Lyapunov exponents in the space of all $r$-Hölder cocycles.

# Chapter 3

# Proof of Theorem C

We start with a simple observation. Let $\mathcal{P}(\mathcal{X})$ be the space of probability measures on $\mathcal{X}$ and let $\mathcal{G}(\mathcal{X})$ and $\mathcal{S}(\mathcal{X})$ denote the spaces of bounded measurable functions from $\mathcal{X}$ to $\mathrm{GL}(2,\mathbb{C})$ and $\mathrm{SL}(2,\mathbb{C})$, respectively. Given any $A \in \mathcal{G}(X)$ let $B \in \mathcal{S}(\mathcal{X})$ and $c : \mathcal{X} \to \mathbb{C}$ be such that $A_x = c_x B_x$ for every $x \in \mathcal{X}$. Although $c_x = (\det A_x)^{1/2}$ and $B_x$ are determined up to sign only, choices can be made consistently in a neighborhood, so that $B$ and $c$ depend continuously on $A$. It is also easy to see that the Lyapunov exponents are related by

$$\lambda_\pm(A,p) = \lambda_\pm(B,p) + \int \log |c_x| \, dp(x)$$

Thus, since the last term depends continuously on $(A,p)$ relative to the topology defined by (2.6), continuity of the Lyapunov exponents on $\mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ yields continuity on the whole $\mathcal{G}(X) \times \mathcal{P}(\mathcal{X})$. So, we may suppose from the start that $A \in \mathcal{S}(\mathcal{X})$. Observe also that in this case one has

$$\lambda_+(A,p) + \lambda_-(A,p) = 0.$$

From here on the proof has two main steps. First, we reduce the problem to the case when the matrices are simultaneously diagonalizable:

**Proposition 3.0.1.** *If $(A,p) \in \mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ is a point of discontinuity for $\lambda_+$ then there is $P \in \mathrm{SL}(2,\mathbb{C})$ and $\theta : \mathcal{X} \to \mathbb{C} \setminus \{0\}$ such that*

$$PA_xP^{-1} = \begin{pmatrix} \theta_x & 0 \\ 0 & \theta_x^{-1} \end{pmatrix}$$

*for all $x \in \mathcal{Y}$, where $\mathcal{Y} \subset \mathcal{X}$ is a full $p$-measure set. In particular, $A_xA_y = A_yA_x$ for all $x, y \in \mathcal{Y}$.*

Then we rule out the diagonal case as well:

**Proposition 3.0.2.** *Let $(A,p) \in \mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ be such that $A$ is as in the conclusion of Proposition 3.0.1. Then $(A,p)$ is a point of continuity for $\lambda_+$.*

The proofs of these two propositions are given in the next couple of chapters. In view of the previous observations, they contain the proof of Theorem C.

## 3.1 Reducing to the diagonal case

The proof of Proposition 3.0.1 is a simplified version of ideas of Avila, Viana [3], partly inspired by Bonatti, Gomez-Mont, Viana [10]. For the sake of completeness, and also because our setting is not strictly contained in [3], we present the complete arguments. The definitions and preliminary results apply to functions $A$ with values in $\mathrm{GL}(d, \mathbb{C})$, for any $d \geq 2$.

The *local stable set* $W^s_{loc}(\mathbf{x})$ of $\mathbf{x} \in M$ is the set of all $\mathbf{y} = (y_n)_{n\in\mathbb{Z}}$ such that $x_n = y_n$ for all $n \geq 0$. The *local unstable set* $W^u_{loc}(\mathbf{x})$ is defined similarly, considering $n < 0$ instead. The *projective cocycle* associated to $A : \mathcal{X} \to \mathrm{GL}(d, \mathbb{C})$ is defined by

$$F_A : M \times \mathbb{P}(\mathbb{C}^d) \to M \times \mathbb{P}(\mathbb{C}^d), \quad (\mathbf{x}, [v]) \mapsto (f(\mathbf{x}), [A(\mathbf{x})v])$$

where $A(\mathbf{x}) = A_{x_0}$ for every $\mathbf{x} \in M$.

### 3.1.1 Invariant $u$-states

Let $\mathcal{M}(p)$ denote the set of probability measures in $M \times \mathbb{P}(\mathbb{C}^d)$ that project down to $\mu$. A *disintegration* of $m \in \mathcal{M}(p)$ is a measurable function assigning to each point $\mathbf{x} \in M$ a probability $m_\mathbf{x}$ with $m_\mathbf{x}\big(\{\mathbf{x}\} \times \mathbb{P}(\mathbb{C}^d)\big) = 1$ and such that

$$m(E) = \int m_\mathbf{x}(E)\, d\mu(\mathbf{x}), \quad \text{for every measurable } E \subset M \times \mathbb{P}(\mathbb{C}^d).$$

A disintegration always exists in this setting and it is essentially unique; see Rokhlin [34] and [9, Appendix C.6].

A probability $m \in \mathcal{M}(p)$ is a *$u$-state* if some disintegration $\mathbf{x} \mapsto m_\mathbf{x}$ is constant on every local unstable set, restricted to a full $\mu$-measure subset of $M$. Then the same is true for every disintegration, by essential uniqueness; moreover, one can choose the disintegration so that it is constant on local unstable sets on the whole $M$. If $m$ is an invariant probability then we say that $m$ is an *invariant $u$-state*. The definition of *invariant $s$-states* is analogous, considering local stable sets instead, and the same observations apply.

An *su-state* is a probability which is both a $u$-state and an $s$-state.

**Lemma 3.1.1.** *A probability $m \in \mathcal{M}(p)$ is an invariant su-state if and only if $m = \mu \times \eta$ for some probability measure $\eta$ on $\mathbb{P}(\mathbb{C}^d)$ invariant under the action of $A_x$ for $p$-almost every $x \in \mathcal{X}$.*

*Proof.* The "if" part is not used in this paper, so we leave the proof to the reader. To prove the "only if" part notice that, by assumption, $m$ admits disintegrations $\mathbf{x} \mapsto m^u_\mathbf{x}$, constant on local unstable sets, and $\mathbf{x} \mapsto m^s_\mathbf{x}$, constant on local stable sets. By essential uniqueness, there exists a full $\mu$-measure set $X \subset M$ such that $m^u_\mathbf{x} = m^s_\mathbf{x}$ for all $\mathbf{x} \in X$. The assumption on $\mu$ implies that $\mu = \mu^u \times \mu^s$ where $\mu^u$ is a probability on the set positive one-sided sequences $(x_n)_{n\geq 0}$ and $\mu^s$ is a probability on the set negative one-sided sequences $(x_n)_{n<0}$. Fix $\bar{\mathbf{x}} \in M$ such that $W^u_{loc}(\bar{\mathbf{x}})$ intersects $X$ on a full $\mu^u$-measure set. Then let $\eta = m^u_{\bar{\mathbf{x}}}$. The local stable sets through the points of $X \cap W^u_{loc}(\mathbf{x})$ fill-in a full $\mu$-measure subset of $M$. Thus, $\eta = m^s_\mathbf{x}$ at

$\mu$-almost every point and so the constant family $\mathbf{x} \mapsto m_{\mathbf{x}} = \eta$ is a disintegration of $m$. This means that $m = \mu \times \eta$. Finally, the fact that $\mu$ and $m$ are invariant gives $A(\mathbf{x})_* m_{\mathbf{x}} = m_{f(\mathbf{x})}$ at $\mu$-almost every point and that implies $(A_x)_*\eta = \eta$ for $p$-almost every $x \in \mathcal{X}$, as claimed. $\square$

**Lemma 3.1.2.** *If $\lambda_\pm(A,p) = 0$ then every $F_A$-invariant measure $m$ in $\mathcal{M}(p)$ is an su-state.*

*Proof.* This is a direct consequence of Ledrappier [27, Theorem 1]. Indeed, let $\mathcal{B}^s$ be the $\sigma$-algebra of measurable subsets of $M$ which are unions of entire local stable sets. Clearly, $f$ and $F_A$ are $\mathcal{B}^s$-measurable. Hence, Ledrappier's theorem gives that the disintegration of any $F_A$-invariant probability $m \in \mathcal{M}(p)$ is $\mathcal{B}^s$-measurable modulo zero $\mu$-measure sets. This is the same as saying that $m$ is an $s$-state. Analogously, one proves that $m$ is a $u$-state. $\square$

Let us consider the function $\phi_A : M \times \mathbb{P}(\mathbb{C}^d) \to \mathbb{R}$ defined by

$$\phi_A(\mathbf{x}, [v]) = \log \frac{\|A(\mathbf{x})v\|}{\|v\|}.$$

**Lemma 3.1.3.** *For every $A : \mathcal{X} \to \mathrm{GL}(d, \mathbb{C})$ and every $F_A$-invariant probability measure $m \in \mathcal{M}(p)$,*

$$\lambda_-(A,p) \leq \int \phi_A \, dm \leq \lambda_+(A,p).$$

*Proof.* For every $(\mathbf{x}, [v]) \in M \times \mathbb{P}(\mathbb{C}^d)$ and $n \geq 1$,

$$\sum_{j=0}^{n-1} \phi_A(F_A^j(\mathbf{x}, [v])) \leq \log \|A^n(\mathbf{x})\| \, .$$

Integrating with respect to any probability $m \in \mathcal{M}(p)$,

$$\frac{1}{n} \int \sum_{j=0}^{n-1} \phi_A \circ F_A^j \, dm \leq \frac{1}{n} \int \log \|A^n(\mathbf{x})\| \, d\mu(\mathbf{x}).$$

The right hand side converges to $\lambda_+(A,p)$ and, assuming $m$ is invariant, the left hand side coincides with $\int \phi_A \, dm$. This gives the upper bound in the statement. The lower bound is analogous. $\square$

Now let $A$ take values in $\mathrm{SL}(2, \mathbb{C})$. We want to show that the upper bound in Lemma 3.1.3 is attained at some $u$-state and the lower bound is attained at some $s$-state. When $\lambda_\pm(A,p) = 0$ this is a trivial consequence of Lemma 3.1.2. So, it is no restriction to suppose that $\lambda_+(A,p) > 0 > \lambda_-(A,p)$.

Let $E^u_{\mathbf{x}} \oplus E^s_{\mathbf{x}}$ be the Oseledets splitting of $F_A$, defined at $\mu$-almost every $\mathbf{x}$. Consider the probabilities $m^u$ and $m^s$ defined on $M \times \mathbb{P}(\mathbb{C}^2)$ by

$$m^*(B) = \mu\big(\{\mathbf{x} : (\mathbf{x}, E^*_{\mathbf{x}}) \in B\}\big) = \int \delta_{(\mathbf{x}, E^*_{\mathbf{x}})}(B) \, d\mu(\mathbf{x}) \tag{3.1}$$

for $* \in \{s,u\}$ and any measurable subset $B$. It is clear that $m^u$ and $m^s$ are invariant under $F_A$ and project down to $\mu$. Moreover, their disintegrations are given by

$$\mathbf{x} \mapsto \delta_{(\mathbf{x}, E^*_{\mathbf{x}})} \quad \text{for } * \in \{s,u\}.$$

Since $E^u_{\mathbf{x}}$ depends only on $\{A_{x_n} : n < 0\}$ and $E^s_{\mathbf{x}}$ depends only on $\{A_{x_n} : n \geq 0\}$, we get that $m^u$ is a $u$-state and $m^s$ is an $s$-state.

**Lemma 3.1.4.** *Every $F_A$-invariant probability measure $m \in \mathcal{M}(p)$ is a convex combination $m = \alpha m^u + \beta m^s$, for some $\alpha, \beta \geq 0$ with $\alpha + \beta = 1$.*

*Proof.* Given $\kappa > 0$, define $X_\kappa$ to be the set of all $(\mathbf{x}, [v]) \in M \times \mathbb{P}(\mathbb{C}^2)$ such that the Oseledets splitting $E^u_{\mathbf{x}} \oplus E^s_{\mathbf{x}}$ is defined at $\mathbf{x}$ and $[v]$ splits $v = v^u + v^s$ with $\kappa^{-1}\|v^s\| \leq \|v^u\| \leq \kappa\|v^s\|$. Since the two Lyapunov exponents are distinct, any point of $X_\kappa$ returns at most finitely many times to $X_\kappa$. So, by Poincaré recurrence, $m(X_\kappa) = 0$ for every $\kappa$. This means that $m$ gives full weight to $\{(\mathbf{x}, E^u_{\mathbf{x}}), (\mathbf{x}, E^s_{\mathbf{x}}) : \mathbf{x} \in M\}$ and so it is a convex combination of $m^u$ and $m^s$. □

**Lemma 3.1.5.** $\lambda_+(A,p) = \int \phi_A \, dm^u$ *and* $\lambda_-(A,p) = \int \phi_A \, dm^s$.

*Proof.* Let $v^u_{\mathbf{x}}$ be a unit vector in the Oseledets subspace $E^u_{\mathbf{x}}$. Then

$$\begin{aligned}
\lambda_+(A, \mathbf{x}) &= \lim_{n\to\infty} \frac{1}{n} \log \|A^n(\mathbf{x}) v^u_{\mathbf{x}}\| = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \log \|A(f^j(\mathbf{x})) v^u_{f^j(\mathbf{x})}\| \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \phi_A(f^j(\mathbf{x}), E^u_{f^j(\mathbf{x})}) = \tilde{\phi}_A(\mathbf{x}, E^u_{\mathbf{x}})
\end{aligned}$$

for $\mu$-almost every $\mathbf{x}$, where $\tilde{\phi}_A$ is the Birkhoff average of $\phi_A$ for $F_A$. Hence,

$$\lambda_+(A,p) = \int \tilde{\phi}_A(\mathbf{x}, E^u_{\mathbf{x}}) \, d\mu(\mathbf{x}) = \int \tilde{\phi}_A \, dm^u = \int \phi_A \, dm^u.$$

Analogously, $\lambda_-(A,p) = \int \phi_A dm^s$. This completes the proof. □

**Remark 3.1.6.** It follows from Lemma 3.1.4 that $m^u$ is the unique invariant measure $m$ such that $\lambda_+(A,p) = \int \phi_A \, dm$.

### 3.1.2 Stationary measures

Given $(B,q)$ in $\mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$, a probability $\eta$ on $\mathbb{P}(\mathbb{C}^2)$ is called $(B,q)$-*stationary* if

$$\eta = \int (B_x)_* \eta \, dq(x). \tag{3.2}$$

The next lemma asserts that the stationary measures are the projections to $\mathbb{P}(\mathbb{C}^2)$ of the $u$-states of the corresponding cocycle. We are going to denote $M^u = \mathcal{X}^{\mathbb{Z}_+}$ and $M^s = \mathcal{X}^{\mathbb{Z}_-}$. Notice that $q^{\mathbb{Z}} = \mu^s \times \mu^u$ where $\mu^*$ is a measure on $M^*$, for $* \in \{s,u\}$.

**Lemma 3.1.7.** *If $m$ is an invariant u-state for $(B, q)$ then its projection $\eta$ to $\mathbb{P}(\mathbb{C}^2)$ is a $(B, q)$-stationary measure. Conversely, given any $(B, q)$-stationary $\eta$ there exists an invariant u-state that projects to $\eta$.*

*Proof.* Let $\mathbf{x} \mapsto m_{\mathbf{x}}$ be a disintegration of $m$ constant along unstable leaves. For any measurable set $I \subset \mathbb{P}(\mathbb{C}^2)$,

$$\eta(I) = m(M \times I) = \int m_{\mathbf{x}}(M \times I)\, d\mu(\mathbf{x}) = \int m_{f(\mathbf{x})}(M \times I)\, d\mu(\mathbf{x})$$

because $\mu$ is $f$-invariant. Since $m$ is $F_B$-invariant, the expression on the right hand side may be rewritten as

$$\int B(\mathbf{x})_* m_{\mathbf{x}}(M \times I)\, d\mu(\mathbf{x})$$
$$= \int_{M^s} \Big( \int_{M^u} B(\mathbf{x})_* m_{\mathbf{x}}(M \times I)\, d\mu^u(\mathbf{x}^u) \Big)\, d\mu^s(\mathbf{x}^s).$$

Since the disintegration is constant on local unstable sets and $B(\mathbf{x}^s, \mathbf{x}^u)$ depends only on $\mathbf{x}^s$ (we write $B(\mathbf{x}^s)$ instead), this last expression coincides with

$$\int_{\Sigma^s} B(\mathbf{x}^s)_* \Big( \int_{\Sigma^u} m_{\mathbf{x}^u}(M \times I)\, d\mu^u(\mathbf{x}^u) \Big)\, d\mu^s(\mathbf{x}^s)$$
$$= \int_{\Sigma^s} B(\mathbf{x}^s)_* \eta(I)\, d\mu^s(\mathbf{x}^s) = \int B(\mathbf{x})_* \eta(I)\, d\mu(\mathbf{x}) = \int (B_x)_* \eta(I)\, dq(x).$$

Thus, $\eta = \int (B_x)_* \eta\, dq(x)$ as claimed.

Conversely, given any $(B, q)$-stationary measure $\eta$, consider the sequence of functions

$$m^n : \mathbf{x} \mapsto m^n_{\mathbf{x}} = B^n(f^{-n}(\mathbf{x}))_* \eta$$

with values in the space of probabilities on $\mathbb{P}(\mathbb{C}^2)$. It is clear from the definition that each $m^n$ is measurable with respect to the $\sigma$-algebra $\mathcal{F}^n$ of subsets of $M$ generated by the cylinders

$$[-n : \Delta_{-n}, \dots, \Delta_{-1}] = \{\mathbf{x} \in M : x_i \in \Delta_i \text{ for } i = -n, \dots, -1\},$$

where the $\Delta_i$ are measurable subsets of $\mathcal{X}$. Observe that the $\mathcal{F}^n$ form a non-decreasing sequence. We claim that $(m^n, \mathcal{F}^n)$ is a martingale, that is,

$$\int_C m^{n+1}\, d\mu = \int_C m^n\, d\mu \quad \text{for every } C \in \mathcal{F}^n \text{ and every } n \geq 1. \tag{3.3}$$

To prove this, it suffices to consider the case when $C$ is a cylinder $[-n : \Delta_{-n}, \dots, \Delta_{-1}]$. Then, for any $n \geq 1$,

$$\int_C A^{n+1}(f^{-n-1}(\mathbf{x}))_* \eta\, d\mu(\mathbf{x}) = \int_C A^n(f^{-n}(\mathbf{x}))_* A(f^{-n-1}(\mathbf{x}))_* \eta\, d\mu(\mathbf{x})$$
$$= \int_C A^n(f^{-n}(\mathbf{x}))_* \Big[ \int_{\mathcal{X}} (A_y)_* \eta\, dp(y) \Big]\, d\mu(\mathbf{x})$$
$$= \int_C A^n(f^{-n}(\mathbf{x}))_* \eta\, d\mu$$

because $\eta$ is stationary. This proves the claim (3.3). Then, by the martingale convergence theorem (see [14, Chapter 5]), there exists a function $\mathbf{x} \mapsto m_{\mathbf{x}}$ such that $m_{\mathbf{x}}^n$ converges $\mu$-almost everywhere to $m_{\mathbf{x}}$ in the weak* topology. Let $m$ be the probability measure on $M \times \mathbb{P}(\mathbb{C}^2)$ defined by

$$m(E) = \int m_{\mathbf{x}}\big(E \cap (\{\mathbf{x}\} \times \mathbb{P}(\mathbb{C}^2))\big)\, d\mu(\mathbf{x})$$

for any measurable set $E$. By construction, the disintegration $\mathbf{x} \mapsto m_{\mathbf{x}}$ is constant on every $\{\mathbf{x}^s\} \times M^u$. This means that $m$ is a $u$-state. Also by construction, $m_{f(\mathbf{x})} = A(\mathbf{x})_* m_{\mathbf{x}}$ for $\mu$-almost every $\mathbf{x} \in M$. This proves that the $u$-state $m$ is invariant. Moreover, by (3.3) and the assumption that $\eta$ is stationary,

$$m^n(M \times I) = m^1(M \times I) = \int_M (A_x)_* \eta(I)\, dp(x) = \eta(I)$$

for every $n \geq 1$ and any measurable set $I \subset \mathbb{P}(\mathbb{C}^2)$. This means that $m^n$ projects to $\eta$ for every $n \geq 1$. Then so does the limit $m$. This completes the proof of the lemma. □

We are also going to show that the projection of $m^u$ to the projective space $\mathbb{P}(\mathbb{C}^2)$ completely determines the Lyapunov exponents:

**Lemma 3.1.8.** *Let $m$ be a $u$-state realizing $\lambda_+(A,p)$ and let $\eta$ be its projection to $\mathbb{P}(\mathbb{C}^2)$. Then*

$$\lambda_+(A,p) = \int \int_{\mathbb{P}(\mathbb{C}^2)} \log \frac{\|A_x v\|}{\|v\|} d\eta([v])\, dp(x).$$

*Proof.* Suppose first that $\lambda_+(A,p) = 0$. By Lemmas 3.1.2 and 3.1.1, every $F_A$-invariant probability $m$ which project down to $\mu$ realizes the largest exponent and is a product measure $m = \mu \times \eta$. Thus, in this case, the lemma follows immediately from Fubini's Theorem.

If $\lambda_+(A,p) > 0$, then $m^u$ is the unique $u$-state which realizes $\lambda_+$ and the lemma follows from a straightforward calculation:

$$\begin{aligned}
\lambda_+(A,p) &= \int_M \log \|A(\mathbf{x}) E^u_{\mathbf{x}}\| d\mu = \int_{M^s} \int_{M^u} \log \|A(\mathbf{x}^s) E^u_{\mathbf{x}^u}\| d\mu(\mathbf{x}^u)\, d\mu^s \\
&= \int_{\mathcal{X}} \int_{M^u} \log \|A_y E^u_{\mathbf{x}^u}\| d\mu(\mathbf{x}^u)\, dp(y) \\
&= \int_{\mathcal{X}} \int_{M^u} \int_{\mathbb{P}(\mathbb{C}^2)} \log \frac{\|A_y v\|}{\|v\|}\, d\delta_{E^u_{x^u}}\, d\mu(\mathbf{x}^u)\, dp(y) \\
&= \int_{\mathcal{X}} \int_{\mathbb{P}(\mathbb{C}^2)} \log \frac{\|A_y v\|}{\|v\|}\, d\eta([v])\, dp(y)
\end{aligned}$$

as claimed. □

**Lemma 3.1.9.** *If $(A^k, p^k)_k$ converges to $(A, p)$ and $\eta^k$ is a sequence of $(A^k, p^k)$-stationary measure converging to $\eta$ then $\eta$ is an $(A, p)$-stationary measure.*

*Proof.* We have to show that

$$\lim_k \int (A^k_x)_* \eta^k \, dp^k = \int (A_x)_* \eta \, dp$$

in the weak$^*$ sense. Let $\phi : \mathbb{P}(\mathbb{C}^2) \to \mathbb{R}$ be a continuous function. Then

$$\left| \int \int \phi(A^k_x v) \, d\eta^k \, dp^k - \int \int \phi(A_x v) \, d\eta \, dp \right| \leq a_k + b_k + c_k$$

where

$$\begin{aligned}
a_k &= \left| \int \int \phi(A^k_x v) \, d\eta^k \, dp^k - \int \int \phi(A_x v) \, d\eta^k \, dp^k \right| \\
b_k &= \left| \int \int \phi(A_x v) \, d\eta^k \, dp^k - \int \int \phi(A_x v) \, d\eta \, dp^k \right| \\
c_k &= \left| \int \int \phi(A_x v) \, d\eta \, dp^k - \int \int \phi(A_x v) \, d\eta \, dp \right|
\end{aligned}$$

It is clear that $(a_k)_k$ converges to zero, because $\|A^k_x - A_x\|$ converges uniformly to zero and $\phi$ is uniformly continuous. To prove that $b_k$ converges to zero we argue as follows. Given $\varepsilon > 0$, fix $\delta > 0$ such that $|\phi(v) - \phi(w)| < \varepsilon/3$ for all $v, w \in \mathbb{P}(\mathbb{C}^2)$ such that $d(v, w) < \delta$. Since the image of $A$ is contained in a compact subset of $\mathrm{SL}(2, \mathbb{C})$, there are $B_1, \ldots, B_n \in \mathrm{SL}(2, \mathbb{C})$ such that their $\delta$-neighborhoods cover $A(\mathcal{X})$. The assumption that $(\eta^k)_k$ converges to $\eta$ in the weak$^*$ topology implies that there exists $k_0 \in \mathbb{N}$ such that

$$\left| \int \phi(B_i v) \, d\eta^k - \int \phi(B_i v) \, d\eta \right| < \varepsilon/3$$

for all $k > k_0$ and for all $i = 1, \ldots, n$. Then we can use the triangle inequality to conclude that

$$\left| \int \phi(A_x v) \, d\eta^k - \int \phi(A_x v) \, d\eta \right| \leq \varepsilon$$

for al $k > k_0$. Integrating with respect to $p^k$ we conclude that $b_k \leq \varepsilon$ for all $k > k_0$. This proves that $b_k$ converges to 0. Finally, it is clear that $a_k$ converges to zero, because our assumptions imply that $(p^k)_k$ converges strongly to $p$. The proof of the lemma is complete. □

### 3.1.3 Proof of Proposition 3.0.1

Notice that $\lambda_+$ is non-negative and, cf. (2.3),

$$(A, p) \mapsto \lambda_+(A, p) = \inf_n \frac{1}{n} \int \log \|A^n(\mathbf{x})\| \, d\mu(\mathbf{x}) \tag{3.4}$$

is upper-semicontinuous for the topology defined by (2.6). Thus, if $(A, p) \in \mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ is a discontinuity point for the largest Lyapunov exponent then $\lambda_+(A, p) > 0$ and there is a sequence $(A^k, p^k)_k$ converging to $(A, p)$ as $k \to \infty$ such that

$$\lim_k \lambda_+(A^k, p^k) < \lambda_+(A, p).$$

As we have seen, for each $k$ there exists some $(A^k, p^k)$-stationary measure $\eta^k$ satisfying

$$\int_{\mathcal{X}} \int_{\mathbb{P}(\mathbb{C}^2)} \log \|A^k_x v\| \, d\eta^k(v) dp^k(x) = \lambda_+(A^k, p^k).$$

Up to restricting to a subsequence, we may assume that $(\eta^k)_k$ converges in the weak* topology to some probability measure $\eta$ on $\mathbb{P}(\mathbb{C}^2)$. Then $\eta$ is an $(A, p)$-stationary measure, by Lemma 3.1.9. Using Lemma 3.1.8 we see that

$$\begin{aligned} \int_{\mathcal{X}} \int_{\mathbb{P}(\mathbb{C}^2)} \log \|A_x v\| \, d\eta(v) dp(x) &= \lim_k \int_{\mathcal{X}} \int_{\mathbb{P}(\mathbb{C}^2)} \log \|A^k_x v\| \, d\eta^k(v) dp^k(x) \\ &< \lambda_+(A, p) = \int_{\mathcal{X}} \int_{\mathbb{P}(\mathbb{C}^2)} \log \|A_x v\| \, d\eta^u(v) dp(x) \end{aligned}$$

where $\eta^u$ is the projection of $m^u$. In particular, by Lemma 3.1.7, there exists an invariant $u$-state $m \neq m^u$. It follows, using Lemma 3.1.4, that

$$m = \alpha m^u + \beta m^s \quad \text{with } \alpha + \beta = 1 \text{ and } \beta \neq 0.$$

This implies that $m^s$ is a $u$-state, because it is a linear combination of $m$ and $m^u$. Hence $m^s$ is an $su$-state. In view of Lemma 3.1.1 this means that the Oseledets subspace $E^s_{\mathbf{x}}$ is constant on a full $\mu$-measure set. Let $F^s \in \mathbb{P}(\mathbb{C}^2)$ denote this constant. Analogously, using that $(A, p)$ is a discontinuity point for the smallest Lyapunov exponent, we find $F^u \in \mathbb{P}(\mathbb{C}^2)$ such that $E^u_{\mathbf{x}} = F^u$ for $\mu$-almost every $\mathbf{x}$. It is clear that $F^u$ and $F^s$ are both invariant under $A_x$, for $p$-almost every $x \in \mathcal{X}$, because $\mu = p^{\mathbb{Z}}$. This means that there exists $\mathcal{Y} \subset \mathcal{X}$ with $p(\mathcal{Y}) = 1$ such that the linear operators defined by the $A_y$, $y \in \mathcal{Y}$ have a common eigenbasis, which is precisely the first claim in the proposition. The last claim (commutativity) is a trivial consequence. This completes the proof of Proposition 3.0.1.

## 3.2 Handling the diagonal case

Here we prove Proposition 3.0.2. Let $(A, p) \in \mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ and $\mathcal{Y}$ be as in the conclusion of Proposition 3.0.1 and consider any $p \in \mathcal{P}(\mathcal{X})$. Since conjugacies preserve the Lyapunov exponents, we may suppose $P = \mathrm{id}$ and

$$A_x = \begin{pmatrix} \theta_x & 0 \\ 0 & \theta_x^{-1} \end{pmatrix} \quad \text{for all} \quad x \in \mathcal{Y}. \tag{3.5}$$

Notice that the Lyapunov exponents of $(A, p)$ are

$$\pm \int_{\mathcal{Y}} \log |\theta_x| \, dp(x). \tag{3.6}$$

If they vanish then $(A, p)$ is automatically a continuity point, and so there is nothing to prove. Otherwise, it is no restriction to suppose

$$\int_{\mathcal{Y}} \log |\theta_x| > 0. \tag{3.7}$$

Let $V_\varepsilon$ be the $\varepsilon$-neighborhood of the horizontal direction in $\mathbb{P}(\mathbb{C}^2)$ and $\mathcal{Y}$ be as given in Proposition 3.0.1. The key step in the proof of Theorem C is the following

**Proposition 3.2.1.** *Given $\varepsilon > 0$ and $\delta > 0$ there exists $\gamma > 0$ such that if $(B, q) \in V(A, p, \gamma, \mathcal{Y})$ and there is no one-dimensional subspace invariant under all $B_x$ for $x$ in a full $q$-measure then $\eta(V_\varepsilon^c) < \delta$ for any $(B, q)$-stationary measure $\eta$.*

The proof of Proposition 3.2.1 will be given in Chapter 4. Right now, let us conclude the proof of Proposition 3.0.2.

Let $(B, q) \in \mathcal{S}(\mathcal{X}) \times \mathcal{P}(\mathcal{X})$ be close to $(A, p)$ in the sense of (2.6). First, suppose there exists some one-dimensional subspace $r \subset \mathbb{C}^2$ invariant under all the $B_x$, $x$ in a $q$-full measure. Then $r$ must be close to either the vertical axis or the horizontal axis: that is because (3.7) implies $|\theta_x| \neq 1$ for some $q$-positive measure subset. Then the Lyapunov exponent of $(B, q)$ along $r$ is close to one of the exponents (3.6). Since the other exponent is symmetric, this proves that the Lyapunov exponents of $(B, q)$ are close to the Lyapunov exponents of $(A, p)$. Now assume $B$ does not admit any invariant one-dimensional subspace. Let $M > 0$ such that $M^{-1}\|v\| < \|B_x v\| < M\|v\|$ for $p$-almost every $x \in \mathcal{X}$, all $v \in \mathbb{C}^2$ and $d(A, B) < 1$. Let $0 \ll \varepsilon \ll \delta \ll \rho \ll 1$. Let $m$ be any $u$-state realizing the largest Lyapunov exponent of $(B, q)$, and $\eta$ its projection on $P(\mathbb{C}^2)$. By Proposition 3.2.1

$$\begin{aligned}\int_{\mathbb{P}(\mathbb{C}^2)} \log \frac{\|B_x v\|}{\|v\|} d\eta([v]) &= \int_{V_\varepsilon^c} \log \frac{\|B_x v\|}{\|v\|} d\eta([v]) + \int_{V_\varepsilon} \log \frac{\|B_x v\|}{\|v\|} d\eta([v]) \\ &\geq -\delta \log M + \eta(V_\varepsilon)(\log |\theta_x| - \delta)\end{aligned}$$

for $q$-almost every $x \in \mathcal{X}$. Together with Lemma 3.1.8, this implies

$$\lambda_+(B, q) > \eta(V_\varepsilon)\lambda_+(A, p) - \delta(\log M + \eta(V_\varepsilon)) > \lambda_+(A, p) - \rho.$$

Upper semi-continuity gives $\lambda_+(B, q) \leq \lambda_+(A, p) + \rho$. Thus, we have shown that $(A, p)$ is indeed a continuity point for the Lyapunov exponents.

We have reduced the proof of Proposition 3.0.2 and Theorem C to proving Proposition 3.2.1.

# Chapter 4

# Proof of the Key Proposition

Here we give a convenient reformulation of Proposition 3.2.1 and reduce its proof to two technical estimates, Propositions 4.2.5 and 4.2.7, whose proof will be presented in the next chapter.

## 4.1 Preliminary observations

As a first step we note that under the assumptions of the proposition all stationary measures are non-atomic.

**Lemma 4.1.1.** *There exists $\gamma > 0$ such that if $(B, q) \in V(A, p, \mathcal{Y}, \gamma)$ and there is no one-dimensional subspace of $\mathbb{R}^2$ invariant under $B_x$ for every $x$ in a full $q$-measure, then every $(B, q)$-stationary measure is non-atomic.*

*Proof.* By assumption, $A$ is diagonal and the Lyapunov exponents do not vanish. So, we may take $\gamma > 0$ so that if $(B, q) \in V(A, p, \mathcal{Y}, \gamma)$ then $B_x$ is hyperbolic and its eigenspaces are close to the horizontal and vertical directions, for every $x$ is some set $\mathcal{Z} \subset \mathcal{X}$ with $q(\mathcal{Z}) > 0$. Then any finite set of one-dimensional subspaces invariant under any $B_x$, $x \in \mathcal{Z}$ has at most two elements. Moreover, they must coincide with the eigenspaces of $B_x$ and, consequently, are fixed under $B_x$. Since we assume there is no one-dimensional subspace fixed by $B_x$ for $\mu$-almost every $x$, it follows that there is no finite set of one-dimensional subspaces invariant under $B_x$ for $\mu$-almost every $x$.

Now let us suppose $\eta$ has some atom. Let $z_1, \ldots, z_N$ be the atoms with the largest mass, say, $\eta(\{z_i\}) = a$ for $i = 1, \ldots, N$. Since $\eta$ is a stationary measure,

$$\eta\big(\{B_x^{-1}(z_1), \ldots, B_x^{-1}(z_N)\}\big) = \eta\big(\{z_1, \ldots, z_N\}\big) = Na$$

for $q$-almost every $x \in \mathcal{X}$. Moreover, in view of the previous paragraph, we have $\{B_x^{-1}(z_1), \ldots, B_x^{-1}(z_N)\} \neq \{z_1, \ldots, z_N\}$ for a positive $q$-measure subset of points $x$. This implies that there exists $z \neq z_i$ for $i = 1, \ldots, N$ such that $\eta(\{z\}) = a$. That contradicts the choice of the $z_i$ and so the lemma is proved. $\square$

Let $\phi : \mathbb{P}(\mathbb{C}^2) \to \mathbb{C}^2 \cup \{\infty\}$, $\phi([z_1, z_2]) = z_1/z_2$ be the standard identification between the complex projective space and the Riemann sphere. Then the projective action of a linear map

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

corresponds to a Möbius transformation on the sphere

$$\hat{B} : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\} \quad \hat{B}(z) = \frac{az+b}{cz+d},$$

in the sense that $\phi \circ B = \hat{B} \circ \phi$. It follows that a measure $\xi$ in projective space is $(B, q)$-stationary if and only if the measure $\eta = \phi_* \xi$ on the sphere satisfies $\eta = \int (\hat{B_x})_* \eta \, dq(x)$. Then the measure $\eta$ is also said to be $(B, q)$-stationary. Clearly, $\eta$ is non-atomic if and only if $\xi$ is.

This means that the key Proposition 3.2.1 may be restated as

**Proposition 4.1.2.** *Given $\varepsilon > 0$ and $\delta > 0$ there exist $\gamma > 0$ such that if $(B, q) \in V(A, p, \gamma, \mathcal{Y})$ and $q(\{x \in \mathcal{X} : \hat{B}_x(z) = z\}) < 1$ for all $z \in \mathbb{C} \cup \{\infty\}$ then*

$$\eta(B(0, \varepsilon^{-1})) < \delta$$

*for any $(B, q)$-stationary probability measure $\eta$ on $\mathbb{C} \cup \{\infty\}$.*

Let us give a brief outline of the proof. Complete arguments will appear in the next chapter.

Clearly, there are infinitely many $(A, p)$-stationary measures, namely, all convex combinations of the Dirac masses $\delta_0$ and $\delta_\infty$ corresponding, respectively, to the vertical direction and the horizontal direction. The point with the proof is that we need to show that stationary measures of nearby cocycles approach the one $(A, q)$-stationary measure, $\delta_\infty$, that realizes the Lyapunov exponent $\lambda_+(A, p)$. The crucial property that singles out $\delta_\infty$ among all $(A, q)$-stationary measures is the fact that it is an attractor for the dynamics

$$f_A : \eta \mapsto \int (A_x)_* \eta \, dp(x) \tag{4.1}$$

induced by $A$ in the space of the probability measures of $\mathbb{P}(\mathbb{C}^2)$. In particular, $\delta_\infty$ can be recovered as the limit of forward iterates under $f_A$ of any Dirac mass other than $\delta_0$. Now, let $(B, q)$ be close to $(A, p)$. In view of what we just said, one may expect $f_B$ to also have an attractor, strongly concentrated near $\infty$. Moreover, assuming there is no one-dimensional subspace invariant under $q$-almost every $B_x$, one may expect this attractor to draw the forward iterates of every Dirac mass. In particular, every fixed point $\eta$ of the operator $f_B$ should be strongly concentrated near $\infty$. This is precisely the contents of the proposition.

The first step in the argument is to estimate the measure of a corona $\mathfrak{C} = B(0, \epsilon^{-1}) \setminus B(0, r)$ for small $\varepsilon > 0$ and $r > 0$. Note that $\mathfrak{C}$ has zero weight for any $(A, p)$-stationary measure: indeed, the $f_A$-iterates of any measure eventually leave $\mathfrak{C}$. This is no longer true for nearby cocycles, but we are able to conclude that any $(B, q)$-stationary measure gives small weight to the corona: given any $\varepsilon > 0$, we have

$$\eta\big(B(0, \epsilon^{-1}) \setminus B(0, r_0)\big) < \varepsilon$$

for any $(B, q)$ close to $(A, p)$ and any $(B, q)$-stationary measure $\eta$, where $r_0 = r_0(B)$ goes to zero when $(B, q)$ approaches $(A, p)$.

The neighborhoods $B(0, r)$ require a different argument is because of the "souvenir" of the repeller $\delta_0$ of $f_A$ imprinted in the dynamics of $f_B$. The hypothesis that $B$ has no subspace that is fixed by $q$-almost every $B_x$ plays an important role at this stage. In the special case when the probability $p$ has finite support plays a key role. In the special case when the measure $p$ has finite support we use it to find $r_1 = r_1(B) > 0$ and $h$ in the support such that $B_h^{-1}(B(0, r_1)$ is disjoint from $B(0, r_1(B))$. This allows us to obtain

$$\eta(B(0, r_1)) < \text{const}\, \eta(B(0, \epsilon^{-1}) \setminus B(0, r_1))$$

and then, using the previous stage, to conclude that $\eta(B(0, r_1))$ is small. In the general case, the argument is a bit more complicated because we need to consider the possible existence of a subset of points $x \in M$, with $\mu$-measure close to 1, and such that the corresponding $B_x$ have a common fixed point. The two alternatives that can arise here are handled by Proposition 4.2.5 and Corollary 4.2.6.

In general, $r_1 < r_0$ and so the previous two inequalities do not quite solve our problem. However, the quotient $r_0/r_1$ is bounded by some constant that depends on $A$ only. This fact allows us to show that

$$\eta(B(0, r_0) \setminus B(0, r_1)) \leq \text{const}\, \eta(B(0, \epsilon^{-1}) \setminus B(0, r_0))$$

and so the term on the left is small. This finishes our outline of the proof. In the sequel we fill-in the details of the argument.

## 4.2 Proof of Proposition 4.1.2

For simplicity, we assume that $\mathcal{Y}$ given in the proof of Proposition 3.0.1 is equals to $\mathcal{X}$, because as $p(\mathcal{X} \setminus \mathcal{Y})$ as $q(\mathcal{X} \setminus \mathcal{Y})$ is equals to zero for all $q$ such that $(B, q) \in V(A, p, \varepsilon, \mathcal{Y})$ for all $\varepsilon > 0$.

Recall, from (3.5) and (3.7), that

$$A_x = \begin{pmatrix} \theta_x & 0 \\ 0 & \theta_x^{-1} \end{pmatrix} \quad \text{with} \quad \int \log |\theta_x| dp(x) > 0 \tag{4.2}$$

Let $B$, $q$, and $\eta$ be as in the statement of Proposition 4.1.2.

**Lemma 4.2.1.** *There exist $\beta, \sigma \in (0, 1)$, positive numbers $(\sigma_x)_{x \in \mathcal{X}}$, integers $(s_x)_{x \in \mathcal{X}}$ and $k \in \mathbb{N}$ satisfying:*

*(a)* $0 < \|A\|^{-1}/4 \leq \sigma_x \leq \beta|\theta_x|$ *for all* $x \in \mathcal{X}$

*(b)* $\sigma_x = \sigma^{s_x}$ *for all* $x \in \mathcal{X}$

*(c)* $\int \log \sigma_x \, dp(x) > 4/k$.

*Proof.* Fix $k \in \mathbb{N}$ large enough so that $\int \log|\theta_x| \, dp(x) > 7/k$. Define $\log \beta = \log \sigma = -1/k$
For each $x \in \mathcal{X}$, define

$$r_x = \left[k \log|\theta_x|\right], \quad s_x = \begin{cases} r_x - 1 & \text{if } r_x \neq 1 \\ r_x - 2 & \text{if } r_x = 1 \end{cases} \qquad \log \sigma_x = -\frac{s_x}{k}.$$

Properties (a) and (b) follow immediately. Moreover,

$$\int \log \sigma_x dp(x) \geq \int \log|\theta_x| - 3/k \, dp(x) > 4/k$$

as claimed in (c). ∎

In what follows, let $\sigma$, $\beta$, $\sigma_x$ and $s_x$, as in Lemma 4.2.1. We partition $\mathcal{X} = \mathcal{X}_- \cup \mathcal{X}_+$, where $\mathcal{X}_-$ be the subset of $x \in \mathcal{X}$ with $s_x < 0$ (i.e. $\sigma_x > 1$) and $\mathcal{X}_+$ is the subset of $x \in \mathcal{X}$ with $s_x > 0$ (i.e. $\sigma_x < 1$). For each $x \in \mathcal{X}$, let

$$D_x = \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_x^{-1} \end{pmatrix} \quad \text{and} \quad \hat{D}_x(z) = \sigma_x^2 z. \tag{4.3}$$

Define another matrix and its associated Möbius transformation

$$D_{sp} = \begin{pmatrix} \sigma^\tau & 0 \\ 0 & \sigma^{-\tau} \end{pmatrix} \quad \text{and} \quad \hat{D_{sp}}(z) = \sigma^{2\tau} z. \tag{4.4}$$

where $\tau$ is the smallest natural such that $\sigma^\tau \leq \|A\|^{-1}/4$. For each $\mathcal{K} \subset \mathcal{X}$ let $K$ be the cocycle defined by

$$K_x = \begin{pmatrix} k_x & 0 \\ 0 & k_x^{-1} \end{pmatrix} \quad \text{where} \quad k_x = \begin{cases} \sigma_x & \text{if } x \in \mathcal{K} \\ \sigma^\tau & \text{if } x \in \mathcal{X} \setminus \mathcal{K} \end{cases} \tag{4.5}$$

**Lemma 4.2.2.** *There exists* $\alpha > 0$ *such that*

$$\int \log k_x \, dp(x) \geq 2/k$$

*for any measurable set* $\mathcal{K}$ *with* $p(\mathcal{K}) \geq 1 - \alpha$.

*Proof.* Taking $\alpha = (-k \log \sigma^\tau)^{-1}$, we have

$$\begin{aligned} \int \log k_x \, dp &\geq \int \log \sigma_x \, dp - \int_{\mathcal{X}\setminus\mathcal{K}} \log \sigma_x \, dp + \int_{\mathcal{X}\setminus\mathcal{K}} \log \sigma^\tau \, dp \\ &\geq 4/k + 2 \log \sigma^\tau p(\mathcal{X} \setminus \mathcal{K}) \geq 2/k. \end{aligned}$$

∎

For $z_0 \in \mathbb{C}$ and $r > 0$, we denote $B(z_0, r) = \{z \in \mathbb{C} : |z - z_0| \le r\}$. Given $C, B \in \mathcal{S}(\mathcal{X})$ and $\mathcal{Y} \subset \mathcal{X}$ we say that $r > 0$ is $(B, \mathcal{Y})$-*centered* with respect to $C$ if

$$\hat{B}_x^{-1}(B(0,r)) \subset \hat{C}_x^{-1}(B(0,r)) \quad \text{for every } x \in \mathcal{Y}. \tag{4.6}$$

When $\mathcal{Y} = \mathcal{X}$ we say only $B$-*centered* with respect to $C$. Given $C, B \in \mathcal{S}(\mathcal{X})$, $q \in \mathcal{P}(\mathcal{X})$, and a $(B, q)$-stationary measure $\eta$, we say that $r > 0$ is $(B, q, \eta)$-*pseudo-centered* with respect to $C$ if

$$\int \eta\big(\hat{B}_x^{-1}(B(0,r))\big)\, dq(x) \le \int \eta\big(\hat{C}_x^{-1}(B(0,r))\big)\, dq(x) \tag{4.7}$$

**Remark 4.2.3.** If $r > 0$ is $B$-centered (respectively, $(B, q, \eta)$-pseudo-centered) with respect to $D$ then it is also $B$-centered (respectively, $(B, q, \eta)$-pseudo-centered) with respect to the cocycle $K$ defined in (4.5), because $\hat{D}_x^{-1}(B(0,r)) \subset \hat{D_{sp}}^{-1}(B(0,r))$ for any $x \in \mathcal{X}$.

**Lemma 4.2.4.** *Given $\rho > 0$ there is $\gamma > 0$ such every $r \in [\rho, \rho^{-1}]$ is $B$-centered with respect to $D$ for every $B \in \mathcal{S}(\mathcal{X})$ with $d(A, B) < \gamma$.*

*Proof.* By assumption, $\pm \log |\theta_x|$, $x \in \mathcal{X}$ is bounded. Write

$$B_x^{-1} = \begin{pmatrix} a_x & b_x \\ c_x & d_x \end{pmatrix}.$$

The condition $d(A, B) < \gamma$ implies that $|a_x - \theta_x^{-1}|$, $|b_x|$, $|c_x|$, $|d_x - \theta_x|$ are all less than $c_1\gamma$ for some constant $c_1$ independent of $x$ and $\gamma$. Given $\rho > 0$, assume first that $\gamma \le \rho^2$. Then, for any $|z| \in [\rho, \rho^{-1}]$,

$$|\hat{B}_x^{-1}(z)| \le \frac{|a_x z| + |b_x|}{|d_x| - |c_x z|} \le \frac{|a_x| + c_1\sqrt{\gamma}}{|d_x| - c_1\sqrt{\gamma}}|z| \le \frac{|\theta_x^{-1}|}{|\theta_x|}\frac{1 + c_2\sqrt{\gamma}}{1 - c_2\sqrt{\gamma}}|z|$$

where $c_2$ is also independent of $x$ and $\gamma$. Thus, there exists $\gamma_0 > 0$ independent of $x \in \mathcal{X}$ such that,

$$|\hat{B}_x^{-1}(z)| \le (\beta|\theta_x|)^{-2}|z| \le \sigma_x^{-2}|z| = |\hat{D}_x^{-1}(z)|$$

for every $x \in \mathcal{X}$ and every $|z| \in [\rho, \rho^{-1}]$, as long as $\gamma \le \gamma_0$. This gives that every $r \in [\rho, \rho^{-1}]$ is $B$-centered with respect to $D$, as claimed. ☐

The proof of Proposition 4.1.2 relies on a couple of technical propositions that we state in the sequel and whose proofs we postpone to Chapter 5. The first one gives a bound on the mass of the stationary measure away from the vertical (and the horizontal) direction:

**Proposition 4.2.5.** *Given $\varepsilon > 0$ and $\delta > 0$ there exists $\gamma > 0$ such that if $d(A, B) < \gamma$ and $d(p, q) < \gamma$ then*

$$\eta\big(B(0, \varepsilon^{-1}) \setminus B(0, r_0)\big) < \delta$$

*for any $(B, q)$-stationary measure $\eta$ and $r_0 < 1$ such that every $r \in [r_0, \varepsilon^{-1}]$ is $(B, \mathcal{K})$-centered with respect to $D$ for some measurable set $\mathcal{K}$ satisfying $p(\mathcal{X} \setminus \mathcal{K}) \le \alpha$, where $\alpha$ is like in the Lemma 4.2.2.*

**Corollary 4.2.6.** *Given $\varepsilon > 0$ and $\delta > 0$ there exist $\gamma > 0$ such that if $d(A,B) < \gamma$ and $d(p,q) < \gamma$ then either $\eta\big(B(0,\varepsilon^{-1})\big) \leq \delta$ or there exist $0 < r_0 < 1$ such that*

$$\eta\big(B(0,\varepsilon^{-1}) \setminus B(0,r_0)\big) \leq \delta$$

*and $p(\{x \in \mathcal{X};\ \hat{B}_x^{-1}(B(0,r_0)) \not\subset \hat{D}_x^{-1}(B(0,r_0))\}) \geq \alpha$.*

*Proof.* Let $r_1$ be the infimum of $0 < r < 1$ such that $\eta\big(B(0,\epsilon^{-1}) \setminus B(0,r)\big) < \delta$. If $r_1 = 0$ then $\eta\big(B(0,\epsilon^{-1}) \setminus \{0\}\big) \leq \delta$. Since $\eta$ has no atoms, by Lemma 4.1.1, it follows that $\eta(B(0,\epsilon^{-1})) \leq \delta$. Now, assume that $r_1 > 0$. Then, $\eta(B(0,\epsilon^{-1}) \setminus B(0,r_1)) \geq \delta$ and thus, by Proposition 4.2.5, $\mathfrak{F}_1 = \{x \in \mathcal{X};\ \hat{B}_x^{-1}(B(0,r_1)) \not\subset \hat{D}_x^{-1}(B(0,r_1))\}$ has $p$-measure greater than $\alpha$. Define for $k = 2, 3, \ldots$

$$\mathfrak{F}_k = \{x \in \mathcal{X};\ \hat{B}_x^{-1}(B(0,r_k)) \not\subset \hat{D}_x^{-1}(B(0,r_k))\},$$

where $1 > r_2 > r_3 > \ldots$ is a decreasing sequence converging to $r_1$. Note that $\mathfrak{F}_k \subset \mathfrak{F}_{k+1}$ for all $k \geq 2$ and $\cup_{k=2}^{\infty} \mathfrak{F}_k = \mathfrak{F}_1$. Thus $p(\mathfrak{F}_k)$ converges to $p(\mathfrak{F}_1)$ and this implies that there is $N \in \{2, 3, \ldots\}$ such that $p(\mathfrak{F}_N) \geq \alpha$. The proof is complete, taking $r_0 = r_N$. $\square$

The next proposition, together with the series of lemmas that follow, lead to a bound on the mass of the stationary measure close to the vertical direction:

**Proposition 4.2.7.** *There is $\gamma > 0$ and $N \in \mathbb{N}$ such that if $d(A,B) < \gamma$, $u \in [0,1]$ and $x \in \mathcal{X}$ are such that $\hat{B}_x^{-1}(B(0,u)) \not\subset \hat{D}_x^{-1}(B(0,u))$ then*

$$\mathcal{D} \cap \hat{B}_x^{-1}(\mathcal{D}) = \emptyset, \quad \text{where } \mathcal{D} = \begin{cases} \hat{D}_x^{-N}(B(0,u)) & \text{if } x \in \mathcal{X}_- \\ \hat{D}_x^{N}(B(0,u)) & \text{if } x \in \mathcal{X}_+. \end{cases}$$

*In particular, $B(0,\sigma^{N\tau}u) \cap \hat{B}_x^{-1}(B(0,\sigma^{N\tau}u)) = \emptyset$.*

Now, since $\lambda_+(A,p) > 0$ there exist $\alpha_0, \rho_0 > 0$ such that if we define $X_0 = \{x \in \mathcal{X}/|\theta_x| > 1 + \rho_0\}$ then $p(X_0) \geq \alpha_0$.

**Lemma 4.2.8.** *There exists $c = c(A)$ and $\gamma > 0$ such that if $\rho < c^{-1}$, $x \in X_0$, $d(A,B) < \gamma$ and $\hat{B}_x$ has a fixed point in $B(0,\rho)$ then*

$$\hat{B}_x^{-1}(B(0,r)) \subset \hat{D}_x^{-1}(B(0,r)) \quad \text{for all} \quad r \in [c\rho, 1]. \tag{4.8}$$

*Proof.* First, take $\gamma > 0$ such that $\hat{B}_x^{-1}$ is a $\lambda_x$-contraction with

$$\frac{\|A\|^{-1}}{2} = \lambda \leq \lambda_x \leq (1+\rho_0)^{-2}$$

and $\hat{D}_x^{-1}(z) = \Lambda_x z$ with $\lambda_x \leq \beta\Lambda_x$. Now, take $c \in \mathbb{N}$ large enough such that $\lambda^{-1} < c(\beta^{-1} - 1 - c^{-1})$. So, $|\hat{B}_x^{-1}(z)| \leq [c^{-1} + \lambda_x(1 + c^{-1})]r \leq \Lambda_x r$. $\square$

Define $\Gamma(z,\rho)=\{x\in X_0 : \hat{B}_x \text{ has a fixed point in } B(z,\rho)\}$ for each $z\in\mathbb{C}$ and $1>\rho>0$ and let $\beta_0=\alpha_0/(1+4c^2\sigma^{-4\tau N})$.

**Lemma 4.2.9.** *If $d(A,B)<\gamma$ and $p(\{x\in X_0 :\ \hat{B}_x(z)=z\})<\beta_0$ for all $z\in B(0,1)$ then for each $\varsigma>0$ there exist $z_0\in B(0,1)$ and $\rho>0$ such that*

*(a) $p(\Gamma(z,\rho))\le p(\Gamma(z_0,\rho))+\varsigma$ for all $z\in B(0,1)$;*

*(b) $\frac{\beta_0}{4}\le p(\Gamma(z_0,\rho))\le\beta_0$*

*Proof.* We begin observing that there is $r>0$ such that $p(\Gamma(z,r))<\beta_0$ for all $z\in B(0,1)$. For this, suppose for contradiction that for each $n\in\mathbb{N}$ there exists $z_n$ such that $p(\Gamma(z_n,\frac{1}{n}))>\beta_0$. By compactness we may suppose $z_n\to z_0$. So, for any $\xi>0$, $p(\Gamma(z_0,\xi))>\beta_0$ and, consequently, $p(\{x\in X_0 :\ \hat{B}_x(z_0)=z_0\})\ge\beta_0$, this contradicts the hypothesis. Now, take $0<\varrho=\inf\{r>0 : p(\Gamma(z,r))\ge\beta_0$ for some $z\in B(0,1)\}$. By the choice of $\varrho$, $p(\Gamma(z,\rho))<\beta_0$ for all $z\in B(0,1)$ and $\rho<\varrho$. However, there is $z_1$ and $\rho<\varrho$ such that $p(\Gamma(z_1,\rho))\ge\beta_0/4$. Note that if this is false then $p(\Gamma(z,(1.1)\varrho))<\beta_0$ for all $z\in B(0,1)$, because we may cover $\Gamma(z,(1.1)\varrho)$ with four sets $\Gamma(z,(0.9)\rho)$ and this contradicts the choice of $\rho$. Fixe some $\rho$ (for instance, $\rho=0.9\varrho$) with this property. Taking $l=\sup\{p(\Gamma(z,\rho)) : z\in B(0,1)\}$, we have for each $\varsigma>0$ there is $z_0\in B(0,1)$ such that $\beta_0/4\le l\le p(\Gamma(z_0,\rho))+\varsigma$. This proves the lemma. □

**Remark 4.2.10.** In the conditions of Lemma 4.2.9 if we take $\varsigma$ small enough then $p(\Gamma(z_0,\rho))\ge\beta_0/4$ and $p(X_0\setminus\Gamma(z_0,c\sigma^{-2\tau N}\rho))\ge\beta_0/2$, because it is easy to see that there is $4c^2\sigma^{-4\tau N}$ subsets of the form $\Gamma(z,\rho)$ covering $\Gamma(z_0,c\sigma^{-2\tau N}\rho)$.

**Lemma 4.2.11.** *There exists $0<\lambda_0<1$ such that if $d(A,B)<\gamma$, $x\in\Gamma(0,\rho)$ and $1>r\ge c\rho$ then $\hat{B}_x^{-1}(B(0,r))\subset B(0,\lambda_0 r)$. In particular, there exists $\kappa>0$ such that if $1>r\ge c\sigma^{-\tau}\rho$ then $\hat{B}_x^{-\kappa}(B(0,r))\subset B(0,\sigma^{2\tau}r)$*

*Proof.* Take $\gamma>0$ sufficiently small such that $\hat{B}_x^{-1}$ is a $\lambda$-contraction for some $\lambda\in(0,1)$ in the ball $B(0,1)$ for all $x\in X_0$. Now, let $z_0\in B(0,1)$ the unique fix point of $\hat{B}_x^{-1}$ and take $\lambda_0=\lambda(1+c^{-1})+c^{-1}$. So, we have

$$|\hat{B}_x^{-1}(z)|\le\rho+\lambda|z-z_0|\le[c^{-1}+\lambda(1+c^{-1})]r$$

for all $z$ with $|z|=r\ge c\rho$. To complete the proof it is enough to take $\kappa$ as the smallest natural such that $\lambda_0^\kappa\le\sigma^{2\tau}$. □

Assuming Propositions 4.2.5 and 4.2.7 for a while, we can give the

*Proof of Proposition 4.1.2.* Suppose first that $z_0=0$ ($z_0$ and $\rho>0$ given by Lemma 4.2.9). Define $X_1=X_0\setminus\Gamma(0,c\sigma^{-2\tau N}\rho)$. Then, by Remark 4.2.10, $p(X_1)\ge\beta_0/2$ and $p(\Gamma(0,\rho))\ge\beta_0/4$. Let $r_0>0$ be as in Corollary 4.2.6 and take $u=\max\{r_0,c\sigma^{-2\tau N}\rho\}$. Define $Y=\{x\in$

$\mathcal{X};\ \hat{B}_x^{\ -1}(B(0,u))\not\subset\hat{D}_x^{\ -1}(B(0,u))\}$. Then, using Lemma 4.2.9 and Corollary 4.2.6, we get $p(Y)\geq\beta_1=\min\{\alpha,\beta_0/2\}$ and

$$\eta(B(0,\varepsilon^{-1})\setminus B(0,u))<\delta \tag{4.9}$$

Note also that, since $d(p,q)$ is close to zero, $q(Y)>\beta_1/2$ and $q(\Gamma(0,\rho))>\beta_0/8$.

Since $\eta$ is stationary,

$$\int_{\mathcal{X}}\eta(B(0,u))\,dq(x)=\eta(B(0,u))=\int_{\mathcal{X}}\eta(\hat{B}_x^{\ -1}(B(0,u)))\,dq(x).$$

This, together with Lemma 4.2.11, implies

$$\begin{aligned}
q(\Gamma(0,\rho))&\eta\big(B(0,u)\setminus(B(0,\lambda_0 u))\big)\\
&\leq\int_{\Gamma(0,\rho)}\eta\big(B(0,u)\setminus\hat{B}_x^{-1}(B(0,u))\big)\,dq(x)\\
&=\int_{\mathcal{X}\setminus\Gamma(0,\rho)}\big(\eta(\hat{B}_x^{-1}(B(0,u)))-\eta(B(0,u))\big)\,dq(x)\\
&\leq\int_{\mathcal{X}\setminus\Gamma(0,\rho)}\eta\big(\hat{B}_x^{-1}(B(0,u)\setminus B(0,u)\big)\,dq(x)\\
&\leq\eta\big(B(0,\varepsilon^{-1})\setminus B(0,u)\big)\leq\delta.
\end{aligned}$$

Consequently, using (4.9) once more,

$$\eta\big(B(0,\varepsilon^{-1})\setminus B(0,\lambda_0 u)\big)\leq\delta(1+8\beta_0^{-1})$$

Arguing by induction we get that

$$\eta\big(B(0,\varepsilon^{-1})\setminus B(0,\lambda_0^j u)\big)\leq\delta(1+8\beta_0^{-1})^j\quad\text{for every } j\geq 0,\quad\text{with } \lambda_0^j\geq\sigma^{2\tau N}.$$

In particular,

$$\eta\big(B(0,\varepsilon^{-1})\setminus(B(0,\lambda_0^{\kappa N}u)\big)\leq\delta(1+8\beta_0^{-1})^{\kappa N}. \tag{4.10}$$

Combining Lemma 4.2.11 and (4.10), we get that

$$\eta\big(B(0,\varepsilon^{-1})\setminus B(0,\sigma^{2\tau N}u)\big)\leq\delta(1+8\beta_0^{-1})^{\kappa N}. \tag{4.11}$$

Similarly,

$$\int\eta(B(0,\sigma^{2\tau N}u))\,dq(x)=\eta(B(0,\sigma^{2\tau N}u))=\int\eta\big(\hat{B}_x^{\ -1}(B(0,\sigma^{2\tau N}u))\big)\,dq(x)$$

together with (4.9), (4.11) and Proposition 4.2.7, implies that

$$\begin{aligned}
q(Y)\eta(B(0,\sigma^{2\tau N}u))&\leq\int_Y\eta(\hat{B}_x^{-1}(B(0,\sigma^{2\tau N}u)))\,dq(x)\\
&\quad+\int_{\mathcal{X}\setminus Y}\eta\big(\hat{B}_x^{-1}(B(0,\sigma^{2\tau N}u))\setminus B(0,\sigma^{2\tau N}u)\big)\,dq(x)\\
&\leq\eta\big(B(0,\varepsilon^{-1})\setminus B(0,\sigma^{2\tau N}u)\big)\\
&\leq\delta(1+8\beta_0^{-1})^{\kappa N}.
\end{aligned}\tag{4.12}$$

Adding (4.11) and (4.12), we get

$$\eta(B(0,u)) \leq \delta(1+4\beta_1^{-1})(1+8\beta_0^{-1})^{\kappa N} \leq \delta\tilde{c} \tag{4.13}$$

where $\tilde{c}>0$ is some upper bound for $\delta(1+4\beta_1^{-1})(1+8\beta_0^{-1})^{\kappa N}$. Adding (4.9) and (4.13), we obtain

$$\eta(B(0,\varepsilon^{-1})) \leq (1+\tilde{c})\delta.$$

That completes the proof of the proposition in this case. When $p(\{x \in X_0;\ \hat{B}_x(0)=0\}) \geq \beta_0$ the proof is analogous, in fact, it is simpler.

Now we treat the general case. Since $A_x$ is diagonal for all $x \in X_0$, with bigger eigenvalue far from $\{z \in \mathbb{C} : |z|=1\}$ corresponding to the horizontal direction. In particular, $z_0$ given by the Lemma 4.2.9 is near to $z=0$. So, the direction corresponding to $z_0$ is close to the horizontal direction. Let $(a_0,b_0) \approx (1,0)$ be a unitary vector in the direction of $z_0$. Define

$$H = \begin{pmatrix} a_0 & -b_0 \\ b_0 & a_0 \end{pmatrix}$$

Then, the cocycle $C_x = HB_xH^{-1}$ satisfies the hypothesis of the first case. Therefore, if $\eta$ is $(B,q)$-stationary then $H_*\eta$ is $(C,q)$-stationary. Using the previous particular case and the fact that $H$ is close to the identity,

$$\eta(B(0,\varepsilon^{-1})) \leq H_*\eta(B(0,2\varepsilon^{-1})) \leq \delta(1+\tilde{c}).$$

This finishes the proof of the proposition. □

# Chapter 5

# Main estimates

In this chapter we prove Propositions 4.2.5 and 4.2.7.

## 5.1 Mass away from the vertical

Here we prove Proposition 4.2.5. Let $\sigma < 1$ be as in Lemma 4.2.1. For each $\mathcal{K} \subset \mathcal{X}$ consider the cocycle $K$ associated, as defined in (4.5). Clearly,

$$\hat{K}_x(B(0, r\sigma^{2j})) = B(0, r\sigma^{2j+2s_x}) \quad \text{for } r > 0, x \in \mathcal{X}, \text{ and } j \in \mathbb{Z}. \tag{5.1}$$

Define

$$I_j(r) = B(0, r\sigma^{2j-2}) \setminus B(0, r\sigma^{2j}) \tag{5.2}$$

for $j \in \mathbb{Z}$ and

$$L_x(r) = \begin{cases} B(0,r) \setminus \hat{K}_x^{-1}(B(0,r)) & \text{for } x \in \mathcal{X}_- \\ \hat{K}_x^{-1}(B(0,r)) \setminus B(0,r) & \text{for } x \in \mathcal{X}_+. \end{cases} \tag{5.3}$$

Note that the partition $\mathcal{X} = \mathcal{X}_- \cup \mathcal{X}_+$ depends on the cocycle $K$. Moreover, $\mathcal{X} \setminus \mathcal{K} \subset \mathcal{X}_+$, because $k_x = \sigma^\tau$ for all $x \in \mathcal{X} \setminus \mathcal{K}$.

**Lemma 5.1.1.** *If $r$ is $(B, q, \eta)$-pseudo-centered with respect to $K$ then*

1. $\int_{\mathcal{X}_-} \eta(L_x(r))\, dq(x) \le \int_{\mathcal{X}_+} \eta(L_x(r))\, dq(x)$

2. $\int_{\mathcal{X}_-} \sum_{j=1}^{-s_x} \eta(I_j(r))\, dq(x) \le \int_{\mathcal{X}_+} \sum_{j=0}^{s_x-1} \eta(I_j(r))\, dq(x)$.

*More generally, if $r\sigma^{2t}$ is $(B, q, \eta)$-pseudo-centered with respect to $K$ for all $t = 0, 1, \ldots, n$ then*

$$\int_{\mathcal{X}_-} \sum_{j=t+1}^{t-s_x} \eta(I_j(r))\, dq(x) \le \int_{\mathcal{X}_+} \sum_{j=t}^{t+s_x-1} \eta(I_j(r))\, dq(x),$$

*for all $t = 0, 1, \ldots, n$.*

*Proof.* Let $J = B(0,r)$. Using that $r$ is $(B,q,\eta)$-centered and $\eta$ is $(B,q)$-stationary

$$\int \left(\eta(J) - \eta(\hat{K_x}^{-1}(J))\right)\, dq(x) \leq \int \left(\eta(J) - \eta(\hat{B_x}^{-1}(J))\right)\, dq(x) = 0.$$

By definition (5.3), the left hand side coincides with

$$\int_{\mathcal{X}_-} \eta(L_x(r))\, dq(x) - \int_{\mathcal{X}_+} \eta(L_x(r))\, dq(x).$$

This proves the first claim. The second one is a direct consequence: just note that, by (5.1),

$$L_x(r) = \begin{cases} B(0,r) \setminus B(0, r\sigma^{-2s_x}) = \bigsqcup_{j=1}^{-s_x} I_j(r) & \text{for } x \in \mathcal{X}_- \\ B(0, r\sigma^{-2s_x}) \setminus B(0,r) = \bigsqcup_{j=-s_x+1}^{0} I_j(r) & \text{for } x \in \mathcal{X}_+. \end{cases}$$

The last claim follows, noticing $I_j(r\sigma^{2t}) = I_{j+t}(r)$ for all $j$ and $r$. □

**Remark 5.1.2.** If in the previous lemma we replace $(B,q,\eta)$-pseudo-centered by $B$-centered then the result follows for every $\eta$ $(B,q)$-stationary measure.

**Lemma 5.1.3.** *There exists $\gamma > 0$ such that if $d(A,B) < \gamma$, $r \in [0,1]$ and $x \in \mathcal{X}$ is such that $\hat{B_x}^{-1}(B(0,r)) \cap B(0,r) \neq \emptyset$ then $\hat{B_x}^{-1}(B(0,r)) \cup B(0,r) \subset \hat{D_{sp}}^{-1}(B(0,r))$ for all $x \in \mathcal{X}$.*

*Proof.* Take $\gamma > 0$ such that if $d(A,B) < \gamma$ then for all $x \in \mathcal{X}$ the diameter of $\hat{B_x}^{-1}(B(0,r))$ is less than $3\|A\|^2 r$, for all $r \in [0,1]$. So, if $\hat{B_x}^{-1}(B(0,r)) \cap B(0,r) \neq \emptyset$ then $\hat{B_x}^{-1}(B(0,r)) \cup B(0,r) \subset B(0, 4\|A\|^2 r) \subset \hat{D_{sp}}^{-1}(B(0,r))$. □

We also need the following calculus result. In the application, for proving Proposition 4.2.5, we will take $n_i = |s_i|$ and $a_j = \eta(I_j(r))$.

**Lemma 5.1.4.** *Let $(n_x)_{x\in\mathcal{X}}$ be a bounded family positive integers and $(a_j)_{j\in\mathbb{Z}}$ be a sequence of non-negative real numbers. Assume*

*(a)* $0 < S \leq \int_{\mathcal{X}_-} n_x\, dq(x) - \int_{\mathcal{X}_+} n_x\, dq(x)$ *and*

*(b)* $\int_{\mathcal{X}_-} \sum_{j=t+1}^{t+n_x} a_j\, dq(x) \leq \int_{\mathcal{X}_+} \sum_{j=t-n_x+1}^{t} a_j\, dq(x)$ *for* $t = 0, \dots, n$.

*Denote* $n_- = \sup\{n_x : x \in \mathcal{X}_-\}$ *and* $n_+ = \sup\{n_x : x \in \mathcal{X}_+\}$. *Then*

$$\sum_{j=1}^{n} a_j \leq \left(\frac{n_- + n_+}{S}\right) \sum_{j=-n_+ +1}^{0} a_j.$$

*Proof.* Begin by noting that

$$\sum_{t=0}^{n}\sum_{j=t+1}^{t+n_x} a_j = \sum_{l=1}^{n_x}\sum_{j=l}^{n+l} a_j \geq n_x\left(\sum_{j=1}^{n} a_j - \sum_{j=1}^{n_x} a_j\right) \tag{5.4}$$

and that

$$\sum_{t=0}^{n}\sum_{j=t-n_x+1}^{t} a_j = \sum_{l=-n_x+1}^{0}\sum_{j=l}^{n+l} a_j \leq n_x\left(\sum_{j=1}^{n} a_j + \sum_{j=-n_x+1}^{0} a_j\right) \tag{5.5}$$

So, adding the inequalities (b) over all $t = 0, \dots, n$ and using (5.4)-(5.5),

$$\int_{\mathcal{X}_-} n_x \left[\sum_{j=1}^{n} a_j - \sum_{j=1}^{n_x} a_j\right] dq(x) \leq \int_{\mathcal{X}_+} n_x \left[\sum_{j=1}^{n} a_j + \sum_{j=-n_x+1}^{0} a_j\right] dq(x)$$

or, equivalently,

$$S\sum_{j=1}^{n} a_j \leq \int_{\mathcal{X}_-} n_i \sum_{j=1}^{n_x} a_j\, dq(x) + \int_{\mathcal{X}_+} n_x \sum_{j=-n_x+1}^{0} a_j\, dq(x)$$

This implies, using the inequality (b) once more,

$$\begin{aligned} S\sum_{j=1}^{n} a_j &\leq n_- \int_{\mathcal{X}_-} \sum_{j=1}^{n_x} a_j\, dq(x) + n_+ \int_{\mathcal{X}_+} \sum_{j=-n_x+1}^{0} a_j\, dq(x) \\ &\leq (n_- + n_+) \int_{\mathcal{X}_+} \sum_{j=-n_x+1}^{0} a_j\, dq(x). \end{aligned}$$

This last expression is bounded above by $(n_- + n_+)\sum_{j=-n_+ +1}^{0} a_j$. In this way we get the conclusion of the lemma. $\square$

Define $\alpha_s = \sum_{j=(s-1)n_+ +1}^{sn_+} a_j$ for each $s \geq 0$. In the same setting as Lemma 5.1.4, we obtain

**Corollary 5.1.5.** *Let $n = s_0 n_+$ for some integer $s_0 \geq 1$. Then there exists $s \in \{1, \dots, s_0\}$ such that*

$$\alpha_s \leq (\frac{n_- + n_+}{s_0 S})\alpha_0.$$

*Proof.* The conclusion of Lemma 5.1.4 may be rewritten

$$\sum_{s=1}^{s_0} \alpha_j = \sum_{j=1}^{n} a_j \leq (\frac{n_- + n_+}{S}) \sum_{j=-n_+ +1}^{0} a_j = (\frac{n_- + n_+}{S})\alpha_0\,.$$

This implies that $\min_{1\leq s\leq s_0} \alpha_j \leq (n_- + n_+)\alpha_0/(s_0 S)$, as claimed. $\square$

*Proof of Proposition 4.2.5.* This will follow from applying Lemma 5.1.4 and Corollary 5.1.5 to appropriate data.

Let $\mathcal{K}$ such that $p(\mathcal{K}) \geq 1 - \alpha$ and replace $D_x$ by $D_{sp}$ for all $x \in \mathcal{X} \setminus \mathcal{K}$. Consequently, replace $s_x$ by $\tau$ and $\sigma_x$ by $\sigma^\tau$ for every $x \in \mathcal{X} \setminus \mathcal{K}$. Note that this way $\mathcal{K} \setminus \mathcal{X} \subset \mathcal{X}_+$. Take $n_x = |s_x|$ for $x \in \mathcal{X}$. Define $S(p) = \int_{\mathcal{X}_-} n_x \, dp(x) - \int_{\mathcal{X}_+} n_x \, dp(x)$. By construction (Lemma 4.2.1) and Lemma 4.2.2,

$$\begin{aligned} -S(p) \log \sigma &= \int_{\mathcal{X}} s_x \log \sigma \, dp(x) \\ &= \int_{\mathcal{K}} \log \sigma^{s_x} \, dp(x) + \sigma^\tau p(\mathcal{X} \setminus \mathcal{K}) > 2/k > 0. \end{aligned}$$

It follows that $S(p) > 0$ and, consequently, there exist $\gamma > 0$ and $S > 0$ such that $S(q) > S$ for every $q$ with $d(p, q) < \gamma$. This corresponds to condition (a) in Lemma 5.1.4.

Given $\varepsilon > 0$ and $\delta > 0$, let $n = s_0 n_+ = s_0 \tau$ for some integer $s_0$ satisfying

$$s_0 \geq (\frac{n_- + n_+}{S})^2 \delta^{-1}$$

and fix also $R > \sigma^{-2n} \varepsilon^{-1}$. By Lemma 4.2.4, there exists $\gamma > 0$ such that if $d(A, B) < \gamma$ for all then every $r \in [(R\sigma^{-2})^{-1}, R\sigma^{-2}]$ is $B$-centered with respect to $D$. In particular, this applies to $y\sigma^{2j}$ for every $j = 0, 1, \ldots, n$ and any $y \in [R, R\sigma^{-2}]$, since these points are in $[\varepsilon^{-1}, R\sigma^{-2}] \subset [(R\sigma^{-2})^{-1}, R\sigma^{-2}]$. Fix $y \in [R, R\sigma^{-2}]$ and define $a_j = \eta(I_j(y))$ for $j \in \mathbb{Z}$. Then Lemma 5.1.1 gives

$$\int_{\mathcal{X}_-} \sum_{j=t+1}^{t+n_x} a_j \, dq(x) \leq \int_{\mathcal{X}_+} \sum_{j=t-n_x+1}^{t} a_j \, dq(x)$$

for all $t = 0, \ldots, n$. This corresponds to condition (b) in Lemma 5.1.4.

Therefore, we are in a position to apply Corollary 5.1.5: we conclude that there exists $s \in \{1, \ldots, n\}$ such that

$$\alpha_s \leq (\frac{n_- + n_+}{s_0 S}) \alpha_0 \leq (\frac{S}{n_- + n_+}) \delta \alpha_0 \leq (\frac{S}{n_- + n_+}) \delta \tag{5.6}$$

Notice that, by definition,

$$\alpha_s = \eta\big(B(0, y\sigma^{2(s-1)n_+}) \setminus B(0, y\sigma^{2sn_+})\big)$$

Let $r_1$ the infimum of $\tilde{r} \in [r_0, 1]\}$ such that all $r \in [\tilde{r}, 1]$ is $(B, q, \eta)$-centered with respect to $K$ and take $y = r_1 \sigma^{-2\tilde{n}}$ for some $\tilde{n} \in \mathbb{N}$. We claim that $\eta(B(0, \varepsilon^{-1})) \setminus B(0, r_1) < \delta$. In fact, $y\sigma^{2t}$ is $(B, q, \eta)$-pseudo-centered for every $t = 0, 1, \ldots, \tilde{n}$. The other two conditions in Lemma 5.1.4 are also satisfied in this context: (a) is just the same as before and (b) follows from Lemma 4.2.4 in the same way as in above. So, from Lemma 5.1.4 we conclude that

$$\sum_{j=1}^{\tilde{n}} a_j \leq (\frac{n_- + n_+}{S}) \sum_{j=-n_+ + 1}^{0} a_j. \tag{5.7}$$

The left hand side coincides with

$$\eta\big(B(0,y)\setminus B(0,y\sigma^{2\tilde{n}})\big)\geq \eta\big(B(0,\varepsilon^{-1})\setminus B(0,r_0)\big).$$

Moreover, the sum on the right hand side of (5.7) coincides with

$$\eta\big(B(0,y\sigma^{-2sn_+})\setminus B(0,y)\big)=\alpha_s.$$

Hence, (5.6) and (5.7) yield $\eta\big(B(0,\varepsilon^{-1})\setminus B(0,r_1)\big)\leq\delta$ and this proves the proposition when $r_1=r_0$. If $r_1>r_0$ then there exists $x\in\mathcal{X}\setminus\mathcal{K}$ such that $\eta(\hat{B_x}^{-1}B(0,r_1))\geq\hat{D_{sp}}^{-1}B(0,r_1))$. So, Lemma 5.1.3 implies that

$$B(0,r_1)\cap\hat{B_x}^{-1}(B(0,r_1))=\emptyset$$

and

$$\eta(B(0,r_1))\leq\eta(\hat{D_{sp}}^{-1}B(0,r_1)))\leq\eta(\hat{B_x}^{-1}B(0,r_1))$$

Therefore $\eta(B(0,r_1))<\delta$. Consequently, $\eta(B(0,\varepsilon^{-1}))<2\delta$ and this completes the proof of proposition. □

## 5.2 Estimates close to the vertical

We prove Proposition 4.2.7. We begin with a couple of auxiliary lemmas. A Möbius transformation $h$ is a $\gamma_0$-*deformation* of $f(z)=\lambda z$ if $h(z)=(az+b)/(cz+d)$ with

$$\max\big\{\big||a|-|\lambda|\big|,\big|b\big|,\big|c\big|,\big||d|-1\big|\big\}<\gamma_0|\lambda|.$$

**Lemma 5.2.1.** *Given $\beta_0,\sigma_0\in(0,1)$ there exist $\gamma_0>0$ and $N_0\in\mathbb{N}$ such that for any $f(z)=\lambda z$ and $g(z)=\Lambda z$ with $|\lambda|\leq\beta_0|\Lambda|$ and $|\Lambda|\leq\sigma_0$, and for any $\gamma_0$-deformation $\tilde{f}$ of $f$ if $r\in[0,1]$ and $\tilde{f}(B(0,r))\not\subset g(B(0,r))$ then*

$$\tilde{f}(g^{N_0}(B(0,r)))\cap g^{N_0}(B(0,r))=\emptyset$$

*Proof.* Fix $N_0\in\mathbb{N}$ such that

$$|\Lambda|^{N_0-1}\leq\sigma_0^{N_0-1}\leq\frac{1-\beta_0}{100}\tag{5.8}$$

and $\gamma_0>0$ given by

$$\gamma_0=\frac{1-\beta_0}{100}<\frac{1}{100}\tag{5.9}$$

Write $\tilde{f}(z)=(az+b)/(cz+d)$. If $\tilde{f}(0)=0$ then $b=0$ and (5.9) gives

$$|\tilde{f}(z)|\leq\frac{|az|}{|d|-|c|}\leq\frac{1+\gamma_0}{1-2\gamma_0}|\lambda z|\leq\beta_0^{-1}|\lambda z|\leq|\Lambda||z|$$

for every $|z| \leq 1$. This means $\tilde{f}(B(0,r)) \subset g(B(0,r))$ for all $r \leq 1$ and we have nothing to do.

Next, suppose $\tilde{f}(0) \neq 0$ or, equivalently, $b \neq 0$. Take

$$r_0 = \frac{10|b|}{|\Lambda|(1-\beta_0)}.$$

Then $|\tilde{f}(z)| \leq |\Lambda z|$ for every $|z| \in [r_0, 1]$. Indeed,

$$|\tilde{f}(z)| \leq \frac{|az|+|b|}{|d|-|c|} \leq \frac{|\lambda|(1+\gamma_0)+|\Lambda|(1-|\beta_0|)/10}{1-2\gamma_0}|z|$$

and, in view of (5.9), the right hand side is bounded by

$$\frac{\beta_0(1+\gamma_0)+10\gamma_0}{1-2\gamma_0}|\Lambda z| \leq \frac{\beta_0+20\gamma_0}{1-2\gamma_0}|\Lambda z| \leq |\Lambda z|.$$

This gives $\tilde{f}(B(0,r)) \subset g(B(0,r))$ for $r \in [r_0, 1]$. Moreover, if $s \in [0, r_0]$, by (5.8),

$$|\Lambda|^N s \leq |\Lambda|^N r_0 \leq \frac{|\Lambda|(1-\beta_0)}{100}\frac{10|b|}{|\Lambda|(1-\beta_0)} \leq \frac{|b|}{10} \leq \frac{|b|}{5|d|}$$

and that means that

$$g^N(B(0,s)) \subset B\big(0, \frac{|b|}{5|d|}\big). \tag{5.10}$$

The relation (5.9) also leads to

$$|\tilde{f}'(z)| \leq \frac{|ad|+|bc|}{(|d|-|cz|)^2} \leq \frac{|\lambda|(1+\gamma_0)^2+(\gamma_0|\lambda|)^2}{(1-2\gamma_0)^2}$$

for all $|z| \leq 1$. Hence, using (5.9) once more,

$$|\tilde{f}'(z)| \leq \frac{1+4\gamma_0}{1-4\gamma_0}|\lambda| \leq \beta_0^{-1}|\lambda| \leq |\Lambda| \leq 1$$

That implies

$$\tilde{f}\big(B\big(0, \frac{|b|}{5|d|}\big)\big) \subset B\big(\frac{b}{d}, \frac{|b|}{5|d|}\big) \tag{5.11}$$

From (5.10) and (5.11) we get that $g^N(B(0,s)) \cap \tilde{f}(g^N(B(0,s))) = \emptyset$, which completes the proof of the lemma. $\square$

**Lemma 5.2.2.** *Given $\beta_0, \sigma_0 \in (0,1)$ there exist $\gamma_0 > 0$ and $N_0 \in \mathbb{N}$ such that for any $f(z) = \lambda z$ and $g(z) = \Lambda z$ with $|\lambda| \leq \beta_0|\Lambda|$ and $|\lambda| \leq \sigma_0$, and for any $\gamma_0$-deformation $\tilde{g}$ of $g$, if $r \in [0,1]$ and $\tilde{g}^{-1}(B(0,r)) \not\subset f^{-1}(B(0,r))$ then*

$$f^{N_0}(B(0,r)) \cap \tilde{g}^{-1}(f^{N_0}(B(0,r))) = \emptyset$$

*Proof.* Begin by noting that the hypothesis $\tilde g^{-1}(B(0,r))\not\subset f^{-1}(B(0,r))$ is equivalent to $f(f^{-1}(B(0,r)))\not\subset \tilde g f^{-1}(B(0,r))$. So, there is no loss of generality in supposing that $f(B(0,r))\not\subset\tilde g(B(0,r))$. Fix $N_0 \geq 1$ such that

$$|\lambda|^{N_0-1} \leq \sigma_0^{N_0-1} \leq \frac{1-\beta_0}{12} \leq \frac{1-|\lambda|/|\Lambda|}{12}. \tag{5.12}$$

Fix $\gamma > 0$ such that

$$\frac{1+2\gamma_0}{1-2\gamma_0} \leq \beta_0^{-1} \tag{5.13}$$

Write $\tilde g(z) = (az+b)/(cz+d)$. If $\tilde g(0) = 0$ then $b = 0$ and so

$$|\tilde g(z)| \geq \frac{|a||z|}{|d|+|c|} \geq \frac{1+\gamma_0}{1-2\gamma_0}|\Lambda||z| \geq |\lambda||z|$$

for every $|z| \leq 1$. This implies $f(B(0,r)) \subset \tilde g(B(0,r))$ for all $r \leq 1$ and we have nothing to do.

Next, suppose $\tilde g(0) \neq 0$. We have

$$|\tilde g(z)| \geq \frac{|az|-|b|}{|d|+|cz|}| \geq |\lambda||z| \quad \text{whenever } w_- \leq |z| \leq w_+$$

where $w_-$ and $w_+$ are the solutions of $|\lambda c|w^2 + (|\lambda d| - |a|)w + |b| = 0$. A direct calculation shows that

$$w_- \leq \frac{2|b|}{|a|-|\lambda d|} \quad \text{and} \quad w_+ \geq 1$$

if $\gamma$ is small enough. This gives $\tilde g(B(0,r)) \supset f(B(0,r))$ for $r \in [r_0, 1]$ with

$$r_0 = \frac{2|b|}{|a|-|\lambda d|}.$$

Notice that $(1-|\lambda|/|\Lambda|) \leq 2(1-|\lambda d|/|a|)$ if $\gamma$ is small enough. Then (5.12) gives

$$|\lambda|^N r_0 \leq \frac{|\Lambda|-|\lambda|}{12|\Lambda|}\frac{2|b\lambda|}{|a|-|\lambda||d|} \leq \frac{|b\lambda|}{3|a|}$$

and that means that

$$f^N(B(0,r_0)) \subset B\big(0,\frac{|b\lambda|}{3|a|}\big) \subset B\big(0,\frac{|b|}{3|a|}\big) \tag{5.14}$$

On the other hand,

$$|(\tilde g^{-1})'(z)| \leq \frac{|ad|+|bc|}{(|cz|-|a|)^2} \leq \lambda^{-1}$$

for all $|z| \leq 1$, as long as $\gamma$ is small enough, and that implies

$$\tilde g^{-1}\big(B\big(0,\frac{|b\lambda|}{3|a|}\big)\big) \subset B\big(-\frac{b}{a},\frac{|b|}{3|a|}\big) \tag{5.15}$$

From (5.14) e (5.15) we conclude that $f^N(B(0,r_0)) \cap \tilde g^{-1}(f^N(B(0,r_0))) = \emptyset$. With greater reason $f^N(B(0,s)) \cap \tilde g^{-1}(f^N(B(0,s))) = \emptyset$ for every $s \in [0, r_0]$ and this completes the proof of the lemma. $\square$

*Proof of Proposition 4.2.7.* Note that if $d(A,B)<\gamma$ then every $\hat{B}_x^{-1}$ is a $(C\gamma)$-deformation of $f=\hat{A}_x^{-1}$, where the constant $C=\sup_{x\in\mathcal{X}}|\theta_x|$ depends only on $A$. Indeed,

$$B_x=\begin{pmatrix} a_x & b_x \\ c_x & d_x \end{pmatrix} \quad \text{yields} \quad \hat{B}_x^{-1}=\frac{d_x\theta_x^{-1}z-b_x\theta_x^{-1}}{-c_x\theta_x^{-1}z+a_x\theta_x^{-1}}$$

and then $\|A_x-B_x\|<\gamma$ implies

$$\left|d_x\theta_x^{-1}-\theta_x^{-2}\right|,\ \left|b_x\theta_x^{-1}\right|,\ \left|c_x\theta_x^{-1}\right|,\ \left|a_x\theta_x^{-1}-1\right|\leq\gamma|\theta_x|^{-1}\leq C\gamma|\theta_x|^{-2}.$$

Take $f=\hat{A}_x{}^{-1}$ and $g=\hat{D}_x{}^{-1}$ for each $x\in\mathcal{X}_-$. Notice $f(z)=|\theta_x|^{-2}|z|$ and $g(z)=\sigma_x^{-2}|z|$. Since $\sigma_x\leq\beta|\theta_x|$ and $\sigma_x\geq\sigma^{-1}$ (cf. Lemma 4.2.1), we are in the setting of Lemma 5.2.1, with $\beta_0=\beta^2$ and $\sigma_0=\sigma^2$. From the lemma, and the observation in the previous paragraph, we get that there exists $\gamma_->0$ and $N_-\in\mathbb{N}$ such that if $d(A,B)<\gamma_-$, $x\in\mathcal{Y}_-$ and $r\in[0,1]$ are such that $\hat{B}_x{}^{-1}(B(0,r))\not\subset\hat{D}_x{}^{-1}(B(0,r))$ then $\hat{B}_x{}^{-1}(\hat{D}_x{}^{-N_-}(B(0,r)))\cap\hat{D}_x{}^{-N_-}(B(0,r))=\emptyset$.

Now take $f=\hat{D}_x$ and $g=\hat{A}_x$ for $x\in\mathcal{X}_+$. Then $f(z)=\sigma_x^2|z|$ and $g(z)=|\theta_x|^2|z|$ and so we are in the setting of Lemma 5.2.2, with $\beta_0=\beta^2$ and $\sigma_0=\sigma^2$. In this way we get $\gamma_+>0$ and $N_+\in\mathbb{N}$ such that if $d(A,B)<\gamma_+$, $x\in\mathcal{Y}_+$ and $r\in[0,1]$ are such that $\hat{B}_x{}^{-1}(B(0,r))\not\subset\hat{D}_x{}^{-1}(B(0,r))$ then

$$\hat{B}_x{}^{-1}(\hat{D}_x{}^{N_+}(B(0,r)))\cap\hat{D}_x{}^{N_+}(B(0,r))=\emptyset.$$

Take $\gamma=\min\{\gamma_-,\gamma_+\}$ and $N=\max\{N_-,N_+\}$. This completes the proof of the proposition. $\square$

# Chapter 6

# Consequences of Theorem C

In this chapter we deduce Theorem B and Theorem D.

## 6.1 Proof of Theorem B

Now we deduce Theorem B from Theorem C. The main step is contained in the lemma that follows. Let $\lambda$ denote Lebesgue measure on the unit interval $I$. We use $\|\eta\|$ to denote the total variation of a signed measure $\eta$.

**Lemma 6.1.1** (Avila). *Let $\nu$ be a Borel probability measure with compact support in some metric space $X$ such that all bounded and closed subset of $X$ is compact. For every $\varepsilon > 0$ there exists $\delta > 0$ and a weak$^*$ neighborhood $V$ of $\nu$ such that every probability measure $\mu \in V$ whose support is contained in $B_\delta(\operatorname{supp}\nu)$ may be written as*

$$\phi_* q = \mu$$

*for some probability measure $q$ in $\operatorname{supp}\nu \times I$ and some measurable map $\phi : \operatorname{supp}\nu \times I \to X$ with $\|q - (\nu \times \lambda)\| < \varepsilon$ and $d(\phi(x,t), x) < \varepsilon$ for all $x \in \operatorname{supp}\nu$.*

*Proof.* Denote $Z = \operatorname{supp}\nu$. First, we claim that for any $\delta > 0$ there exist a cover of $B_\delta(Z)$ by disjoint sets $Q_i$, $i = 1, \ldots, N$ with $\nu(Q_i) > 0$, $\nu(\partial Q_i) = 0$, and $\operatorname{diam} Q_i < 12\delta$. This can be seen as follows.

Firstly, take for each $x \in Z$, $r_x \in (\delta, 2\delta)$ such that $\nu(\partial B(x, r_x)) = 0$. So, $\{B(x, r_x) : x \in Z\}$ is a cover of the compact set, $\overline{B_\delta(Z)}$. Take $\{V_1, V_2, \ldots, V_k\}$ a finite subcover of $\overline{B_\delta(Z)}$ and consider the associated partition $\mathcal{P}$ of $\cup_{i=1}^k V_i$, whose atoms are the sets $P = V_1^* \cap \cdots \cap V_k^*$, where $V_i^*$ is equals to $V_i$ or $V_i^c$. Note that, by construction, $\operatorname{diam} V_i < 4\delta$, $\nu(V_i) > 0$ and $\nu(\partial V_i) = 0$. From $\mathcal{P}$, we are going to obtain a partition $\mathcal{Q}$ with the property in our claim. To this end, define

$$B_1 = V_1 \cup \{P \in \mathcal{P} : \nu(P) = 0 \quad \text{and} \quad P \subset V_i \quad \text{with} \quad V_i \cap V_1 \neq \emptyset\}$$

Now, if $V_2 \subset B_1$ then we define $B_2 = \emptyset$. Otherwise, $\nu(V_2 \setminus B_1) > 0$ and then we define

$$B_2 = V_2 \cup \{P \in \mathcal{P} : \nu(P) = 0 \quad \text{and} \quad P \subset V_i \quad \text{with} \quad V_i \cap V_2 \neq \emptyset\} \setminus B_1$$

Now we repeat this procedure inductively: for each $l \geq 2$, we define $B_l = \emptyset$ if $V_l \subset \cup_{i=1}^{l-1} B_i$; otherwise, $\nu(V_l \setminus \cup_{i=1}^{l-1} B_i) > 0$ and then we define

$$B_l = V_l \cup \{P \in \mathcal{P} : \nu(P) = 0 \text{ and } P \subset V_i \text{ with } V_i \cap V_l \neq \emptyset\} \setminus D_{l-1}$$

where $D_{l-1} = \cup_{i=1}^{l-1} B_i$. The claim follows by taking as $Q_i$ the non-empty sets $B_i$.

Proceeding with the proof of the lemma, take $\delta = \varepsilon/12$ and assume the neighborhood $V$ is small enough that

$$\sum_{i=1}^{N} |\mu(Q_i) - \nu(Q_i)| < \varepsilon \quad \text{for every } \mu \in V.$$

Let $Z_i = \operatorname{supp} \nu \cap Q_i$ for each $i = 1, \ldots, N$. Clearly, $\nu(Z_i) = \nu(Q_i)$. Let $q$ be the measure on $Z \times I$ that coincides with

$$\frac{\mu(Q_i)}{\nu(Q_i)}(\nu \times \lambda)$$

restricted to each $Z_i \times I$. For each $i$, let $a_{i,j}$, $j \in J(i)$ be the atoms of $\nu'$ contained in $Q_{i,j}$ (note that $J(i)$ may be empty). Moreover, let $I_{i,j}$, $j \in J_i$ be disjoint subsets of $I$ such that

$$\lambda(I_{i,j}) = \frac{p_{i,j}}{\mu(Q_i)} \quad \text{for all } j \in J_i,$$

where $p_{i,j} = \nu'(a_{i,j})$. Denote by $I_i$ the complement of the union of all $I_{i,j}$, $j \in J_i$ inside $I$. Then

$$q\big(Z_i \times I_i\big) = \mu(Q_i) - \sum_{j \in J_i} p_{i,j} = \mu\big(Q_i \setminus \cup_{j \in J_i} a_{i,j}\big).$$

Since all non-atomic Lebesgue probability spaces are equivalent to the unit interval endowed with Lebesgue measure (see [34]), the previous equality ensures that there exists an invertible measurable map

$$\phi_i : Z_i \times I_i \to Q_i \setminus \cup_{j \in J_i} a_{i,j}$$

mapping the restriction of $q$ to the restriction of $\mu$. By setting $\phi \equiv a_{i,j}$ on each $Z_i \times I_{i,j}$ we extend $\phi_i$ to a measurable map $Z_i \times I \to Q_i$ that still sends the restriction of $q$ to the restriction of $\mu$. Gluing all these extensions we obtain a measurable map $\phi : Z \times I \to X$ such that $\phi_* q = \mu$. By construction, $\phi(x,t) \in Q_i$ for every $x \in Z_i$ and $t \in I$. This implies that $d(\phi(x,t), x) \leq \operatorname{diam} Q_i < \varepsilon$ for all $(x,t) \in Z \times I$. Finally,

$$\begin{aligned} \|q - (\nu \times \lambda)\| &= \sum_{i=1}^{n} \Big\| \Big( \frac{\mu(Q_i)}{\nu(Q_i)} - 1 \Big) (\nu \times \lambda) \mid (Z_i \times I) \Big\| \\ &= \sum_{i=1}^{n} |\mu(Q_i) - \nu(Q_i)| < \varepsilon. \end{aligned}$$

The proof of the lemma is complete $\square$

Now Theorem B can be obtained as follows. Given $\rho > 0$, let $\nu$ be a probability measure in $\mathrm{GL}(2,\mathbb{C})$ with compact support. Consider $\mathcal{X} = \operatorname{supp}\nu \times I$, $p = \nu \times \lambda$ and $A : \mathcal{X} \to \mathrm{GL}(2,\mathbb{C})$ given by $A(x,t) = x$. From Theorem C, there is $\varepsilon > 0$ such that $|\lambda_\pm(A,p) - \lambda_\pm(B,q)| < \rho$ for all $(B,q)$ such that $d(p,q) < \varepsilon$ and $d(A,B) < \varepsilon$. On the other hand, Lemma 6.1.1 implies that there exist a weak* neighborhood $V$ and $\delta$ such that if $\nu' \in V$ and $\operatorname{supp}\nu' \subset B_\delta(\operatorname{supp}\nu)$ then there exist $B : \mathcal{X} \to \mathrm{GL}(2,\mathbb{C})$ and a probability measure $q$ on $\mathcal{X}$ such that $d(p,q) < \varepsilon$, $d(A,B) < \varepsilon$ and $\nu' = B_* q$. Noting that $\lambda_\pm(\nu) = \lambda_\pm(A,p)$ and $\lambda_\pm(\nu') = \lambda_\pm(B,q)$, we get Theorem B.

## 6.2 Proof of Theorem D

We need the following proposition, whose proof we postpone for a while.

**Proposition 6.2.1.** *Suppose that $\lambda_+(A,p) > 0$ and let $m^u$ be its a canonical u-state. If $(A^k,p)_k$ converges to $(A,p)$ and $m^u_k$ is a canonical u-state for $(A^k,p)$ for each $k$ then $(m^u_k)_k$ converges to $m^u$ weakly*.*

To prove Theorem D, it is enough to show that

$$\mu(\{x \in M : \angle(E^u_{A,x}, E^u_{A^k,x}) < \epsilon\})$$

converges to 1 when $k$ goes to $\infty$. Let $(A^k,p)_k$ converge to $(A,p)$, and let $(m^{k,u})_k$ and $m^u$ be the canonical $u$-states for $(A^k,p)$ and $(A,p)$, respectively, for $k \geq 1$. By using Proposition 6.2.1, we have that $(m^{k,u})_k$ converges to $m^u$. Note that $\psi : M \ni \mathbf{x} \mapsto E^u_{A,\mathbf{x}}$ is a measurable map and its graphic has full $m^u$-measure. Given $\varepsilon > 0$, by the theorem of Lusin, see Theorem 1 in Loeb [28], we may take a compact set $K \subset M$ such that the map $\psi_K : K \ni \mathbf{x} \mapsto E^u_{A,\mathbf{x}}$ is continuous and its graphic has $m^u$-measure greater than $1-\varepsilon$. Now, given $\delta > 0$, take a open neighborhood $V$ of the graphic of $\psi_K$, such that the diameter of $V \cap \{\mathbf{x}\} \times \mathbb{P}(\mathbb{C}^2)$ is less than $\delta$ for all $\mathbf{x} \in K$, that is, $V \cap K \times \mathbb{P}(\mathbb{C}^2) \subset V_\delta := \{(\mathbf{x},\xi) \in K \times \mathbb{P}(\mathbb{C}^2) : d(\psi(\mathbf{x}),\xi) < \delta\}$, where $d$ stands for a distance on the projective $\mathbb{P}(\mathbb{C}^2)$. From the weak* convergence, we have that

$$\liminf m^{k,u}(V) \geq m^u(V) \geq 1 - \varepsilon.$$

On the other hand, we have that $m^{k,u}(K \times \mathbb{P}(\mathbb{C}^2)) = \mu(K) \geq 1-\varepsilon$ for all $k$. Thus, $m^{k,u}(V \cap K \times \mathbb{P}(\mathbb{C}^2)) \geq 1 - 2\varepsilon$ and consequently $m^{k,u}(V_\delta) \geq 1 - 3\varepsilon$ for all $k \geq k_0$ for some $k_0$. Nevertheless, $m^{k,u}(V_\delta) = \mu(\{\mathbf{x} \in K : d(E^u_{A_k,\mathbf{x}}, E^u_{A,\mathbf{x}}) < \delta\})$ and, therefore, $\mu(\{\mathbf{x} \in M : d(E^u_{A_k,\mathbf{x}}, E^u_{A,\mathbf{x}}) < \delta\}) \geq 1 - 3\varepsilon$, for all $k \geq k_0$. Since $\delta$ and $\varepsilon$ are arbitrary, this proves Theorem D.

To prove Proposition 6.2.1, we begin by showing that the space $\mathcal{M}(p)$ of the probability measure that project down to $\mu = p^{\mathbb{Z}}$ is compact. More precisely,

**Lemma 6.2.2.** *Let $(\mu^k)_k$ converge to $\mu$ in the weak* topology of $\mathcal{P}(\mathcal{X}^{\mathbb{Z}})$. Let $(m^k)_k$ be a sequence of probabilities in $\mathcal{P}(\mathcal{X}^{\mathbb{Z}} \times \mathbb{P}(\mathbb{C}^2))$ projecting down to $(\mu_k)_k$. Then there exists a* *subsequence of $(m^k)_k$ converging in the weak$^*$ topology to some probability $m$ on $\mathcal{X}^{\mathbb{Z}} \times \mathbb{P}(\mathbb{C}^2)$ that projects down to $\mu$.*

*In particular, the space of probabilities measures on $\mathcal{P}(\mathcal{X}^{\mathbb{Z}} \times \mathbb{P}(\mathbb{C}^2))$ that project down to $\mu$ is compact for the weak$^*$ topology.*

*Proof.* The key result we use is Prohorov's theorem, see, for instance, Billingsley [4]. We begin by noting that $\mathcal{X}^{\mathbb{Z}}$ and $\mathcal{X}^{\mathbb{Z}} \times \mathbb{P}(\mathbb{C}^2)$ are polish spaces. Therefore, the sequence of probabilities $(\xi^k)_k$ on any of these two spaces have converging subsequences, in the weak$^*$ sense, if and only if for each $\varepsilon > 0$ there is a compact set $K_\varepsilon$ such that $\xi^k(K_\varepsilon) > 1 - \varepsilon$. So, it is enough to prove that for any $\varepsilon > 0$ there exists a compact set $\hat{K}_\varepsilon \subset \mathcal{X}^{\mathbb{Z}} \times \mathbb{P}(\mathbb{C}^2)$ such that $m^k(\hat{K}) > 1 - \varepsilon$ for each $k \geq 1$. From Prohorov's theorem, we have that for any $\varepsilon > 0$, there is a compact set $K_\varepsilon \subset \mathcal{X}^{\mathbb{Z}}$ such that $\mu^k(K_\varepsilon) > 1 - \varepsilon$, for any $k \geq 1$. However, $\hat{K}_\varepsilon = K_\varepsilon \times \mathbb{P}(\mathbb{C}^2)$ is compact and also, $m^k(\hat{K}_\varepsilon) = \mu^k(K_\varepsilon) > 1 - \varepsilon$ for all $k \geq 1$. The lemma thus follows from an application of Prohorov's theorem. $\square$

*Proof of Proposition 6.2.1.* Let $(m^k)_k$ a sequence of probability measures such that $m^k$ project down to $\mu$ for all $k$. We claim that it is enough to show that if $(m^k)_k$ converges to $m$, then $((F_{A^k})_* m^k)_k$ converges to $(F_A)_* m$. In fact, the claim implies that all limit point $m$ of the sequence $(m^u_k)_k$ is $F_A$-invariant, because $m^u_k$ is $F_{A^k}$ for all $k$. Moreover, by Theorem C and Lemma 3.1.5,

$$\lim \lambda_+(A^k, p) = \lim \int \phi_{A^k}\, dm^u_k = \int \phi_A\, dm^u = \lambda_+(A, p)$$

So, using Remark 3.1.6, we conclude that $m = m^u$. Furthermore, using Lemma 6.2.2, we conclude that $(m^u_k)_k$ converges to $m^u$. To finish, let us prove the claim. To this end, given $\varepsilon > 0$, by the theorem of Lusin, there is a compact set $K \subset M$ such that $\mu(K) > 1-\varepsilon$ and the transformation $A : M \to \mathrm{SL}(2, \mathbb{C})$ is continuous when restricted to $K$. So, if $\varphi : M \times \mathbb{P}(\mathbb{C}^2) \to \mathbb{R}$ is a bounded and uniformly continuous function (this is enough to characterize weakly$^*$ convergence, see, for instance, Billingsley [4]), the function $\varphi \circ F_A : M \times \mathbb{P}(\mathbb{C}^2) \to \mathbb{R}$ is also continuous when restricted to $K \times \mathbb{P}(\mathbb{C}^2)$. Now, using the theorem of extension of Titze (see Kelley [22]), take $\tilde{\varphi} : M \times \mathbb{P}(\mathbb{C}^2) \to \mathbb{R}$ a continuous extension of the restriction of $\varphi \circ F_A$ to $K \times \mathbb{P}(\mathbb{C}^2)$ such that $\|\tilde{\varphi}\| \leq 2\|\varphi\|$. We have

$$\begin{aligned}
|\int \varphi\, d(F_{A^k})_* m^k - \int \varphi\, d(F_A)_* m| &= |\int \varphi \circ F_{A^k}\, dm^k - \int \varphi \circ F_A\, dm| \\
\leq |\int \varphi \circ F_{A^k}\, dm^k - \int \varphi \circ F_A\, dm^k| &+ |\int \varphi \circ F_A\, dm^k - \int \varphi \circ F_A\, dm|
\end{aligned}$$

The first term converges to zero, because $\varphi \circ F_{A^k} - \varphi \circ F_A$ converges to zero uniformly. Moreover, using triangle inequality and the fact that $m^k$ and $m$ project down to $\mu$, we see that the second one is bounded by the sum $|\int \tilde{\varphi}\, dm^k - \int \tilde{\varphi}\, dm|$ with $6\|\varphi\|\varepsilon$. Since $\varepsilon > 0$ is taken arbitrary and $\tilde{\varphi}$ is a bounded and continuous function, we conclude that $\int \varphi\, d(F_{A^k})_* m^k$ converges to $\int \varphi\, d(F_A)_* m$, for all $\varphi$ bounded and continuous function, that is, $(F_{A^k})_* m^k$ converges weakly$^*$ to $(F_A)_* m$. This completes the proof of the proposition. $\square$

# Chapter 7

# Further considerations

## 7.1 An example of discontinuity

In this chapter we describe a construction of points of discontinuity of the Lyapunov exponents as functions of the cocycle, relative to some Hölder topology. This builds on and refines [5, 6, 8, 29], where it is shown that Lyapunov exponents are often discontinuous relative to the $C^0$ topology.

Let $M = \Sigma_2$ be the shift with 2 symbols, endowed with the metric $d(\mathbf{x}, \mathbf{y}) = 2^{-N(\mathbf{x},\mathbf{y})}$, where

$$N(\mathbf{x}, \mathbf{y}) = \sup\{n \geq 0 : x_n = y_n \text{ whenever } |n| < N\}.$$

For any $r \in (0, \infty)$, the $C^r$ norm in the space of $r$-Hölder continuous functions $L : M \to \mathcal{L}(\mathbb{C}^d, \mathbb{C}^d)$ is defined by

$$\|L\|_r = \sup_{\mathbf{x}\in M} \|L(\mathbf{x})\| + \sup_{\mathbf{x}\neq\mathbf{y}} \frac{\|L(\mathbf{x}) - L(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r}.$$

Consider on $M$ the Bernoulli measure associated to any probability vector $(p_1, p_2)$ with positive entries and $p_1 \neq p_2$. Given any $\sigma > 1$, consider the (locally constant) cocycle $A : M \to \mathrm{SL}(2, \mathbb{R})$ defined by

$$A(\mathbf{x}) = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \quad \text{if } x_0 = 1$$

and

$$A(\mathbf{x}) = \begin{pmatrix} \sigma^{-1} & 0 \\ 0 & \sigma \end{pmatrix} \quad \text{if } x_0 = 2.$$

**Theorem 7.1.1.** *For any $r > 0$ such that $2^{2r} < \sigma$ there exist cocycles $B : M \to \mathrm{SL}(2, \mathbb{R})$ with vanishing Lyapunov exponents and such that $\|A - B\|_r$ is arbitrarily close to zero.*

Since the Lyapunov exponents $\lambda_\pm(A) = \pm|p_1 - p_2| \log \sigma$ of the cocycle $A : M \to \mathrm{SL}(2, \mathbb{R})$ are non-zero, this gives that $A$ is a point of discontinuity for the Lyapunov exponents relative to the $C^r$ topology.

The proof of Theorem 7.1.1 is an adaptation of ideas of Knill [25] and Bochi [5, 6]. Here is an outline. Notice that the unperturbed cocycle $A$ preserves both the horizontal line bundle $H_{\mathbf{x}} = \{\mathbf{x}\} \times \mathbb{R}(1,0)$ and the vertical line bundle $V_{\mathbf{x}} = \{\mathbf{x}\} \times \mathbb{R}(0,1)$. Then, the Oseledets subspaces must coincide with $H_{\mathbf{x}}$ and $V_{\mathbf{x}}$ almost everywhere. We choose cylinders $Z_n \subset M$ whose first $n$ iterates $f^i(Z_n)$, $0 \leq i \leq n-1$ are pairwise disjoint. Then we construct cocycles $B_n$ by modifying $A$ on some of these iterates so that

$$B_n^n(x)H_{\mathbf{x}} = V_{f^n(\mathbf{x})} \quad \text{and} \quad B_n^n(x)V_{\mathbf{x}} = H_{f^n(\mathbf{x})} \quad \text{for all } \mathbf{x} \in Z_n.$$

We deduce that the Lyapunov exponents of $B_n$ vanish. Moreover, by construction, each $B_n$ is constant on every atom of some finite partition of $M$ into cylinders. In particular, $B_n$ is Hölder continuous for every $r > 0$. From the construction we also get that

$$\|B_n - A\|_r \leq \text{const}\left(2^{2r}/\sigma\right)^{n/2} \tag{7.1}$$

decays to zero as $n \to \infty$. This is how we get the claims in the theorem. Now let us fill-in the details of the proof.

Let $n = 2k+1$ for some $k \geq 1$ and $Z_n = [0; 2, \dots, 2, 1, \dots, 1, 1]$ where the symbol 2 appears $k$ times and the symbol 1 appears $k+1$ times. Notice that the $f^i(Z_n)$, $0 \leq i \leq 2k$ are pairwise disjoint. Let

$$\varepsilon_n = \sigma^{-k} \quad \text{and} \quad \delta_n = \arctan \varepsilon_n. \tag{7.2}$$

Define $R : M \to \mathrm{SL}(2, \mathbb{R})$ by

$$\begin{aligned}
&R(\mathbf{x}) = \text{rotation of angle } \delta_n \quad \text{if } \mathbf{x} \in f^k(Z_n) \\
&R(\mathbf{x}) = \begin{pmatrix} 1 & 0 \\ \varepsilon_n & 1 \end{pmatrix} \quad \text{if } \mathbf{x} \in Z_n \cup f^{2k}(Z_n) \\
&R(\mathbf{x}) = \text{id} \quad \text{in all other cases.}
\end{aligned}$$

and then take $B_n = AR_n$.

**Lemma 7.1.2.** *$B_n^n(\mathbf{x})H_{\mathbf{x}} = V_{f^n(\mathbf{x})}$ and $B_n^n(\mathbf{x})V_{\mathbf{x}} = H_{f^n(\mathbf{x})}$ for all $\mathbf{x} \in Z_n$.*

*Proof.* Notice that for any $\mathbf{x} \in Z_n$,

$$\begin{aligned}
B_n^k(\mathbf{x})H_{\mathbf{x}} &= \mathbb{R}(\varepsilon_n, 1) \quad \text{and} \quad B_n^k(\mathbf{x})V_{\mathbf{x}} = V_{f^k(\mathbf{x})} \\
B_n^{k+1}(\mathbf{x})H_{\mathbf{x}} &= V_{f^{k+1}(\mathbf{x})} \quad \text{and} \quad B_n^{k+1}(\mathbf{x})V_{\mathbf{x}} = \mathbb{R}(-\varepsilon_n, 1) \\
B_n^{2k}(\mathbf{x})H_{\mathbf{x}} &= V_{f^{2k}(\mathbf{x})} \quad \text{and} \quad B_n^{2k}(\mathbf{x})V_{\mathbf{x}} = \mathbb{R}(-1, \varepsilon_n).
\end{aligned}$$

The claim follows by iterating one more time. □

**Lemma 7.1.3.** *There exists $C > 0$ such that $\|B_n - A\|_r \leq C\,(2^{2r}/\sigma)^k$ for every $n$.*

*Proof.* Let $L_n = A - B_n$. Clearly, $\sup \|L\| \le \sup \|A\| \, \| \mathrm{id} - R_n\|$ and this is bounded by $\sigma\varepsilon_n$. Now let us estimate the second term in the definition (7.1). If $\mathbf{x}$ and $\mathbf{y}$ are not in the same cylinder $[0; a]$ then $d(\mathbf{x}, \mathbf{y}) = 1$, and so

$$\frac{\|L_n(\mathbf{x}) - L_n(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r} \le 2 \sup \|L_n\| \le 2\sigma\varepsilon_n. \tag{7.3}$$

From now on we suppose $\mathbf{x}$ and $\mathbf{y}$ belong to the same cylinder. Then, since $A$ is constant on cylinders,

$$\frac{\|L_n(\mathbf{x}) - L_n(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r} = \frac{\|A(\mathbf{x})(R_n(\mathbf{x}) - R_n(\mathbf{y}))\|}{d(\mathbf{x}, \mathbf{y})^r} \le \sigma \frac{\|R_n(\mathbf{x}) - R_n(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r}.$$

If neither $\mathbf{x}$ nor $\mathbf{y}$ belong to $Z_n \cup f^k(Z_n) \cup f^{2k}(Z_n)$ then $R_n(\mathbf{x})$ and $R_n(\mathbf{y})$ are both equal to id, and so the expression on the right vanishes. If $\mathbf{x}$ and $\mathbf{y}$ belong to the same $f^i(Z_n)$ then $R_n(\mathbf{x}) = R_n(\mathbf{y})$ and so, once more, the expression on the right vanishes. We are left to consider the case when one of the points belongs to some $f^i(Z_n)$ and the other one does not. Then $d(\mathbf{x}, \mathbf{y}) \ge 2^{-2k}$ and so, using once more that $\| \mathrm{id} - R_n\| \le \varepsilon_n$ at every point,

$$\frac{\|L_n(\mathbf{x}) - L_n(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r} \le \sigma \frac{\|R_n(\mathbf{x}) - R_n(\mathbf{y})\|}{d(\mathbf{x}, \mathbf{y})^r} \le 2\sigma\varepsilon_n 2^{2kr}.$$

Noting that this bound is worst than (7.3), we conclude that

$$\|L_n\|_r \le \sigma\varepsilon_n + 2\sigma\varepsilon_n 2^{2kr} \le 3\sigma \left(2^{2r}/\sigma\right)^k$$

Now it suffices to take $C = 3\sigma$. □

Now we want to prove that $\lambda_\pm(B_n) = 0$ for every $n$. Let $\mu_n$ be the normalized restriction of $\mu$ to $Z_n$ and $f_n : Z_n \to Z_n$ be the first return map (defined on a full measure subset). Indeed,

$$Z_n = \bigsqcup_{b \in \mathcal{B}} [0; w, b, w] \quad \text{(up to a zero measure subset)}$$

where $w = (1, \dots, 1, 2, \dots, 2, 2)$ and the union is over the set $\mathcal{B}$ of all finite words $b = (b_1, \dots, b_s)$ not having $w$ as a sub-word. Moreover,

$$f_n \mid [0; w, b, w] = f^{n+s} \mid [0; w, b, w] \quad \text{for each } b \in \mathcal{B}.$$

Thus, $(f_n, \mu_n)$ is a Bernoulli shift with an infinite alphabet $\mathcal{B}$ and probability vector given by $p_b = \mu_n([0; w, b, w])$. Let $\hat{B}_n : Z_n \to \mathrm{SL}(2, \mathbb{R})$ be the cocycle induced by $B$ over $f_n$, that is,

$$\hat{B}_n \mid [0; w, b, w] = B_n^{n+s} \mid [0; w, b, w] \quad \text{for each } b \in \mathcal{B}.$$

It is a well known basic fact (see [37, Proposition 2.9], for instance) that the Lyapunov spectrum of the induced cocycle is obtained multiplying the Lyapunov spectrum of the original cocycle by the average return time. In our setting this means

$$\lambda_\pm(\hat{B}_n) = \frac{1}{\mu(Z_n)} \lambda_\pm(B_n).$$

Therefore, it suffices to prove that $\lambda_\pm(\hat{B}_n) = 0$ for every $n$.

Indeed, suppose the Lyapunov exponents of $\hat{B}_n$ are non-zero and let $E^u_{\mathbf{x}} \oplus E^s_{\mathbf{x}}$ be the Oseledets splitting (defined almost everywhere in $Z_n$). Consider the probability measures $m^u$ and $m^s$ for the cocycle $\hat{B}_n$ defined as in (3.1). The key observation is that, as a consequence of Lemma 7.1.2, the cocycle $\hat{B}_n$ permutes the vertical and horizontal subbundles:

$$\hat{B}_n(\mathbf{x})H_{\mathbf{x}} = V_{f_n(\mathbf{x})} \quad \text{and} \quad \hat{B}_n(\mathbf{x})V_{\mathbf{x}} = H_{f_n(\mathbf{x})} \quad \text{for all } \mathbf{x} \in Z_n. \tag{7.4}$$

Let $m$ be the measure defined on $M \times \mathbb{P}(\mathbb{R}^2)$ by

$$m_n(X) = \frac{1}{2}(\mu_n(\{\mathbf{x} \in Z_n : V_{\mathbf{x}} \in X\}) + \mu_n(\{\mathbf{x} \in Z_n : H_{\mathbf{x}} \in X\}).$$

In other words, $m_n$ projects down to $\mu_n$ and its disintegration is given by $\mathbf{x} \mapsto (\delta_{H_{\mathbf{x}}} + \delta_{V_{\mathbf{x}}})/2$. It is clear from (7.4) that $m_n$ is $\hat{B}_n$-invariant.

**Lemma 7.1.4.** *The probability measure $m_n$ is ergodic.*

*Proof.* Suppose there is an invariant set $\mathcal{X} \subset M \times \mathbb{P}(\mathbb{R}^2)$ with $m_n(\mathcal{X}) \in (0,1)$. Let $X_0$ be the set of $\mathbf{x} \in Z_n$ whose fiber $\mathcal{X} \cap (\{\mathbf{x}\} \times \mathbb{P}(\mathbb{R}^2))$ contains neither $H_{\mathbf{x}}$ nor $V_{\mathbf{x}}$. In view of (7.4), $X_0$ is an $f_n$-invariant set and so its $\mu_n$-measure is either 0 or 1. Since $m_n(\mathcal{X}) > 0$, we must have $\mu_n(X_0) = 0$. The same kind of argument shows that $\mu_n(X_2) = 0$, where $X_2$ is the set of $\mathbf{x} \in Z_n$ whose fiber contains both $H_{\mathbf{x}}$ and $V_{\mathbf{x}}$. Now let $X_H$ be the set of $\mathbf{x} \in Z_n$ whose fiber contains $H_{\mathbf{x}}$ but not $V_{\mathbf{x}}$, and let $X_V$ be the set of $\mathbf{x} \in Z_n$ whose fiber contains $V_{\mathbf{x}}$ but not $H_{\mathbf{x}}$. The previous observations show that $X_H \cup X_V$ has full $\mu_n$-measure and it follows from (7.4) that

$$f_n(X_H) = X_V \quad \text{and} \quad f_n(X_V) = X_H.$$

Thus, $\mu_n(X_H) = 1/2 = \mu_n(X_V)$ and $f_n^2(X_H) = X_H$ and $f_n^2(X_V) = X_V$. This is a contradiction because $f_n$ is Bernoulli and, in particular, the second iterate is ergodic. □

By Lemma 3.1.4, the invariant measure $m_n$ is a linear combination of $m^u$ and $m^s$. Then, in view of Lemma 7.1.4, $m_n$ must coincide with either $m^s$ and $m^u$. This is a contradiction, because the conditional probabilities of $m_n$ are supported on exactly two points on each fiber, whereas the conditional probabilities of either $m^u$ and $m^s$ are Dirac masses on a single point. This contradiction proves that the Lyapunov exponents of $\hat{B}_n$ do vanish for every $n$, and that concludes the proof of Theorem 7.1.1.

The same kind of argument shows that, in general, one can expect continuity to hold when some of the probabilities $p_i$ vanishes:

**Remark 7.1.5.** (Kifer [23]) Take $d = 2$, a probability vector $p = (p_1, p_2)$ with non-negative coefficients, and a cocycle $A = (A_1, A_2)$ defined by

$$A_1 = \begin{pmatrix} \sigma & 0 \\ 0 & \sigma^{-1} \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

where $\sigma > 1$. By the same arguments as we used before, $\lambda_\pm(A, p) = 0$ if $p_2 > 0$. In this regard, observe that the cocycle induced by $A$ over the cylinder $[0; 2]$ exchanges the vertical and horizontal directions, just as in (7.4). Now, it is clear that $\lambda_\pm(A, (1, 0)) = \pm \log \sigma$. Thus, the Lyapunov exponents are discontinuous at $(A, (1, 0))$.

## 7.2 Open problems

While our results give a very complete answer to the continuity problem for two-dimensional matrices some interesting problems remain, that we pose here:

**Problem 7.2.1.** Does continuity extend to unbounded cocycles satisfying an integrability condition, for instance, $\log \|A^{\pm 1}\| \in L^1(\mu)$ ? Notice that this condition involves both the cocycle and the probability measure. So, in this context the topology should be defined in the space of pairs $(A, p)$.

**Problem 7.2.2.** Does continuity extend to *locally constant* cocycles over Bernoulli shifts, that is, such that $A(\mathbf{x})$ depends on a bounded number of coordinates of $f$ ? Notice that we have handled the case when $A(\mathbf{x})$ depends only on the zeroth coordinate of $\mathbf{x}$. What about for locally cocycles over Markov systems.

**Problem 7.2.3.** Does continuity extend to extremal Lyapunov exponents of $\mathrm{GL}(d, \mathbb{C})$-cocycles for any dimension $d$ ? Then, using exterior powers in a well-known way (see e.g. [33]), one would get continuity for *all* Lyapunov exponents. An interesting special case to look at are symplectic cocycles, that is, such that every $A(\mathbf{x})$ preserves some given symplectic form.

**Problem 7.2.4.** Can we say more about the regularity of the Lyapunov exponents as functions of the cocycle: Hölder or even Lipschitz continuity ? differentiability ? Some partial answers and related results can be found in [32, 33], for instance.

# Bibliography


[1] A. Arbieto and J. Bochi. $L^p$-generic cocycles have one-point Lyapunov spectrum. *Stoch. Dyn.*, 3:73–81, 2003.

[2] L. Arnold and N. D. Cong. On the simplicity of the Lyapunov spectrum of products of random matrices. *Ergod. Th. & Dynam. Sys.*, 17:1005–1025, 1997.

[3] A. Avila and M. Viana. Extremal Lyapunov exponents of smooth cocycles. Preprint www.preprint.impa.br 2008.

[4] P. Billingsley. *Convergence of probability measures.* J. Wiley & Sons, 1968.

[5] J. Bochi. Discontinuity of the Lyapunov exponents for non-hyperbolic cocycles. Preprint www.mat.puc-rio.br/∼jairo/.

[6] J. Bochi. Genericity of zero Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*, 22:1667–1696, 2002.

[7] J. Bochi. $C^1$-generic symplectic diffeomorphisms: partial hyperbolicity and zero centre Lyapunov exponents. *J. Inst. Math. Jussieu*, 8, 2009.

[8] J. Bochi and M. Viana. The Lyapunov exponents of generic volume-preserving and symplectic maps. *Ann. of Math.*, 161:1423–1485, 2005.

[9] C. Bonatti, L. J. Díaz, and M. Viana. *Dynamics beyond uniform hyperbolicity*, volume 102 of *Encyclopaedia of Mathematical Sciences.* Springer-Verlag, 2005.

[10] C. Bonatti, X. Gómez-Mont, and M. Viana. Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 20:579–624, 2003.

[11] C. Bonatti and M. Viana. Lyapunov exponents with multiplicity 1 for deterministic products of matrices. *Ergod. Th. & Dynam. Sys*, 24:1295–1330, 2004.

[12] J. Bourgain. Positivity and continuity of the Lyapounov exponent for shifts on $\mathbb{T}^d$ with arbitrary frequency vector and real analytic potential. *J. Anal. Math.*, 96:313–355, 2005.

[13] J. Bourgain and S. Jitomirskaya. Continuity of the Lyapunov exponent for quasiperiodic operators with analytic potential. *J. Statist. Phys.*, 108:1203–1218, 2002.

[14] L. Breiman. *Probability.* Classics in applied mathematics. Addison-Weslley Pub. Co., 1968.

[15] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.

[16] H. Furstenberg and Yu. Kifer. Random matrix products and measures in projective spaces. *Israel J. Math*, 10:12–32, 1983.

[17] I. Ya. Gol'dsheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Uspekhi Mat. Nauk.*, 44:13–60, 1989.

[18] R. M. Gray. *Probability, random processes, and ergodic properties.* Springer, 2009.

[19] Y. Guivarc'h and A. Raugi. Products of random matrices : convergence theorems. *Contemp. Math.*, 50:31–54, 1986.

[20] H. Hennion. Loi des grands nombres et perturbations pour des produits réductibles de matrices aléatoires indépendantes. *Z. Wahrsch. Verw. Gebiete*, 67:265–278, 1984.

[21] R. Johnson. Lyapounov numbers for the almost periodic Schrödinger equation. *Illinois J. Math.*, 28:397–419, 1984.

[22] J. L. Kelley. *General Topology.* D. Van Nostrand Company, Inc., 1955.

[23] Yu. Kifer. Perturbations of random matrix products. *Z. Wahrsch. Verw. Gebiete*, 61:83–95, 1982.

[24] J. Kingman. The ergodic theorem of subadditive stochastic processes. *J. Royal Statist. Soc.*, 30:499–510, 1968.

[25] O. Knill. The upper Lyapunov exponent of $\mathrm{SL}(2,\mathbb{R})$ cocycles: discontinuity and the problem of positivity. In *Lyapunov exponents (Oberwolfach, 1990)*, volume 1486 of *Lecture Notes in Math.*, pages 86–97. Springer, 1991.

[26] O. Knill. Positive Lyapunov exponents for a dense set of bounded measurable $\mathrm{SL}(2,\mathbb{R})$-cocycles. *Ergod. Th. & Dynam. Sys.*, 12:319–331, 1992.

[27] F. Ledrappier. Positivity of the exponent for stationary sequences of matrices. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lect. Notes Math.*, pages 56–73. Springer, 1986.

[28] P. A. Loeb and E. Talvila. Lusin's Theorem and Bochner Integration. In *Scientiae Mathematicae Japonicae Online*, volume 10, 2004, 55-62.

[29] R. Mañé. Oseledec's theorem from the generic viewpoint. In *Procs. International Congress of Mathematicians, Vol. 1, 2 (Warsaw, 1983)*, pages 1269–1276, Warsaw, 1984. PWN Publ.

[30] V. I. Oseledets. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231, 1968.

[31] É. Le Page. Théorèmes limites pour les produits de matrices aléatoires. In *Probability measures on groups (Oberwolfach, 1981)*, volume 928 of *Lecture Notes in Math.*, pages 258–303. Springer, 1982.

[32] É. Le Page. Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. *Ann. Inst. H. Poincaré Probab. Statist.*, 25:109–142, 1989.

[33] Y. Peres. Analytic dependence of Lyapunov exponents on transition probabilities. In *Lyapunov exponents (Oberwolfach, 1990)*, volume 1486 of *Lecture Notes in Math.*, pages 64–80. Springer, 1991.

[34] V. A. Rokhlin. On the fundamental ideas of measure theory. *A. M. S. Transl.*, 10:1–52, 1952. Transl. from Mat. Sbornik 25 (1949), 107–150.

[35] D. Ruelle. Analyticity properties of the characteristic exponents of random matrix products. *Adv. in Math.*, 32:68–80, 1979.

[36] B. Simon and M. Taylor. Harmonic analysis on $\mathrm{SL}(2,\mathbb{R})$ and smoothness of the density of states in the one-dimensional Anderson model. *Comm. Math. Phys.*, 101:1–19, 1985.

[37] M. Viana. Lyapunov exponents of Teichmüller flows. In *Partially hyperbolic dynamics, laminations, and Teichmüller flow*, volume 51 of *Fields Inst. Commun.*, pages 139–201. Amer. Math. Soc., 2007.