Instituto Nacional de Matematica Pura e Aplicada
submitted for the approval of doctorate of philosophy

Supervisor: Marcelo Viana

Moduli of continuity of Lyapunov exponents for random $\mathrm{GL}(2)$-cocycles

El Hadji Yaya Tall

Rio de Janeiro , February 24, 2018

Abstract

The Lyapunov exponents of random GL(2)-cocycles are Hölder continuous functions of the underlying probability distribution at each point with simple Lyapunov spectrum. Moreover, they are log-Hölder continuous at every point.

Contents

Chapter 1

Introduction

Let $A_1, \dots, A_N$ be invertible 2-by-2 matrices and $p_1, \dots, p_N$ be positive numbers such that $p_1 + \cdots + p_N = 1$. Consider the probability measure

$$\nu = p_1 \delta_{A_1} + \cdots + p_N \delta_{A_N}$$

on the group GL(2) of invertible 2-by-2 matrices. A theorem of Furstenberg, Kesten [FK60] asserts that there exist numbers $\lambda_-(\nu) \leq \lambda_+(\nu)$ such that

$$\frac{1}{n} \log \|g_{n-1} \cdots g_0\| \to \lambda_+(\nu) \text{ and } -\frac{1}{n} \log \|g_0^{-1} \cdots g_{n-1}^{-1}\| \to \lambda_-(\nu)$$

for $\nu^{\mathbb{Z}}$-almost all sequences $(g_j)_j$ with values in $\{A_1, \dots, A_N\}$. They are called the *Lyapunov exponents* of ν.

It was proven by Bocker, Viana [BV17] that the Lyapunov exponents are continuous functions of the probability weights and the matrices coefficients, at every point. Avila, Eskin and Viana [AEV] announced that the statement extends to arbitrary dimension, which also provides an alternative proof of the 2-dimensional case (see Viana [Via14, Chapter 10]).

Here we investigate moduli of continuity of these functions, in dimension 2:

Theorem A. *The Lyapunov exponents are Hölder continuous functions of the probability weights $p_1, \dots, p_N$ and the coefficients of the matrices $A_1, \dots, A_N$ at every point such that $\lambda_-(\nu) < \lambda_+(\nu)$. Moreover, they are log-Hölder continuous at every point.*

It is a classical result of Le Page [Pag89] that, in any dimension, the largest Lyapunov exponent is Hölder continuous on every compact set of probability measures ν satisfying strong irreducibility and the contraction property. These assumptions can not be removed, as shown by a construction of Duarte, Klein, Santos [DKS] that we adapt to our setting in Section 5.

Theorem A means that Hölder continuity still holds under the weaker assumption that the Lyapunov spectrum is simple, at least in dimension 2 and pointwise. Moreover, there is an explicit modulus of continuity – log-Hölder continuity – that holds at every point. More precise statements, for probability distributions that are not necessarily finitely supported, will be given in the next section.

Independently, and at about the same time, Duarte, Klein [**?**] proved that the Lyapunov exponents are weak-Hölder continuous functions of the probability weights and matrix coefficients on any compact domain with $\lambda_-(\nu) < \lambda_+(\nu)$. *Weak-Hölder* is defined by replacing $d(x,y)^\beta$ with $\exp(-\beta(-\log d(x,y))^\theta)$ with $\theta < 1$ in the definition of Hölder continuity. They use a very different approach, based on large deviations and the so-called avalanche principle [GS01, DK16].

Chapter 2

Main results

Let ν be a probability measure on the group $G = \mathrm{GL}(d)$, $d \geq 2$ such that the functions $g \mapsto \log \|g^{\pm 1}\|$ are in $L^1(\nu)$. The *extremal Lyapunov exponents* of ν are the numbers $\lambda_-(\nu) \leq \lambda_+(\nu)$ defined by

$$\lambda_+(\nu) = \lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0\| \text{ and } \lambda_-(\nu) = \lim_n -\frac{1}{n} \log \|g_0^{-1} \cdots g_{n-1}^{-1}\| \tag{2.0.1}$$

for $\nu^{\mathbb{Z}}$-almost every $(g_j)_j \in G^{\mathbb{Z}}$. See Furstenberg, Kesten [FK60].

Let $\mathcal{M}_c(G)$ be the space of compactly supported probability measures on G equipped with the smallest topology $\mathcal{T}$ such that:

1. $\nu \mapsto \int_G \psi \, d\nu$ is continuous for every continuous $\psi : G \to \mathbb{R}$ and
2. $\nu \mapsto \operatorname{supp} \nu$ is continuous relative to the Hausdorff topology in the space of compact subsets of G.

In other words, $\mathcal{T}$ is the smallest topology that contains both the weak* topology and the pull-back of the Hausdorff topology under the support map.

Let $\delta(\cdot, \cdot)$ be the distance defined on G by $\delta(g, g') = \|g - g'\| + \|g^{-1} - g'^{-1}\|$. For each $\theta \in (0, 1]$, the function $\delta(\cdot, \cdot)^\theta$ is also a distance on G. Let $L_\theta(G)$ be the associated space of 1-Lipschitz continuous functions, that is, functions $\psi : G \to \mathbb{R}$ such that

$$|\psi(g) - \psi(g')| \leq \delta(g, g')^\theta \text{ for every } g, g' \in G.$$

Denote by δ_θ the corresponding *Wasserstein distance* on $\mathcal{M}_c(G)$, that is,

$$\begin{aligned} \delta_\theta(\nu, \nu') &= \sup \left\{ \left| \int_G \psi \, d(\nu - \nu') \right| : \psi \in L_\theta(G) \right\} \\ &= \inf \left\{ \int_G \delta(g, g')^\theta \, d\xi(g, g') : \xi \text{ is a coupling of } \nu \text{ and } \nu' \right\} \end{aligned} \tag{2.0.2}$$

(see Villani [Vil09, Chapter I.6]). A *coupling* is a measure on $G \times G$ whose projections to the two coordinates coincide with ν and ν'. The absolute value in the first part of (2.0.2) is innocuous, as $\psi \in L_\theta(G)$ if and only if $-\psi \in L_\theta(G)$. Also, the integral is not affected when ψ is replaced by $\psi + \text{constant}$. This observation will be used several times.

It is well known (see [Vil09, Theorem 6.9]) that every δ_θ generates the weak* topology on $\mathcal{M}_c(G)$. Let δ_H be any distance generating the Hausdorff topology in the space of compact subsets of G. Then

$$\delta_{\mathcal{T},\theta}(\nu, \nu') = \delta_\theta(\nu, \nu') + \delta_H(\operatorname{supp} \nu, \operatorname{supp} \nu').$$

is a distance and it generates the topology of $\mathcal{M}_c(G)$. Here we take

$$\delta_H(K, K') = \inf\{r > 0 : K \subset B_r(K') \text{ and } K' \subset B_r(K)\}, \tag{2.0.3}$$

where the neighborhoods are with respect to the distance δ on G.

2.1 Statements

We are going to prove that, in dimension $d = 2$, the functions $\nu \mapsto \lambda_\pm(\nu)$ are pointwise Hölder continuous at every point whose Lyapunov exponents are distinct:

Theorem B. *Let $G = \mathrm{GL}(2)$. For every $\nu \in \mathcal{M}_c(G)$ with $\lambda_-(\nu) < \lambda_+(\nu)$ there exists a neighborhood $U \subset \mathcal{M}_c(G)$ and for every $\theta \in (0, 1]$ there exist constants $C > 0$ and $\beta > 0$ such that*

$$|\lambda_\pm(\nu) - \lambda_\pm(\nu')| \le C\delta_{\mathcal{T},\theta}(\nu, \nu')^\beta \text{ for every } \nu' \in U.$$

In some cases, for instance when ν is diagonal (Section 3.4), we even get local *Lipschitz* continuity, meaning that $\beta = 1$. On the other hand, the conclusion of Theorem B need not hold when $\lambda_-(\nu) = \lambda_+(\nu)$, as we comment upon in Section 5. Instead,

Theorem C. *Let $G = \mathrm{GL}(2)$. For every $\nu \in \mathcal{M}_c(G)$ there exists a neighborhood $U \subset \mathcal{M}_c(G)$ and for every $\theta \in (0, 1]$ there exist constants $C > 0$ and $\beta > 0$ such that*

$$|\lambda_\pm(\nu) - \lambda_\pm(\nu')| \le C \left(\log \frac{1}{\delta_\theta(\nu, \nu')}\right)^{-\beta} \text{ for every } \nu' \in U.$$

The statement remains true if we replace δ_θ with $\delta_{\mathcal{T},\theta}$.

For measures with finite support,

$$\nu = p_1\delta_{A_1} + \cdots + p_N\delta_{A_N} \text{ and} \nu' = p'_1\delta_{A'_1} + \cdots + p'_N\delta_{A'_N},$$

the definition (2.0.2) means that

$$\delta_\theta(\nu,\nu') = \sup\left\{\sum_{j=1}^N p_j\psi(A_j) - p'_j\psi(A'_j) : \psi \in L_\theta(G)\right\}$$
$$\leq \sup\left\{\sum_{j=1}^N |p_j - p'_j|\|\psi(A_j)\| + p'_j\|\psi(A_j) - \psi(A'_j)\| : \psi \in L_\theta(G)\right\}.$$

The expression on the right-hand side of the identity does not change when one adds a constant to ψ. Thus, we may restrict ourselves to functions such that $\psi(A_1) = 0$. Then $\|\psi(A_j)\| \leq D^\theta_\nu$ for every j, where D^θ_ν denotes the δ^θ-diameter of $\operatorname{supp}\nu = \{A_1, \dots, A_N\}$. It follows that

$$\delta_\theta(\nu,\nu') \leq \sum_{j=1}^N D^\theta_\nu |p_j - p'_j| + \|A_j - A'_j\|^\theta \tag{2.1.1}$$

for any $p_1, \dots, p_N$ and $A_1, \dots, A_N$. This shows that Theorem A is a special case of Theorems B and C.

2.2 Comments on the proofs

Denote $M = G^{\mathbb{Z}}$ and $\mu = \nu^{\mathbb{Z}}$ and let $f : M \to M$ be the shift map. The generic point $(g_j)_{j\in\mathbb{Z}}$ of M will also be represented as x. Moreover, we write $x_+ = (g_j)_{j\geq 0}$ and $x_- = (g_j)_{j<0}$. We are going to consider the projective cocycle

$$\mathbb{P}F : M \times P \to M \times P, \quad \mathbb{P}F(x,v) = (f(x), g_0 v).$$

Observe that $\mathbb{P}F^n(x,v) = (f^n(x), g_{n-1}\cdots g_0 v)$ for any $n \geq 1$.

Define $\phi : M \times P \to \mathbb{R}$ through $\phi(x,[v]) = \log\|g_0 v\|/\|v\|$. Since the image depends only on g_0 and $[v]$, we may also view this as a function $\phi : G \times P \to \mathbb{R}$. The Furstenberg formula (see [Via14, Chapter 6]) asserts that

$$\lambda_+(\nu) = \sup\left\{\int_{G\times P} \phi\, d\mu\, d\eta : \eta \text{ is a } \nu\text{-stationary measure}\right\} \tag{2.2.1}$$

and the supremum is attained. A ν-stationary measure η is called *maximal* if it realizes the supremum in (2.2.1).

The strategy to prove Theorem B is to show that if $\lambda_-(\nu) < \lambda_+(\nu)$ then the (unique) maximal ν-stationary measure is a kind of hyperbolic attractor for the dynamics of the associated random walk: for any compact set K outside the support of the stationary measures and any initial point, the probability that the random orbit hits K at time n decays exponentially fast. See Propositions 3.3.1 and 3.4.1. The proof uses a special case of a result of Avila, Eskin, Viana [AEV] that we state in Theorem 3.2.1.

Such exponential estimates are generally false when $\lambda_-(\nu) = \lambda_+(\nu)$. Instead, we prove that the random walk has power-law "diffusion": there is $\beta > 0$ such that the probability mentioned in the previous paragraph decreases as $O(n^{-\beta})$. This is expressed more precisely in (4.2.6)–(4.2.8) in Section 4.2. The proof is based on various quantitative versions of the central limit theorem, with ideas from the theory of perpetuities intervening also in the degenerate triangular case. The conclusion of Theorem C is readily seen to follow from such a behavior (Section 4.2).

In detailing the arguments we will use a few simplifying observations:

Remark 2.2.1. Theorems B and C hold for λ_+ if and only if they hold for λ_-. That is because (see [Via14, Chapter 4])

$$\lambda_+(\nu) + \lambda_-(\nu) = \int_G \log|\det g|\, d\nu(g), \tag{2.2.2}$$

and the function $\nu \mapsto \int_G \log|\det g|\, d\nu(g)$ is locally Lipschitz with respect to every distance δ_θ: the latter claim is a rather straightforward consequence of the definitions.

Remark 2.2.2. For each $n \geq 2$, we use $\nu^{(n)} = \nu * \cdots * \nu$ to denote the *n-convolution* of ν, that is, the image of ν^n under the map $(g_0, \ldots, g_{n-1}) \mapsto g_{n-1} \cdots g_0$. It is easy to deduce from the definitions that

$$\lambda_+\left(\nu^{(n)}\right) = n\lambda_+(\nu) \text{ and } \delta_\theta\left(\nu^{(n)}, \nu'^{(n)}\right) \leq n\delta_\theta(\nu, \nu')$$

for any $\nu, \nu' \in \mathcal{M}_c(G)$. So, ν is a point of Hölder (respectively, log-Hölder) continuity for the Lyapunov exponents if $\nu^{(n)}$ is.

Remark 2.2.3. Given $h \in G$, let $h\nu h^{-1}$ denote the image of any $\nu \in \mathcal{M}_c(G)$ under the map $g \mapsto hgh^{-1}$. It is clear that $\lambda_\pm(h\nu h^{-1}) = \lambda_\pm(\nu)$ and it is straightforward to check that the map $\nu \mapsto h\nu h^{-1}$ is locally Lipschitz relative to any distance d_θ. So, ν is a point of Hölder (respectively, log-Hölder) continuity for the Lyapunov exponents if and only if $h\nu h^{-1}$ is.

Remark 2.2.4. It follows directly from the Hölder inequality that the Wasserstein distances are related by $\delta_\theta(\nu, \nu') \leq \delta_{\theta'}(\nu, \nu')^{\theta/\theta'}$ for any $0 < \theta < \theta' \leq 1$ and any $\nu, \nu' \in \mathcal{M}_c(G)$. Thus, for the purpose of proving our results it is no restriction to suppose that θ is close to zero.

Chapter 3

Proof of Theorem B

When ν is strongly irreducible and has the contraction property, the conclusion was first obtained by Le Page [Pag89]. The contraction property means that the *eccentricity* $\|g\|\|g^{-1}\|$ is unbounded on the semigroup generated by the support of ν, and it is clearly necessary for $\lambda_-(\nu) < \lambda_+(\nu)$.

Thus we only need to treat the case when ν is not strongly irreducible, that is, when there exists a finite family of lines which is invariant under every $g \in \operatorname{supp}\nu$. Since we take the Lyapunov exponents to be distinct, these lines must be fixed under every matrix in the support of ν, and there are at most two of them. So, in the remainder of this section we assume that there exists some line $E \subset \mathbb{R}^2$ such that $gE = E$ for every $g \in \operatorname{supp}\nu$.

According to the Oseledets theorem (see [Via14, Chapter 4]), there exists an invariant splitting $\mathbb{R}^2 = E^u_x \oplus E^s_x$ defined at μ-almost every $x \in M$ such that

$$\lim_n \frac{1}{n} \log \|g_{n-1} \cdots g_0 v\| = \begin{cases} \lambda_+(\nu) & \text{if } v \in E^u_x \setminus \{0\} \\ \lambda_-(\nu) & \text{if } v \in E^s_x \setminus \{0\}. \end{cases} \tag{3.0.1}$$

The invariant subspace E must correspond to one of the two Oseledets subbundles: either $E = E^s_x$ for μ-almost every $x \in M$ or $E = E^u_x$ for μ-almost every $x \in M$. We consider the first case and we are going to prove that ν is a point of Hölder continuity for λ_+. A dual argument shows that in the second case ν is a point of Hölder continuity for λ_-. By Remark 2.2.1, that proves the theorem in both cases.

3.1 Markov operators

Let ν be fixed. It is no restriction to suppose that $E = \mathbb{R} \times \{0\}$. Then every $g \in \operatorname{supp}\nu$ may be written in the form

$$g = \begin{pmatrix} \alpha(g) & \gamma(g) \\ 0 & \beta(g) \end{pmatrix}. \tag{3.1.1}$$

We denote $\tau(g) = \alpha(g)/\beta(g)$. Observe that

$$\int_G \log|\alpha(g)|\,d\nu(g) = \lambda_-(\nu) \text{ and } \int_G \log|\beta(g)|\,d\nu(g) = \lambda_+(\nu):$$

the first identity follows from $E^s_x = (1, 0)$ and the second one is a consequence, using (2.2.2). Then

$$\int_G \log|\tau(g)|\,d\nu(g) = \lambda_-(\nu) - \lambda_+(\nu) < 0. \tag{3.1.2}$$

Lemma 3.1.1. *There is $\theta_* \in (0,1]$ such that $\int_G |\tau|^\theta d\nu < 1$ for every $\theta \in (0, \theta_*]$.*

Proof. Observe that

$$\frac{d}{d\theta}\int_G |\tau|^\theta\,d\nu\bigg|_{\theta=0} = \int_G |\tau|^\theta \log|\tau|\,d\nu\bigg|_{\theta=0} = \int_G \log|\tau|\,d\nu < 0.$$

The claim is an immediate consequence. □

Let $d(\cdot,\cdot)$ denote the projective distance on $P = \mathbb{P}(\mathbb{R}^d)$, defined by

$$d(u, v) = \frac{\|u \wedge v\|}{\|u\|\|v\|} = |\sin \measuredangle(u, v)| \tag{3.1.3}$$

and let $L_\theta(P)$ be the space of functions $\varphi : P \to \mathbb{R}$ such that

$$|\varphi(v) - \varphi(v')| \leq d(v, v')^\theta \text{ for every } v, v' \in P.$$

It is clear that $d(\cdot,\cdot) \leq 1$. Let $\mathcal{M}(P)$ denote the space of probability measures on P, equipped with the corresponding Wasserstein distance (see Villani [Vil09, Section I.6])

$$\begin{aligned} d_\theta(\eta, \eta') &= \sup\left\{\left|\int_P \varphi\,d(\eta - \eta')\right| : \varphi \in L_\theta(P)\right\} \\ &= \inf\left\{\int_{P\times P} d(u, v)^\theta\,d\xi(u, v) : \xi \text{ is a coupling of } \eta \text{ and } \eta'\right\}. \end{aligned} \tag{3.1.4}$$

Consider the Markov operator $\mathcal{P}_\nu : \mathcal{M}(P) \to \mathcal{M}(P)$ defined by

$$\mathcal{P}_\nu \eta = \int_G g_*\eta\,d\nu(g). \tag{3.1.5}$$

It is clear that $\mathcal{P}_\nu$ is continuous relative to the weak* topology on $\mathcal{M}(P)$. By definition, its fixed points are precisely the ν-stationary measures.

Lemma 3.1.2. *For any $\nu, \nu' \in \mathcal{M}_c(G)$, $\eta \in \mathcal{M}(P)$ and $\theta \in (0, 1]$,*

$$d_\theta(\mathcal{P}_\nu\eta, \mathcal{P}_{\nu'}\eta) \leq \delta_\theta(\nu, \nu').$$

Proof. Consider any $\varphi \in L_\theta(P)$. For each $v \in P$, let $\psi_v : G \to \mathbb{R}$ be defined by $\psi_v(g) = \varphi(gv)$. Then

$$|\psi_v(g_1) - \psi_v(g_2)| = |\varphi(g_1 v) - \varphi(g_2 v)| \le d(g_1 v, g_2 v)^\theta \le \delta(g_1, g_2)^\theta$$

for every $g_1, g_2 \in G$. This proves that $\psi_v \in L_\theta(G)$. Hence,

$$\left| \int_G \varphi(gv)\, d(\nu - \nu')(g) \right| = \left| \int_G \psi_v \, d(\nu - \nu') \right| \le \delta_\theta(\nu, \nu')$$

for every $v \in P$. Thus

$$\begin{aligned} \left| \int_P \varphi \, d\left(\mathcal{P}_\nu \eta - \mathcal{P}_{\nu'} \eta\right) \right| &= \left| \int_P \int_G \varphi(gv)\, d\left(\nu - \nu'\right)(g)\, d\eta(v) \right| \\ &\le \int_P \left| \int_G \psi_v \, d(\nu - \nu')(g) \right| d\eta(v), \end{aligned}$$

is bounded by $\delta_\theta(\nu, \nu')$ for any $\varphi \in L_\theta(P)$. □

Lemma 3.1.3. *For any $\nu \in \mathcal{M}_c(G)$ there exists $C_1 > 1$ such that*

$$d_\theta(\mathcal{P}_\nu \eta, \mathcal{P}_\nu \eta') \le C_1 d_\theta(\eta, \eta') \text{ for every } \eta, \eta' \in \mathcal{M}(P) \text{ and } \theta \in (0, 1].$$

Proof. Given $\varphi \in L_\theta(P)$, let $\psi : G \to \mathbb{R}$ be given by $\psi(v) = \int_G \varphi(gv)\, d\nu(g)$. Fix $C_1 > 1$ such that $d(gv_1, gv_2) \le C_1 d(v_1, v_2)$ for $v_1, v_2 \in P$ and $g \in \operatorname{supp} \nu$. Then

$$|\psi(v_1) - \psi(v_2)| = \left| \int \left[\varphi(gv_1) - \varphi(gv_2)\right] \, d\nu(g) \right| \le d(gv_1, gv_2)^\theta \le C_1^\theta d(v_1, v_2)^\theta$$

for every $v_1, v_2 \in P$. This proves that $\psi / C_1^\theta \in L_\theta(P)$. Hence,

$$\left| \int_P \varphi \, d(\mathcal{P}_\nu \eta - \mathcal{P}_\nu \eta') \right| = \left| \int_P \psi \, d(\eta - \eta') \right| \le C_1^\theta d_\theta(\eta, \eta')$$

for any $\varphi \in L_\theta(P)$. Note that $C_1^\theta \le C_1$. □

Lemma 3.1.4. *For any $\nu \in \mathcal{M}_c(G)$ there exist $L > 0$ and a neighborhood $U \subset \mathcal{M}_c(G)$ such that*

$$\left|\lambda_+(\nu) - \lambda_+(\nu')\right| \le L \delta_\theta(\nu, \nu') + L d_\theta(\eta, \eta')$$

any $\nu' \in U$, any maximal ν-stationary measure η, any ν'-stationary measure η' and any $\theta \in (0, 1]$.

Proof. Fix any compact neighborhood $K \subset G$ of the support of ν. Clearly, there exists $L' > 0$, depending only on K, such that

$$\begin{aligned} &|\phi(g, v_1) - \phi(g, v_2)| \le L' d(v_1, v_2) \le L' d(v_1, v_2)^\theta \\ &|\phi(g_1, v) - \phi(g_2, v)| \le L' \delta(g_1, g_2) \le L' (\operatorname{diam} K)^{1-\theta} \delta(g_1, g_2)^\theta \end{aligned}$$

for $g, g_1, g_2 \in K$, $v, v_1, v_2 \in P$ and $\theta \in (0,1]$. Fix $L \geq \max\{L', L'(\operatorname{diam} K)^{1-\theta}\}$ for every $\theta \in (0,1]$. Then $\phi(g,\cdot)/L \in L_\theta(P)$ for every $g \in K$ and $\phi(\cdot, v)/L \in L_\theta(G)$ for every $v \in P$. Thus

$$\begin{aligned}\left|\lambda_+(\nu) - \lambda_+(\nu')\right| &= \left|\int_{G\times P} \phi \, d(\nu \times \eta - \nu' \times \eta')\right| \\ &\leq \int_P \left|\int_G \phi \, d(\nu - \nu')\right| d\eta + \int_G \left|\int_P \phi \, d(\eta - \eta')\right| d\nu' \\ &\leq L\delta_\theta(\nu, \nu') \, d\eta + L d_\theta(\eta, \eta'),\end{aligned}$$

as claimed. □

3.2 Stationary measures

Keep in mind that we assume that $\lambda_-(\nu) < \lambda_+(\nu)$ and $E^s_x = E$ for μ-almost every x. Then the ν-stationary measures are precisely the convex combinations of

$$\eta^u = \int_M \delta_{E^u_x} \, d\mu(x) \text{ and } \eta^s = \int_M \delta_{E^s_x} \, d\mu(x) = \delta_E. \tag{3.2.1}$$

In particular, η^u is the unique maximal ν-stationary measure. It follows from Bocker, Viana [BV17] that $\lambda_-(\nu') < \lambda_+(\nu')$, and so there exists a unique maximal ν'-stationary measure η'^u for every ν' in a neighborhood of ν. Moreover, η'^u varies continuously with ν'.

Consider any $\eta' \in \mathcal{M}(P)$. It is clear from (3.1.4) that

$$d_\theta(\eta^u, \eta') \leq \int_{P\times P} d(u,v)^\theta \, d\eta^u(u) \, d\eta'(v) = \int_{P\times M} d(E^u_x, v)^\theta \, d\mu(x) \, d\eta'(v),$$

because $\eta^u \times \eta'$ is a coupling of η^u and η'. The *diagonal push-forward* of a measure ξ on $P \times P$ is the measure $\mathcal{D}_\nu \xi$ defined by

$$\int_{P\times P} \varphi(u,v) \, d(\mathcal{D}_\nu \xi)(u,v) = \int_{P\times P\times G} \varphi(gu, gv) \, d\nu(g) \, d\xi(u,v).$$

If ξ is a coupling of η and η' then $\mathcal{D}_\nu \xi$ is a coupling of $\mathcal{P}_\nu \eta$ and $\mathcal{P}_\nu \eta'$. Thus, the n-th diagonal push-forward $\mathcal{D}^n_\nu(\eta^u \times \eta')$ is a coupling of $\mathcal{P}^n_\nu \eta^u = \eta^u$ and $\mathcal{P}^n_\nu \eta'$ for every $n \geq 1$. So,

$$\begin{aligned}d_\theta(\eta^u, \mathcal{P}^n_\nu \eta') &\leq \int_{P\times P\times G} d(gu, gv)^\theta \, d\nu^{(n)}(g) \, d\eta^u(u) \, d\eta'(v) \\ &= \int_{P\times M\times G} d(gE^u_x, gv)^\theta \, d\nu^{(n)}(g) \, d\mu(x) \, d\eta'(v).\end{aligned} \tag{3.2.2}$$

We are going to estimate the expression on the right-hand side of (3.2.2), by decomposing the domain of integration into suitable sub-domains.

Let $E(r) \subset P$ denote the ϵ-neighborhood of E, for any $r > 0$. The following key estimate is contained in a much more general result of Avila, Eskin, Viana [AEV]. For a simplified proof of this special case, see Viana [Via].

Theorem 3.2.1. *There exist $q, s, \epsilon_0 \in (0,1)$ and a neighborhood $U \subset \mathcal{M}_c(G)$ of ν such that*

$$\eta'(E(q\epsilon)) < \frac{2}{3}\eta'(E(\epsilon)))$$

for every $\epsilon \in (0, \epsilon_0)$, every non-atomic ν'-stationary measure η' and every $\nu' \in U$ whose support is contained in the $s\epsilon$-neighborhood of $\operatorname{supp}\nu$.

As a direct consequence, there exist $C_2 > 1$ and $\beta_2 = \log(2/3)/\log q > 0$ such that

$$\eta'(E(r)) \leq C_2 r^{\beta_2} \tag{3.2.3}$$

for every $r \in (0,1)$, every non-atomic ν'-stationary measure η' and every $\nu' \in U$ whose support is contained in the sr-neighborhood of $\operatorname{supp}\nu$.

Remark 3.2.2. Gol'dsheid, Margulis [GM89, Lemmas 4.2 and 4.5] proved that if ξ is a probability measure on $\mathrm{GL}(d)$ such that $\operatorname{supp}\xi$ is not contained in any proper algebraic submanifold, then there exists a unique ξ-stationary measure ζ on $\mathbb{P}(\mathbb{R}^d)$ and it satisfies $\zeta(V) = 0$ for any proper algebraic submanifold V; in particular, ζ is non-atomic.

It is not difficult to check that every probability measure in $\mathcal{M}_c(G)$ is approximated by open sets of probability measures whose supports have non-empty interior and, consequently, are not contained in any proper algebraic submanifold of G. So, every probability measure in $\mathcal{M}_c(G)$ is approximated by others for which the stationary measure is stably unique and non-atomic.

In view of Remark 2.2.4, it is no restriction to suppose that $\theta \in (0, \theta_*]$. We do so in the remainder of Section 3.

3.3 Triangular case

First, we prove Theorem B when η^u is non-atomic. By Lemma 3.1.1, the number $\sigma = \int_G |\tau|^\theta \, d\nu$ is less than 1.

Proposition 3.3.1. *Assume that η^u is non-atomic. Then there exist $C_3 > 1$ and $\tau_3 \in (0,1)$ such that*

$$d_\theta(\eta^u, \mathcal{P}_\nu^n \eta') \leq C_3 \tau_3^n$$

for any $n \geq 1$, any non-atomic ν'-stationary measure η', and any $\nu' \in U$ whose support is contained in the $s\sigma^{n/2}$-neighborhood of $\operatorname{supp}\nu$.

Proof. Define

$$K_n = \Big\{(u,v) \in P \times P : u \notin E\big(\sigma^{n/2}\big) \text{ and } v \notin E\big(\sigma^{n/2}\big)\Big\}.$$

Then, since $d(\cdot,\cdot) \le 1$ and $\nu^{(n)}$ is a probability measure,

$$\begin{aligned}\int_{K_n^c \times G} d(gu,gv)^\theta \, d\nu^{(n)}(g)\, d\eta^u(u)\, d\eta'(v) &\le \big(\eta^u \times \eta'\big)\left(K_n^c\right) \\ &\le \eta^u\left(E\big(\sigma^{n/2}\big)\right) + \eta'\left(E\big(\sigma^{n/2}\big)\right) \le 2C_2\sigma^{\beta_2 n/2}\end{aligned} \tag{3.3.1}$$

(using (3.2.3) in the last step). Next we estimate the integral over $K_n \times G$.

For each $x \in M$, take $h(x) \in (0,+\infty)$ such that $(h(x),1) \in E^u_x$. Let $g = g_0$ denote the 0-th coordinate of x. Note that

$$h(f(x)) = \frac{\alpha(g)}{\beta(g)} h(x) + \frac{\gamma(g)}{\beta(g)} \text{ for } \mu\text{-almost every } x, \tag{3.3.2}$$

because the sub-bundle E^u is invariant, that is, $gE^u_x = E^u_{f(x)}$ for μ-almost every x. Write each $v \in P$ as $v = v_E(1,0) + v_u(h(x),1)$. The definition (3.1.3) gives

$$\begin{aligned} d(E^u_x, E) &= \frac{|(h(x),1) \wedge (1,0)|}{\|(h(x),1)\|\|(1,0)\|} = \frac{1}{\sqrt{h(x)^2+1}} \\ d(v,E) &= \frac{|(v_E + v_u h(x), v_u) \wedge (1,0)|}{\|(v_E + v_u h(x), v_u)\|\|(1,0)\|} = \frac{1}{\sqrt{(z_v + h(x))^2 + 1}}, \end{aligned} \tag{3.3.3}$$

where $z_v = v_E/v_u$. Therefore,

$$\begin{aligned} (E^u_x, v) \in K_n &\Rightarrow d(E^u_x, E) \ge \sigma^{n/2} \text{ and } d(v,E) \ge \sigma^{n/2} \\ &\Rightarrow |h(x)| \le \sigma^{-n/2} \text{ and } |z_v + h(x)| \le \sigma^{-n/2} \\ &\Rightarrow |z_v| \le 2\sigma^{-n/2}. \end{aligned} \tag{3.3.4}$$

Using (3.1.1) and (3.3.2), we see that

$$\begin{aligned} gv &= v_E(\alpha(g),0) + v_u(\alpha(g)h(x) + \gamma(g), \beta(g)) \\ &= v_E\alpha(g)(1,0) + v_u\beta(g)(h(f(x)),1), \end{aligned}$$

and so

$$\begin{aligned} d(gv, gE^u_x) &= \frac{|\big(v_E\alpha(g) + v_u\beta(g)h(f(x)), v_u\beta(g)\big) \wedge \big(h(f(x),1)\big)|}{\|(v_E\alpha(g) + v_u\beta(g)h(f(x)), v_u\beta(g))\|\|(h(f(x)),1)\|} \\ &= \frac{|\tau(g)z_v|}{\sqrt{(\tau(g)z_v + h(f(x)))^2+1}\sqrt{h(f(x))^2+1}} \le |\tau(g)z_v|. \end{aligned}$$

Define $\tilde{K}_n = \{(x,v) \in M \times P : (E^u_x, v) \in K_n\}$. Using Lemma 3.1.1 and (3.3.4),

$$\begin{aligned}\int_{K_n\times G} & d(gu,gv)^\theta \, d\nu^{(n)}(g)\, d\eta^u(u)\, d\eta'(v) \\ &= \int_{\tilde{K}_n\times G} d(gE^u_x, gv)^\theta \, d\nu^{(n)}(g)\, d\mu(x)\, d\eta'(v) \\ &\le \int_{\tilde{K}_n\times G} |\tau(g) z_v| \, d\nu^{(n)}(g)\, d\mu(x)\, d\eta'(v) \\ &\le \int_{\tilde{K}_n} \sigma^n |z_v| \, d\mu(x)\, d\eta'(v) \le 2\sigma^{n/2}.\end{aligned} \tag{3.3.5}$$

Adding (3.3.1) and (3.3.5), we find that

$$\int_{P\times P\times G} d(gu,gv)^\theta \, d\nu^{(n)}(g)\, d\eta^u(u)\, d\eta'(v) \le 2C_2\sigma^{\beta_2 n/2} + 2\sigma^{n/2}.$$

This proves the proposition, with $C_3 = 2C_2 + 2$ and $\tau_3 = \sigma^{\beta_2 n/2}$. □

Corollary 3.3.2. *$d_\theta(\eta^u, \mathcal{P}^n_\nu \eta'^u) \le C_3 \tau_3^n$ for any $n \ge 1$ and $\nu' \in U$ whose support is contained in the $s\sigma^{n/2}$-neighborhood of* $\operatorname{supp}\nu$.

Proof. Let ν' be as in Proposition 3.3.1. By Remark 3.2.2, we may find a sequence $(\nu'_k)_k \to \nu'$ such that the ν'_k-stationary measure η'_k is unique and non-atomic. Then, by Proposition 3.3.1,

$$d_\theta(\eta^u, \mathcal{P}^n_\nu \eta'_k) \le C_3 \tau_3^n \text{ for any } n \ge 1 \text{ and } k \ge 1.$$

As observed previously, it follows from [BV17] that $(\eta'_k)_k \to \eta'^u$. To conclude, just use the fact that $\mathcal{P}_\nu$ is continuous. □

Corollary 3.3.3. *There exist $C_4 > 1$ and $\beta_4 \in (0,1)$ such that*

$$d_\theta(\eta^u, \eta'^u) \le C_4 \delta_{\mathcal{T},\theta}(\nu,\nu')^{\beta_4} \textit{ for any } \nu' \in U.$$

Proof. By Lemmas 3.1.2 and 3.1.3, and the triangle inequality,

$$d_\theta(\mathcal{P}^n_\nu \eta'^u, \mathcal{P}^n_{\nu'} \eta'^u) \le \delta_\theta(\nu,\nu') + C_1 d_\theta(\mathcal{P}^{n-1}_\nu \eta'^u, \mathcal{P}^{n-1}_{\nu'} \eta'^u),$$

for every $n \ge 1$. So, by induction,

$$d_\theta(\mathcal{P}^n_\nu \eta'^u, \mathcal{P}^n_{\nu'} \eta'^u) \le \left[1 + C_1 + \cdots + C_1^{n-1}\right] \delta_\theta(\nu,\nu') \le C_1^n \delta_\theta(\nu,\nu') \tag{3.3.6}$$

(it is no restriction to suppose that $C_1 \ge 2$).

Fix $\zeta = \min\{\sigma^{1/2}, C_1^{-2}\}$ and let n be the smallest integer such that

$$\delta_{\mathcal{T},\theta}(\nu,\nu') < s\zeta^n. \tag{3.3.7}$$

From the definitions of $\delta_{\mathcal{T},\theta}$ and ζ, we see that ν' is contained in the $s\sigma^{n/2}$-neighborhood of $\operatorname{supp}\nu$. So, we may use Corollary 3.3.2 to conclude that

$$\begin{aligned} d_\theta(\eta^u, \eta'^u) = d_\theta(\mathcal{P}_\nu^n \eta^u, \mathcal{P}_{\nu'}^n \eta'^u) &\le d_\theta(\mathcal{P}_\nu^n \eta^u, \mathcal{P}_\nu^n \eta'^u) + d_\theta(\mathcal{P}_\nu^n \eta'^u, \mathcal{P}_{\nu'}^n \eta'^u) \\ &\le C_3 \tau_3^n + C_1^n \delta_\theta(\nu, \nu'). \end{aligned} \tag{3.3.8}$$

By the choice of n in (3.3.7),

$$C_3 \tau_3^n \le (C_3/\tau_3) \exp\left(\frac{\log \tau_3}{\log \zeta} \log \frac{\delta_{\mathcal{T},\theta}(\nu, \nu')}{s}\right) = C_4' \delta_{\mathcal{T},\theta}(\nu, \nu')^{\beta_4'},$$

with $\beta_4' = \log \tau_3 / \log \zeta$ and $C_4' = C_3/(\tau_3 s^{\beta_4'})$. Moreover,

$$C_1^n \delta_\theta(\nu, \nu') \le \zeta^{-n/2} \delta_{\mathcal{T},\theta}(\nu, \nu') \le s^{1/2} \delta_{\mathcal{T},\theta}(\nu, \nu')^{1/2}.$$

So, we may take $C_4 = C_4' + s^{1/2}$ and $\beta_4 = \min\{\beta_4', 1/2\}$. □

The conclusion of Theorem B in the present case follows immediately from Lemma 3.1.4 and Corollary 3.3.3.

3.4 Diagonal case

Now let us suppose that η^u does have some atom. In view of (3.2.1), this means that there exists $F \in P$ such that $M_F = \{x \in M : E_x^u = F\}$ has positive μ-measure. We claim that M_F must have full μ-measure, that is, $E_x^u = F$ for μ-almost every $x \in M$. That can be seen as follows.

Let $M_- = G^{\mathbb{Z}_-}$ and $\mu_- = \nu^{\mathbb{Z}_-}$. Since the unstable subspace E_x^u depends only on $x_- = (g_j)_{j<0}$, we may view M_F as a subset of M_- with positive μ_--measure. Then we may find $(h_{-n,n}, \dots, h_{-1,n}) \in G^n$ such that

$$\mu_- \left(M_F \mid g_{-n} = h_{-n,n}, \dots, g_{-1} = h_{-1,n}\right) \to 1 \text{ as } n \to \infty.$$

For each $n \ge 1$, let M_n be the set of all $(f_j)_{j<0}$ such that

$$(\dots, f_j, \dots, f_{-1}, h_{-n,n}, \dots, h_{-1,n}) \in M_F.$$

By invariance, $E_y^u = (h_{-1,n} \cdots h_{-n,n})^{-1} F$ for every $y \in M_n$. Moreover,

$$\mu_-(M_n) = \mu_- \left(M_F \mid g_{-n} = h_{-n,n}, \dots, g_{-1} = h_{-1,n}\right) \to 1 \text{ as } n \to \infty.$$

In particular, M_n must intersect M_F for every large n. Consequently,

$$E_y^u = (h_{-1,n} \cdots h_{-n,n})^{-1} F = F$$

for every $y \in M_n$ and every large n. That proves the claim.

Recall that we are assuming that there exists $E \in P$ such that $gE = E$ for every $g \in \operatorname{supp}\nu$. We have just shown that if η^u is atomic then there exists $F \neq E$ such that $\eta^u = \delta_F$ and $gF = F$ for every $g \in \operatorname{supp} \nu$. It is no restriction to suppose that $E = \mathbb{R} \times \{0\}$ and $F = \{0\} \times \mathbb{R}$. Then (3.1.1) becomes

$$g = \begin{pmatrix} \alpha(g) & 0 \\ 0 & \beta(g) \end{pmatrix}. \tag{3.4.1}$$

In this setting we are able to prove a stronger version of Proposition 3.3.1:

Proposition 3.4.1. *There exist $\tau_5 < 1$ and a neighborhood $U_5 \subset \mathcal{M}_c(G)$ of ν such that*

$$d_\theta(\eta^u, \mathcal{P}_\nu \eta') \leq \tau_5 d_\theta(\eta^u, \eta')$$

any non-atomic ν'-stationary measure η' and any $\nu' \in U_5$.

Proof. Each $v \in P$ may be written as $v = v_E(1,0) + v_F(0,1)$. Then (3.1.3) gives

$$d(v, F) = \frac{|(v_E, v_F) \wedge (0,1)|}{\|(v_E, v_F)\| \|(0,1)\|} = \frac{1}{\sqrt{1 + z_v^{-2}}} \text{ and}$$
$$d(gv, F) = \frac{|(\alpha(g)v_E, \beta(g)v_F) \wedge (0,1)|}{\|(\alpha(g)v_E, \beta(g)v_F)\| \|(0,1)\|} = \frac{1}{\sqrt{1 + (\tau(g) z_v)^{-2}}}$$

with $z_v = v_E/v_F$ and $\tau(g) = \alpha(g)/\beta(g)$. Note that $z_F = 0$ and $z_E = \infty$.

Given any measure η on P, there exists a unique coupling of η and the Dirac mass $\eta^u = \delta_F$, namely, their product. Thus (3.1.4) gives

$$\begin{aligned} d_\theta(\eta^u, \eta) &= \int_P d(v, F)^\theta \, d\eta(v) \text{ and} \\ d_\theta(\eta^u, \mathcal{P}_\nu \eta) &= \int_{P \times G} d(gv, F)^\theta \, d\nu(g) \, d\eta(v) \end{aligned} \tag{3.4.2}$$

for any probability measure η on P.

Lemma 3.4.2. *There exists a continuous function $\rho : P \to (0, 1]$ such that*

$$\int_G d(gv, F)^\theta \, d\nu(g) \leq \rho(v) \, d(v, F)^\theta \textit{ for every } v \in P$$

and $\rho(v) < 1$ for any $v \neq E$.

Proof. Define

$$\rho(v) = \frac{\int_G d(gv, F)^\theta \, d\nu(g)}{d(v, F)} \text{ for } v \neq F.$$

It is clear that ρ is continuous and satisfies the inequality in the statement.

Note that $\rho(E) = 1$. To prove that $\rho(v) < 1$ for all other v, consider

$$\Phi : [0, +\infty) \to \mathbb{R}, \quad \Phi(z) = \left(\frac{1}{\sqrt{1 + z^{-2/\theta}}} \right)^\theta.$$

It is straightforward to check that Φ is increasing and concave. By the Jensen inequality and (3.1.2), it follows that

$$\int_G d(gv,F)^\theta \, d\nu(g) = \int_G \Phi(|\tau(g)z_v|^\theta) \, d\nu(g) \le \Phi\left(\int_G |\tau(g)z_v|^\theta \, d\nu(g)\right)$$
$$< \Phi(|z_v|^\theta) = d(v,F) \text{ for any } v \neq E, F.$$

Finally, $\Phi(z) \approx z$ for $z > 0$ close to zero, where $\approx$ means that the quotient goes to 1 as z goes to 0. So,

$$d(v,F)^\theta = \Phi(|z_v|^\theta) \approx |z_v|^\theta \text{ and}$$
$$\int_G d(gv,F)^\theta \, d\nu(g) = \int_G \Phi(|\tau(g)z_v|^\theta) \, d\nu(g) \approx |z_v|^\theta \int_G |\tau(g)|^\theta \, d\nu(g).$$

This ensures that $\rho(v)$ extends continuously to F with

$$\rho(F) = \int_G |\tau(g)|^\theta \, d\nu(g) < 1.$$

That completes the proof. □

Note that $d(v,F) \ge 1 - r$ for every $v \in E(r)$, because $d(E,F) = 1$ and $E(r)$ is the neighborhood of radius r around E.

Lemma 3.4.3. *There exist $\epsilon_5 > 0$ and a neighborhood $U_5 \subset \mathcal{M}_c(G)$ of ν such that*

$$\int_{P \setminus E(\epsilon_5)} d(v,F)^\theta \, d\eta'(v) > \int_{E(\epsilon_5)} d(v,F)^\theta \, d\eta'(v).$$

for any non-atomic ν'-stationary measure η' and any $\nu' \in U_5$.

Proof. It follows from Theorem 3.2.1 that,

$$\eta'\big(E(q^2\epsilon)\big) < \frac{4}{9}\eta'\big(E(\epsilon)\big) \text{ and so } \eta'\big(E(q^2\epsilon)\big) < \frac{4}{5}\eta'\big(E(\epsilon) \setminus E(q^2\epsilon)\big),$$

for any $\epsilon \in (0, \epsilon_0)$, any non-atomic ν'-stationary measure η' and any $\nu' \in U$ whose support is contained in the $sq\epsilon$-neighborhood of $\operatorname{supp}\nu$. Fix ϵ such that $(1-\epsilon)^\theta > 4/5$, take $\epsilon_5 = q^2\epsilon$ and let U_5 be the intersection of U with the $sq\epsilon$-neighborhood of ν relative to the distance $\delta_{\mathcal{T},\theta}$.

On the one hand,

$$\int_{E(\epsilon_5)} d(v,F)^\theta \, d\eta'(v) \le \eta'(E(\epsilon_5)) < \frac{4}{5}\eta'\left(E(\epsilon) \setminus E(\epsilon_5)\right).$$

On the other hand,

$$\int_{P \setminus E(\epsilon_5)} d(v,F)^\theta \, d\eta'(v) \ge \int_{E(\epsilon) \setminus E(\epsilon_5)} d(v,F)^\theta \, d\eta'(v)$$
$$\ge \int_{E(\epsilon) \setminus E(\epsilon_5)} (1-\epsilon)^\theta \, d\eta'(v) > \frac{4}{5}\eta'\left(E(\epsilon) \setminus E(\epsilon_5)\right).$$

These two observations yield the claim. □

Let us conclude the proof of Proposition 3.4.1. Let $\rho : P \to (0, 1]$ be as in Lemma 3.4.2. Then

$$d_\theta(\eta^u, \mathcal{P}_\nu \eta) = \int_P \int_G d(gv, F)^\theta \, d\nu(g) \, d\eta(v) \le \int_P \rho(v) d(v, F)^\theta \, d\eta'(v).$$

Fix $\delta_5 > 0$ such that $\rho(v) \le 1 - 2\delta_5$ for $v \in G \setminus E(\epsilon_5)$. Then, using also Lemma 3.4.3,

$$\begin{aligned}
&\int_P \rho(v) d(v, F)^\theta \, d\eta'(v) \\
&\qquad = \int_{P \setminus E(\epsilon_5)} \rho(v) d(v, F)^\theta \, d\eta'(v) + \int_{E(\epsilon_5)} \rho(v) d(v, F)^\theta \, d\eta'(v) \\
&\qquad \le (1 - 2\delta_5) \int_{P \setminus E(\epsilon_5)} d(v, F)^\theta \, d\eta'(v) + \int_{E(\epsilon_5)} d(v, F)^\theta d\eta'(v) \\
&\qquad \le (1 - \delta_5) \int_P d(v, F)^\theta \, d\eta'(v).
\end{aligned}$$

This proves the claim, with $\tau_5 = 1 - \delta_5$. □

Corollary 3.4.4. *$d_\theta(\eta^u, \mathcal{P}_\nu \eta'^u) \le \tau_5 d_\theta(\eta^u, \eta'^u)$ for any $\nu' \in U_5$.*

Proof. Let ν' be as in Proposition 3.4.1. By Remark 3.2.2, we may find a sequence $(\nu'_k)_k \to \nu'$ such that the ν'_k-stationary measure η'_k is unique and non-atomic. Then, Proposition 3.4.1 gives that $d_\theta(\eta^u, \mathcal{P}_\nu \eta'_k) \le \tau_5 d(\eta^u, \eta'_k)$ for every $k \ge 1$. Recall that $(\eta'_k)_k \to \eta'^u$ as $k \to \infty$, by [BV17]. □

Corollary 3.4.5. *There exists $C_6 > 1$ such that*

$$d_\theta(\eta^u, \eta'^u) \le C_6 \delta_\theta(\nu, \nu') \le C_6 \delta_{\mathcal{T},\delta}(\nu, \nu') \text{ for any } \nu' \in U_5.$$

Proof. Combining Lemma 3.1.2 and Corollary 3.4.4,

$$d_\theta(\eta^u, \eta'^u) \le d_\theta(\eta^u, \mathcal{P}_\nu \eta'^u) + d_\theta(\mathcal{P}_\nu \eta'^u, \mathcal{P}_{\nu'} \eta'^u) \le \tau_5 d_\theta(\eta^u, \eta'^u) + \delta_\theta(\nu, \nu'),$$

and so,

$$d_\theta(\eta^u, \eta'^u) \le \frac{1}{1 - \tau_5} \delta_\theta(\nu, \nu').$$

Take $C_6 = 1/(1 - \tau_5)$. □

From Lemma 3.1.4 and Corollary 3.4.5 we get the conclusion of Theorem B in this second and final case.

Chapter 4

Proof of Theorem C

Since Hölder continuity is stronger than log-Hölder continuity, at this point we only need to consider the cases when $\lambda_-(\nu) = \lambda_+(\nu)$. By the invariance principle [BGMV03, AV10], it follows that all $\mathbb{P}F$-invariant probabilities are of the form $\mu \times \eta$. See [Via14, Chapter 5]. A probability measure η on P is ν-stationary if and only if $\mu \times \eta$ is $\mathbb{P}F$-invariant, and this happens if and only if η is invariant under every $g \in \operatorname{supp} \nu$. So, we have the following cases:

- *Conformal*: there exists some stationary measure η with no atoms of mass larger than or equal to 1/2; then there exists a conformal structure on $\mathbb{R}^2$ relative to which every $g \in \operatorname{supp} \nu$ is a conformal map.

- *Degenerate diagonal*: there exist exactly two subspaces, E and F, that are invariant under every $g \in \operatorname{supp} \nu$, and the stationary measures are the convex combinations $a\delta_E + (1-a)\delta_F$ of the corresponding Dirac masses.

- *Simply reducible*: there exists a pair $\{E, F\}$ of subspaces with $gE = F$ and $gF = E$ for every $g \in \operatorname{supp} \nu$, and the average $(\delta_E + \delta_F)/2$ of the Dirac masses is the only stationary measure.

- *Degenerate triangular*: there exists exactly one subspace, E, invariant under every $g \in \operatorname{supp} \nu$, and the Dirac mass δ_E is the unique stationary measure.

The conformal case follows from rather elementary arguments combined with a classical observation of Douady–Earle [DE86], as we explain in Section 4.1. Actually, in this situation the measure ν is still a point of local Hölder (rather than log-Hölder) continuity, as in the setting of the previous section.

The other three cases are handled in Sections 4.2 through 4.5. As we outlined in Section 2.2, the main point is to prove a power-law "diffusion" property, which is stated in (4.2.6)–(4.2.8) in Section 4.2. In that same section we explain how to deduce the conclusion of Theorem C from that property. In Sections 4.3 we check that (4.2.6)–(4.2.8) do hold in each of the three cases.

Let $\theta \in (0, 1]$ be fixed in all that follows.

4.1 Conformal case

Here we suppose that ν admits a stationary measure η such that $\eta(\{u\}) < 1/2$ for every $u \in P$. We prove that ν is a point of Lipschitz continuity for λ_+.

Proposition 4.1.1. *There exists an isomorphism $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\phi^{-1} \circ g \circ \phi$ is a conformal map, relative to Euclidean metric, for every $g \in \operatorname{supp} \nu$.*

Proof. Let us associate to each

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

in GL(2) the corresponding Möbius automorphism

$$\tilde{g} : \mathbb{H} \to \mathbb{H}, \quad \tilde{g}(z) = \frac{az + b}{cz + d}$$

of the hyperbolic upper half plane $\mathbb{H}$. Note that $\tilde{g}(i) = i$ if and only if $a = d$ and $b = -c$, that is, if and only if g is conformal. The action $[x, y] \mapsto [ax + by, cx + dy]$ of g on the projective space $\mathbb{P}(\mathbb{R}^2)$ corresponds to the restriction of $\tilde{g}$ to the real axis, through the identification

$$\mathbb{P}(\mathbb{R}^2) \to \partial\mathbb{H}, \quad [x, y] \mapsto \frac{x}{y}.$$

Thus we may view η as a measure on $\partial\mathbb{H}$ invariant under $\tilde{g}$ for every $g \in \operatorname{supp} \nu$.

The *conformal barycenter* construction of Douady-Earle [DE86] assigns a point $B(\xi) \in \mathbb{H}$ to each probability measure ξ on $\partial\mathbb{H}$ with $\xi(\{u\}) < 1/2$ for every u, in such a way that

$$B(\tilde{h}_* \xi) = \tilde{h}(B(\xi)) \text{ for any Möbius automorphism } \tilde{h} : \mathbb{H} \to \mathbb{H}. \tag{4.1.1}$$

In particular, the fact that η is invariant under $\tilde{g}$ implies that $B(\eta)$ is a fixed point of $\tilde{g}$ for every $g \in \operatorname{supp} \nu$. Fix $\phi \in \mathrm{GL}(2)$ such that $\tilde{\phi}(i) = B(\eta)$. Then i is a fixed point of $\tilde{\phi}^{-1} \circ \tilde{g} \circ \tilde{\phi}$ for every $g \in \operatorname{supp} \nu$, as claimed. □

Thus it is no restriction to assume that every $g \in \operatorname{supp} \nu$ is conformal or, equivalently, that it satisfies $\|g\| = \|g^{-1}\|$ for every $g \in \operatorname{supp} \nu$. Then

$$\int_G \log \|g\| \, d\nu(g) = \lambda_+(\nu) = \lambda_-(\nu) = -\int_G \log \|g^{-1}\| \, d\nu(g). \tag{4.1.2}$$

In general,

$$\int_G \log\|g\|\,d\nu'(g) \geq \lambda_+(\nu') \geq \lambda_-(\nu') \geq -\int_G \log\|g^{-1}\|\,d\nu'(g). \tag{4.1.3}$$

Let K be a compact neighborhood of the support of ν. Since $g \mapsto \log\|g^{\pm1}\|$ are locally Lipschitz functions, there exists $L > 0$ such that

$$|\log\|g_1\| - \log\|g_2\|| \leq L\delta(g_1, g_2)^\theta \text{ and } \left|\log\|g_1^{-1}\| - \log\|g_2^{-1}\|\right| \leq L\delta(g_1, g_2)^\theta$$

for any $g_1, g_2 \in K$. Then

$$\begin{aligned} \left|\int_G \log\|g\|\,d\nu(g) - \int_G \log\|g\|\,d\nu'(g)\right| &\leq L\delta_\theta(\nu,\nu') \text{ and} \\ \left|\int_G \log\|g^{-1}\|\,d\nu(g) - \int_G \log\|g^{-1}\|\,d\nu'(g)\right| &\leq L\delta_\theta(\nu,\nu') \end{aligned} \tag{4.1.4}$$

for any ν' close enough to ν that $\operatorname{supp}\nu' \subset K$. From (4.1.2)–(4.1.4) we get that

$$\left|\lambda_+(\nu') - \lambda_+(\nu)\right| \leq L\delta_\theta(\nu,\nu') \text{ and } \left|\lambda_-(\nu') - \lambda_-(\nu)\right| \leq L\delta_\theta(\nu,\nu'),$$

which implies the claim in Theorem C in the conformal case.

4.2 Power-law diffusion

Recall that the space $\mathrm{Stat}(\nu)$ of ν-stationary measures coincides with the set $[\delta_E, \delta_F]$ of linear combinations of δ_E and δ_F when ν is degenerate diagonal. Moreover, $\mathrm{Stat}(\nu) = \{\delta_E/2 + \delta_F/2\}$ if ν is simply reducible and $\mathrm{Stat}(\nu) = \{\delta_E\}$ if ν is degenerate triangular. It is no restriction to assume that the subspaces E and F are, respectively, the horizontal axis and the vertical axis, so that $z_E = \infty$ and $z_F = 0$.

The next lemma provides some useful bounds on the distance from $\mathrm{Stat}(\nu)$ to any $\zeta \in \mathcal{M}(P)$:

Lemma 4.2.1. *For any $\zeta \in \mathcal{M}(P)$ and $R > 0$,*

$$d_\theta([\delta_E, \delta_F], \zeta) \leq \zeta\big(\{v \in P : e^{-R} \leq |z_v| \leq e^R\}\big) + e^{-R\theta} \tag{4.2.1}$$

$$\begin{aligned} d_\theta(\delta_E/2 + \delta_F/2, \zeta) \leq{}& \zeta\left(\{v \in P : e^{-R} \leq |z_v| \leq e^R\}\right) + e^{-R\theta} \\ &+ \left|\zeta(\{v \in P : |z_v| > e^R\} - \zeta(\{v \in P : |z_v| < e^{-R}\}\right| \end{aligned} \tag{4.2.2}$$

$$d_\theta(\delta_E, \zeta) \leq \zeta\big(\{v \in P : -R \leq z_v \leq R\}\big) + R^{-\theta}. \tag{4.2.3}$$

Proof. To prove (4.2.1) let $\eta = a\delta_E + (1-a)\delta_F$ with $a = \zeta\big(\{|z_v| > e^R\}\big)$. It follows from the definition that

$$d_\theta([\delta_E, \delta_F], \zeta) \leq d_\theta(\eta, \zeta) = \sup_\varphi \left|\int_P \varphi\,d(\zeta - \eta)\right|$$

where the supremum is over all $\varphi \in L_\theta(P)$ such that $\varphi(F) = 0$. Note that $|\varphi| \le 1$ and $\left|\int_P \varphi \, d(\zeta - \eta)\right|$ is bounded by

$$\int_{\{e^{-R} \le |z_v| \le e^R\}} |\varphi| \, d\zeta + \int_{\{|z_v| < e^{-R}\}} |\varphi| \, d\zeta + \int_{\{|z_v| > e^R\}} |\varphi - \varphi(E)| \, d\zeta. \quad (4.2.4)$$

The first term is bounded by $\zeta\big(\{e^{-R} \le |z_v| \le e^R\}\big)$. By definition,

$$|\varphi(v) - \varphi(E)| \le d(v, E)^\theta = \left(\frac{1}{\sqrt{z_v^2 + 1}}\right)^\theta < e^{-R\theta} \text{ if } |z_v| > e^R$$

$$\text{and } |\varphi(v)| \le d(v, F)^\theta = \left(\frac{1}{\sqrt{1 + z_v^{-2}}}\right)^\theta < e^{-R\theta} \text{ if } |z_v| < e^{-R}.$$

So the sum of the two last terms in (4.2.4) is bounded by e^{-R}.

Next we prove (4.2.2). It follows from the definition that

$$d_\theta\left(\delta_E/2 + \delta_F/2, \zeta\right) = \sup_\varphi \left|\int_P \varphi \, d\zeta\right|$$

where the supremum is over all $\varphi \in L_\theta(P)$ with $\varphi(E) + \varphi(F) = 0$. Note that $|\varphi| \le 1$ and $\left|\int_P \varphi \, d\zeta\right|$ is bounded by

$$\int_{\{e^{-R} \le |z_v| \le e^R\}} |\varphi| \, d\zeta + |\varphi(E)| \left|\zeta(\{|z_v| > e^R\} - \zeta(\{|z_v| < e^{-R}\}\right|$$
$$+ \int_{\{|z_v| > e^R\}} |\varphi - \varphi(E)| \, d\zeta + \int_{\{|z_v| < e^{-R}\}} |\varphi - \varphi(F)| \, d\zeta. \quad (4.2.5)$$

The first term is bounded by $\zeta\big(\{e^{-R} \le |z_v| \le e^R\}\big)$, and the second one is bounded by $\left|\zeta(\{|z_v| > e^R\} - \zeta(\{|z_v| < e^{-R}\}\right|$. Since

$$|\varphi(v) - \varphi(E)| < e^{-R\theta} \text{ if } |z_v| > e^R \text{ and } |\varphi(v) - \varphi(F)| < e^{-R\theta} \text{ if } |z_v| < e^{-R},$$

the sum of the last two terms in (4.2.5) is bounded by $e^{-R\theta}$.

Finally, we prove (4.2.3). It follows from the definition that

$$d_\theta(\delta_E, \zeta) = \sup_\varphi \left|\int_P \varphi \, d\zeta\right|$$

where the supremum is over all $\varphi \in L_\theta(P)$ with $\varphi(E) = 0$. Since $|\varphi| \le 1$ and

$$|\varphi(v)| \le d(v, E)^\theta < R^{-\theta} \text{ if } |z_v| > R,$$

we get that

$$\left|\int_P \varphi \, d\zeta\right| \le \int_{\{-R \le z_v \le R\}} |\varphi| \, d\zeta + \int_{\{|z_v| > R\}} |\varphi| \, d\zeta \le \zeta\left(\{-R \le z_v \le R\}\right) + R^{-\theta}.$$

This completes the proof of the lemma. □

We are going to prove that there exist $D_0 > 0$ and $\beta_0 > 0$ such that some of the following conditions are satisfied for any $\xi \in \mathcal{M}(P)$, $n \geq 1$ and $R \geq 3$:

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : e^{-R} \leq |z_v| \leq e^R\}\big) \leq D_0 R n^{-\beta_0} \tag{4.2.6}$$

$$\big|\mathcal{P}_\nu^n \xi\big(\{v \in P : |z_v| < e^{-R}\}\big) - \mathcal{P}_\nu^n \xi\big(\{v \in P : |z_v| > e^R\}\big)\big| \leq D_0 R n^{-\beta_0} \tag{4.2.7}$$

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \leq z_v \leq R\}\big) \leq D_0 R n^{-\beta_0}. \tag{4.2.8}$$

More precisely, apart from a few special sub-cases that are handled separately, we have (4.2.6) if ν is degenerate diagonal, (4.2.8) if ν is degenerate triangular, and both (4.2.6) and (4.2.7) if ν is simply reducible. We postpone the verification of these facts to Sections 4.3 through 4.5, and proceed directly to check that they imply the conclusion of Theorem C.

Proposition 4.2.2. *There exist $D_1 > 0$ and $\beta_1 > 0$ such that*

$$d_\theta(\mathrm{Stat}(\nu), \mathcal{P}_\nu^n \eta') \leq D_1 n^{-\beta_1} \text{ for every } n \geq 1 \text{ and } \eta' \in \mathcal{M}(P).$$

Proof. In the degenerate diagonal case, use (4.2.1) and (4.2.6) with $\zeta = \mathcal{P}_\nu^n \eta'$ and $R = (\beta_0/\theta)\log n$ to find that

$$d_\theta(\mathrm{Stat}(\nu), \mathcal{P}_\nu^n \eta') \leq D_0 R n^{-\beta_0} + e^{-R\theta} = (D_0\beta_0/\theta) n^{-\beta_0}\log n + n^{-\beta_0}.$$

The right-hand side is bounded by $D_1 n^{-\beta_1}$ for every $n \geq 1$, as long as $\beta_1 < \beta_0$ and D_1 is large enough.

In the simply reducible case, use (4.2.2), (4.2.6) and (4.2.7) with $\zeta = \mathcal{P}_\nu^n \eta'$ and $R = (\beta_0/\theta)\log n$ to get that

$$d_\theta(\mathrm{Stat}(\nu), \mathcal{P}_\nu^n \eta') \leq D_0 R n^{-\beta_0} + e^{-R\theta} + D_0 n^{-\beta_0} \leq D_1 n^{-\beta_1}$$

if $\beta_1 < \beta_0$ and D_1 is large enough.

In the degenerate triangular case, use (4.2.3) and (4.2.8) with $\zeta = \mathcal{P}_\nu^n \eta'$ and $R = n^{\beta_0/(1+\theta)}$ to conclude that

$$d_\theta(\mathrm{Stat}(\nu), \mathcal{P}_\nu^n \eta') \leq D_0 R n^{-\beta_0} + R^{-\theta} = D_1 n^{-\beta_1},$$

where $D_1 = D_0 + 1$ and $\beta_1 = \beta_0\theta/(1+\theta)$.

In all the cases, we have assumed that n is large enough that the relevant conditions (4.2.6)–(4.2.8) are valid for the corresponding value of R. This is no restriction because small values of n may be dealt with by increasing D_1. □

Corollary 4.2.3. *There exists $D_2 > 0$ such that*

$$d_\theta(\mathrm{Stat}(\nu), \eta') \le D_2 \left(\log \frac{1}{\delta_\theta(\nu, \nu')} \right)^{-\beta_1}.$$

for any ν' close to ν and any $\eta' \in \mathrm{Stat}(\nu')$.

Proof. Combining Lemmas 3.1.2 and 3.1.3, we find that

$$\begin{aligned} d_\theta(\mathcal{P}_\nu^n \eta', \mathcal{P}_{\nu'}^n \eta') &\le d_\theta(\mathcal{P}_\nu \mathcal{P}_\nu^{n-1} \eta', \mathcal{P}_{\nu'} \mathcal{P}_\nu^{n-1} \eta') + d_\theta(\mathcal{P}_{\nu'} \mathcal{P}_\nu^{n-1} \eta', \mathcal{P}_{\nu'} \mathcal{P}_{\nu'}^{n-1} \eta') \\ &\le \delta_\theta(\nu, \nu') + C_1 d_\theta(\mathcal{P}_\nu^{n-1} \eta', \mathcal{P}_{\nu'}^{n-1} \eta'), \end{aligned}$$

for every $n \ge 1$. By induction, it follows that

$$d_\theta(\mathcal{P}_\nu^n \eta', \mathcal{P}_{\nu'}^n \eta') \le \delta_\theta(\nu, \nu') \left[1 + C_1 + \cdots + C_1^{n-1}\right] \le C_1^n \delta_\theta(\nu, \nu')$$

(it is no restriction to suppose that $C_1 \ge 2$). Then, using Proposition 4.2.2,

$$\begin{aligned} d_\theta(\mathrm{Stat}(\nu), \eta') &\le d_\theta(\mathrm{Stat}(\nu), \mathcal{P}_\nu^n \eta') + d_\theta(\mathcal{P}_\nu^n \eta', \mathcal{P}_{\nu'}^n \eta') \\ &\le D_1 n^{-\beta_1} + C_1^n \delta_\theta(\nu, \nu') \end{aligned}$$

for $\eta' \in \mathrm{Stat}(\nu')$ and $n \ge 1$. Let $n \ge 1$ be smallest such that $C_1^n \delta_\theta(\nu, \nu') > 2^{-n}$, that is, $n > \log\left(1/\delta_\theta(\nu, \nu')\right)/\log(2C_1)$. Then

$$C_1^n \delta_\theta(\nu, \nu') \le (2C_1) 2^{-n} \le (2C_1) n^{-\beta_1}.$$

Replacing these estimates in the previous inequality,

$$d_\theta(\mathrm{Stat}(\nu), \eta') \le (D_1 + 2C_1) n^{-\beta_1} \le D_2 \left(\log \frac{1}{\delta_\theta(\nu, \nu')} \right)^{-\beta_1}$$

with $D_2 = (D_1 + 2C_1) \log(2C_1)^{\beta_1}$. □

In view of Lemma 3.1.4, Corollary 4.2.3 yields the conclusion of Theorem C.

All that is left now is to check (4.2.6)–(4.2.8) in each of the three cases. In the process a few special sub-cases arise that are dealt with separately.

4.3 Degenerate diagonal case

Take the measure ν to be degenerate diagonal. The setting and terminology are as in Section 3.4, except that presently we have $\lambda_-(\nu) = \lambda_+(\nu)$, that is,

$$\int_G \log |\tau(g)| \, d\nu(g) = 0. \tag{4.3.1}$$

Denote $T(g) = \log |\tau(g)| = \log |\alpha(g)/\beta(g)|$ and then define

$$\sigma^2 = \int_G T(g)^2 \, d\nu(g) \text{ and } \rho = \int_G T(g)^3 \, d\nu(g).$$

If $\sigma = 0$ then $|\alpha(g)| = |\beta(g)|$ and so g is conformal, for every $g \in \operatorname{supp} \nu$. That case was already treated in Section 4.1. In the remainder of this section we assume that $\sigma > 0$.

Every $g \in \operatorname{supp} \nu$ has the form (3.4.1), and so its action on P is given by $\log |z_v| \mapsto \log |z_v| + T(g)$. Then the action of $g_{n-1} \cdots g_0$ is given by

$$\log |z_v| \mapsto \log |z_v| + T(g_{n-1} \cdots g_0) = \log |z_v| + \sum_{i=0}^{n-1} T(g_i). \tag{4.3.2}$$

For each $a < b$ and $n \geq 1$ define

$$\Phi_n(a, b) = \mu \left(\left\{ (g_j)_j \in G^{\mathbb{Z}} : \frac{1}{\sigma \sqrt{n}} \sum_{i=0}^{n-1} T(g_i) \in [a, b] \right\} \right)$$

and

$$\mathcal{N}(a, b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} \, dt. \tag{4.3.3}$$

The central limit theorem asserts that $\Phi_n(a, b) \to \mathcal{N}(a, b)$ when $n \to \infty$. In fact, by the Berry–Esseen theorem (see [Dur10, Theorem 3.4.9] and Petrov [Pet87]),

$$|\Phi_n(a, b) - \mathcal{N}(a, b)| \leq \frac{\rho}{\sigma^3 \sqrt{n}} \text{ for every } n \geq 1 \text{ and any } a < b. \tag{4.3.4}$$

Property (4.2.6) is contained in the next proposition, which is a sort of prototype for the rather more elaborate arguments we will use in the simply reducible and degenerate triangular cases:

Proposition 4.3.1. *There is $D_3 > 0$ such that for any $\xi \in \mathcal{M}(P)$*

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \leq \log |z_v| \leq R\}\big) \leq \frac{D_3 R}{\sqrt{n}} \text{ for every } n \geq 1 \text{ and } R \geq 1.$$

Proof. Since $\mathcal{P}_\nu^n \xi = \int_P \mathcal{P}_\nu^n \delta_u \, d\xi(u)$, it suffices to consider the case when $\xi = \delta_u$ for some $u \in P$. It is also useful to keep in mind the following immediate consequence of the definitions: for any measurable set $Z \subset P$,

$$\begin{aligned} \mathcal{P}_\nu^n \delta_u(Z) &= \nu^{(n)} \left(\{ g \in G : gu \in Z \} \right) \\ &= \mu \left(\left\{ (g_j)_j \in G^{\mathbb{Z}} : g_{n-1} \cdots g_0 u \in Z \right\} \right). \end{aligned}$$

This also shows that, since the lines $E, F \in P$ are fixed under every $g \in \operatorname{supp}\nu$, it suffices to consider $u \in P \setminus \{E, F\}$. Then, recalling (4.3.2),

$$\begin{aligned}
&\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \leq \log|z_v| \leq R\}\big) \\
&\qquad\qquad = \mu\left(\left\{(g_j)_j \in G^{\mathbb{Z}} : \log|z_u| + \sum_{i=0}^{n-1} T(g_i) \in [-R, R]\right\}\right) \\
&\qquad\qquad = \Phi_n\left(\frac{-R - \log|z_u|}{\sigma\sqrt{n}}, \frac{R - \log|z_u|}{\sigma\sqrt{n}}\right).
\end{aligned}$$

Then, noting that $\mathcal{N}(a, b) < (b - a)$ for every $a < b$,

$$\begin{aligned}
\mathcal{P}_\nu^n \xi\big(\{-R \leq \log|z_v| \leq R\}\big) &\leq \mathcal{N}\left(\frac{-R - \log|z_u|}{\sigma\sqrt{n}}, \frac{R - \log|z_u|}{\sigma\sqrt{n}}\right) + \frac{\rho}{\sigma^3\sqrt{n}} \\
&\leq \frac{2R}{\sigma\sqrt{n}} + \frac{\rho}{\sigma^3\sqrt{n}} \leq \left(\frac{2\sigma^2 + \rho}{\sigma^3}\right)\frac{R}{\sqrt{n}}.
\end{aligned}$$

This proves the claim, with $D_3 = (2\sigma^2 + \rho)/(\sigma^3)$. □

4.4 Simply reducible case

Take the measure ν to be simply reducible: there is a pair $E, F \in P$ such that $g(\{E, F\}) = \{E, F\}$ for every $g \in \operatorname{supp}\nu$ but neither E nor F are invariant under every g in the support of ν. This implies that $\lambda_-(\nu) = \lambda_+(\nu)$. We are going to explain how the arguments of Section 4.3 can be extended to this case.

Since we take E to be the horizontal axis and F to be the vertical axis, every $g \in \operatorname{supp}\nu$ has one of the following forms:

$$g = \begin{pmatrix} \alpha(g) & 0 \\ 0 & \beta(g) \end{pmatrix} \text{ or } g = \begin{pmatrix} 0 & \beta(g) \\ \alpha(g) & 0 \end{pmatrix}. \tag{4.4.1}$$

Define $\kappa(g) = 0$ in the first case and $\kappa(g) = 1$ in the second one. Then, in both cases,

$$g = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\kappa(g)} \begin{pmatrix} \alpha(g) & 0 \\ 0 & \beta(g) \end{pmatrix} \tag{4.4.2}$$

and so the action of g on P corresponds to

$$\log|z_v| \mapsto (-1)^{\kappa(g)} \left(\log|z_v| + T(g)\right),$$

with $T(g) = \log|\tau(g)| = \log|\alpha(g)/\beta(g)|$. It follows that the action of $g_{n-1} \cdots g_0$ on P is given by

$$\log|z_v| \mapsto (-1)^{\kappa_n} \left(\log|z_v| + \sum_{i=0}^{n-1} (-1)^{\kappa_i} T(g_i)\right) \tag{4.4.3}$$

where $\kappa_j = \kappa(g_0) + \cdots + \kappa(g_{j-1})$ for $j \geq 1$ and $\kappa_0 = 0$.

Up to conjugating the measure ν (Remark 2.2.3), we may assume that

$$\int_G T \, d\nu = 0. \tag{4.4.4}$$

Indeed, consider $t > 0$ and $h : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (tx, t^{-1}y)$. For g as in (4.4.2), the action of hgh^{-1} on P is given by $z_v \mapsto (-1)^{\hat{\kappa}(hgh^{-1})}(\log |z_v| + \hat{T}(hgh^{-1}))$ with

$$\hat{\kappa}(hgh^{-1}) = \kappa(g) \text{ and } \hat{T}(hgh^{-1}) = \begin{cases} T(g) & \text{if } \kappa(g) = 0 \\ T(g) + 2\log t & \text{if } \kappa(g) = 1. \end{cases}$$

So, since $\nu(\{\kappa = 1\}) > 0$, we may always choose t in such a way that the expected value of $\hat{T}$ is equal to zero.

Let us also get rid of a couple special sub-cases that are covered by situations we treated previously. Suppose that T is identically zero, that is, $|\alpha| = |\beta|$ on the support of ν. Then every $g \in \operatorname{supp} \nu$ is conformal and we already know (Section 4.1) that Theorem C holds in that case. Next, define

$$\Delta = \int_G (-1)^\kappa \, d\nu = \nu\left(\{g \in G : \kappa(g) = 0\}\right) - \nu\left(\{g \in G : \kappa(g) = 1\}\right). \tag{4.4.5}$$

By assumption, $\Delta \in [-1, 1)$. Suppose that $\Delta = -1$, that is,

$$g = \begin{pmatrix} 0 & \beta(g) \\ \alpha(g) & 0 \end{pmatrix} \text{ for every } g \in \operatorname{supp} \nu.$$

Then every h in the support $\operatorname{supp}(\nu * \nu)$ of the 2-convolution of ν has the form,

$$h = \begin{pmatrix} \hat{\alpha}(h) & 0 \\ 0 & \hat{\beta}(h) \end{pmatrix} \text{ with } \int_G \log |\hat{\alpha}/\hat{\beta}| \, d\nu = 0.$$

This implies that $\nu * \nu$ is degenerate diagonal and so, by Section 4.3, Theorem C holds at $\nu * \nu$. Using Remark 2.2.2, we find that the theorem also holds at ν.

In the remainder of this section we assume that $|\Delta| < 1$ and T is not identically zero on the support of ν. Thus, by the Jensen inequality, the number $\sigma^2 = \int_G T^2 \, d\nu$ is strictly positive. The main result in this section is the following variation of (4.3.4) for the present setting:

Proposition 4.4.1. *There exists $D > 0$ such that*

$$\left| \mu \left(\left\{ (g_j)_j \in G^{\mathbb{Z}} : \frac{1}{\sigma\sqrt{n}} \sum_{i=0}^{n-1} (-1)^{\kappa_i} T(g_i) \in [a, b] \right\} \right) - \mathcal{N}(a, b) \right| \leq D \frac{\log n}{\sqrt{n}}$$

for every $n \geq 2$ and any $a < b$.

Proof. The main difference with respect to (4.3.4) is that this time we are dealing with random variables

$$(-1)^{\kappa_i} T(g_i), \quad i = 0, 1, \ldots$$

that are neither independent nor identically distributed, and so the Berry–Esseen theorem (4.3.4) may not apply. Instead, we are going to use the following result for exponentially mixing sequences, due to Schneider [Sch81, Theorem 1]:

Theorem 4.4.2. *Let X_i, $i = 0, 1, \ldots$ be a sequence of random variables in some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ satisfying*

(a) $\mathbb{E}(X_i) = 0$ *for every* $i \geq 0$ *and* $\sup\{\mathbb{E}(|X_i|^3) : i \geq 0\} < \infty$;

(b) $\liminf_n \dfrac{\upsilon_n^2}{n} > 0$, *where* $\upsilon_n^2 = \mathbb{E}\left(\left(\sum_{i=0}^{n-1} X_i\right)^2\right)$;

(c) there exist $C, c > 0$ *such that* $|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| \leq Ce^{-ck}\mathbb{P}(A)$ *for any* $A \in \mathcal{A}_{0,l}$, $B \in \mathcal{A}_{l+k,\infty}$, $l \geq 0$ *and* $k \geq 1$;

where $\mathcal{A}_{m,n} \subset \mathcal{A}$ *denotes the* σ*-algebra generated by* $\{X_i : m \leq i \leq n\}$.

Then there exists $D > 0$ *such that*

$$\left| \mathbb{P}\left(\frac{1}{\upsilon_n} \sum_{i=0}^{n-1} X_i \in [a, b] \right) - \mathcal{N}(a, b) \right| \leq D \frac{\log n}{\sqrt{n}}$$

for every $n \geq 2$ *and any* $a < b$.

We are going to apply Theorem 4.4.2 to the sequence $X_i = (-1)^{\kappa_i} T(g_i)$, with $\mathbb{P} = \mu$ and $\mathbb{E}(X) = \int_G X \, d\mu$. Observe that

$$\begin{aligned} \mathbb{E}(X_i) &= \int_G (-1)^{\kappa(g_0) + \cdots + \kappa(g_{i-1})} T(g_i) \, d\nu(g_0) \cdots d\nu(g_{i-1}) \, d\nu(g_i) \\ &= \Delta^i \int_G T \, d\nu = 0 \text{ for every } i \geq 0. \end{aligned}$$

Moreover, $\sup_i \mathbb{E}(|X_i|^3) \leq (\sup |T|)^3 < \infty$ because $\sup |X_i| \leq \sup |T|$. This gives condition (a).

Next, observe that $\mathbb{E}(X_i^2) = \int_G T^2 \, d\nu$ for any i, and

$$\begin{aligned} \mathbb{E}(X_i X_j) &= \int_G (-1)^{\kappa(g_i)} T(g_i) (-1)^{\kappa(g_{i+1}) + \cdots + \kappa(g_{j-1})} T(g_j) \, d\nu(g_i) \cdots d\nu(g_{j-1}) \\ &= \Delta^{j-i-1} \int_G (-1)^{\kappa} T \, d\nu \int_G T \, d\nu = 0 \text{ for any } i < j. \end{aligned}$$

Then $\upsilon_n^2 = \sum_{i,j=0}^{n-1} \mathbb{E}(X_i X_j) = \sigma^2 n$. Since $\sigma^2 > 0$, this gives condition (b).

Next we check condition (c). Given any (i_r, ξ_r, U_r), $r = 0, \dots, s$ with

$$0 \le i_0 < \cdots < i_s, \quad \xi_r \in \{-1, +1\} \text{ and } U_r \subset \mathbb{R},$$

let $\Sigma(i_0, \xi_0, U_0, \dots, i_s, \xi_s, U_s)$ denote the subset of $(g_j)_j \in M$ such that

$$(-1)^{\kappa_{i_r}} = \xi_r \text{ and } T(g_{i_r}) \in U_r \text{ for } r = 0, \dots, s.$$

Each σ-algebra $\mathcal{A}_{m,n}$ is generated by the sets $\Sigma(i_0, \xi_0, U_0, \dots, i_s, \xi_s, U_s)$ with $i_0 \ge m$ and $i_s < n$. Denote

$$\Theta(i, \xi, U) = \mathbb{P}\big((-1)^{\kappa_i} = \xi, T(g_i) \in U\big) \text{ and}$$
$$\Theta(i, \xi, U \mid V) = \frac{\mathbb{P}\big(T(g_0) \in V, (-1)^{\kappa_i} = \xi, T(g_i) \in U\big)}{\mathbb{P}\big(T(g_0) \in V\big)}.$$

Lemma 4.4.3. *For any $i \ge 1$, ξ, U and V,*

$$\left|\Theta(i, \xi, U \mid V) - \frac{1}{2} T_*\nu(U)\right| \le \frac{1}{2} |\Delta|^{i-1} T_*\nu(U).$$

Proof. Clearly,

$$\begin{aligned} &\mathbb{P}\big((-1)^{\kappa_i} = 1\big) + \mathbb{P}\big((-1)^{\kappa_i} = -1\big) = 1 \text{ and} \\ &\mathbb{P}\big((-1)^{\kappa_i} = 1\big) - \mathbb{P}\big((-1)^{\kappa_i} = -1\big) = \int_G (-1)^{\kappa_i} \, d\mu = \Delta^i. \end{aligned} \tag{4.4.6}$$

This gives that

$$\Theta(i, \xi, U) = \mathbb{P}\big((-1)^{\kappa_i} = \xi\big) \mathbb{P}\big(T(g_i) \in U\big) = \frac{1 + \xi\Delta^i}{2} T_*\nu(U).$$

Then

$$\begin{aligned} \Theta(i, \xi, U \mid V) &= \sum_{\xi' \in \{-1, +1\}} \frac{\mathbb{P}\left(T(g) \in V, (-1)^{\kappa(g)} = \xi'\right)}{\mathbb{P}(T(g) \in V)} \Theta\left(i - 1, \xi/\xi', U\right) \\ &= \sum_{\xi' \in \{-1, +1\}} \frac{\mathbb{P}\left(T(g) \in V, (-1)^{\kappa(g)} = \xi'\right)}{\mathbb{P}(T(g) \in V)} \frac{1 + \xi/\xi' \Delta^{i-1}}{2} T_*\nu(U). \end{aligned}$$

This gives the conclusion of the lemma. □

It follows from the definitions that

$$\begin{aligned} &\mathbb{P}\left(\Sigma(i_0, \xi_0, U_0, \dots, i_s, \xi_s, U_s)\right) \\ &\qquad = \Theta(i_0, \xi_0, U_0) \prod_{r=1}^{s} \Theta(i_r - i_{r-1}, \xi_r/\xi_{r-1}, U_r \mid U_{r-1}). \end{aligned}$$

Thus, given any

$$A = \Sigma(i_0, \xi_0, U_0, \dots, i_s, \xi_s, U_s) \text{ and } B = \Sigma(j_0, \upsilon_0, V_0, \dots, j_t, \upsilon_t, V_t)$$

with $i_s \le l$ and $j_0 \ge l + k$, we have

$$\begin{aligned}\mathbb{P}(A)\mathbb{P}(B) =& \Theta(i_0, \xi_0, U_0) \prod_{r=1}^{s} \Theta(i_r - i_{r-1}, \xi_r/\xi_{r-1}, U_r \mid U_{r-1}) \\ & \Theta(j_0, \upsilon_0, V_0) \prod_{r=1}^{t} \Theta(j_r - j_{r-1}, \upsilon_r/\upsilon_{r-1}, V_r \mid V_{r-1}).\end{aligned}$$

whereas

$$\begin{aligned}\mathbb{P}(A \cap B) =& \Theta(i_0, \xi_0, U_0) \prod_{r=1}^{s} \Theta(i_r - i_{r-1}, \xi_r/\xi_{r-1}, U_r \mid U_{r-1}) \\ & \Theta(j_0 - i_s, \upsilon_0/\epsilon_s, V_0 \mid U_s) \prod_{r=1}^{t} \Theta(j_r - j_{r-1}, \upsilon_r/\upsilon_{r-1}, V_r \mid V_{r-1}).\end{aligned}$$

Then, using Lemma 4.4.3 and recalling that $j_0 - i_s \ge k$,

$$\begin{aligned}|\mathbb{P}(A \cap B)\mathbb{P}(A) \cap \mathbb{P}(B)| &\le \mathbb{P}(A) \left|\Theta(j_0 - i_s, \upsilon_0/\epsilon_s, V_0 \mid U_s) - \Theta(j_0, \upsilon_0, V_0)\right| \\ &\le \frac{1}{2} T_* \nu(V_0) |\Delta|^{j_0 - i_s - 1} + \frac{1}{2} T_* \nu(V_0) |\Delta|^{j_0 - 1} \le |\Delta|^{k-1}.\end{aligned}$$

This implies condition (c), since we are assuming that $|\Delta| < 1$.

This proves that Proposition 4.4.1 is a particular instance of Theorem 4.4.2. □

The following consequence contains property (4.2.6) in this case:

Proposition 4.4.4. *There exists $D_4 > 0$ such that for any $\xi \in \mathcal{M}(P)$*

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \le \log|z_v| \le R\}\big) \le \frac{D_4 R \log n}{\sqrt{n}} \quad \textit{for every } n \ge 2 \textit{ and } R \ge 1.$$

Proof. As in Proposition 4.3.1, it suffices to consider the case when $\xi = \delta_u$ for some $u \in P \setminus \{E, F\}$. Then, using (4.4.3),

$$\begin{aligned}&\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \le \log|z_v| \le R\}\big) \\ &\qquad = \mu\left(\left\{(g_j)_j \in G^{\mathbb{Z}} : \log|z_u| + \sum_{i=0}^{n-1} (-1)^{\kappa_i} T(g_i) \in [-R, R]\right\}\right).\end{aligned}$$

Thus the proof of Proposition 4.4.4 is analogous to that of Proposition 4.3.1, with $T(g_i)$ replaced with $(-1)^{\kappa_i} T(g_i)$ and Proposition 4.4.1 in the place of the Berry–Esseen theorem. □

Next we establish (4.2.7) in this case. For each $\xi \in \mathcal{M}_c(G)$ and $R > 0$, define

$$\delta_n(\xi, R) = \left|\mathcal{P}_\nu^n \xi\big(\{v \in P : \log|z_v| > R\}\big) - \mathcal{P}_\nu^n \xi\big(\{v \in P : \log|z_v| < -R\}\big)\right|.$$

Proposition 4.4.5. *There exists $D_5 > 0$ such that for any $\xi \in \mathcal{M}(P)$*

$$\delta_n(\xi, R) \le \frac{D_5 R}{n^{1/3}}$$

for every $n \ge 1$ and $R \ge 1$.

Proof. Let $S_n = S_n(v, g_0, \ldots, g_{n-1})$ be the expression on the right-hand side of (4.4.3), that is,

$$S_n = (-1)^{\kappa_n}\left(\log|z_v| + \sum_{i=0}^{n-1}(-1)^{\kappa_i} T(g_i)\right).$$

Then $\delta_n(\xi, R) = |\mathbb{P}_n(S_n > R) - \mathbb{P}_n(S_n < -R)|$, where $\mathbb{P}_n$ stands for $\xi \times \nu^n$. Observe that

$$\mathbb{P}_{n+1}(S_{n+1} > R) = \int_{\kappa=0} \mathbb{P}_n\left(S_n > R - T(g_n)\right) d\nu(g_n) + \int_{\kappa=1} \mathbb{P}_n\left(S_n < -R - T(g_n)\right) d\nu(g_n)$$

$$\mathbb{P}_{n+1}(S_{n+1} < -R) = \int_{\kappa=0} \mathbb{P}_n\left(S_n < -R - T(g_n)\right) d\nu(g_n) + \int_{\kappa=1} \mathbb{P}_n\left(S_n > R - T(g_n)\right) d\nu(g_n),$$

since $S_{n+1} = (-1)^{\kappa(g_n)}\big(S_n + T(g_n)\big)$. Let us write

$$\mathbb{P}_n\left(S_n > R - T(g_n)\right) = \mathbb{P}_n\left(S_n > R\right) + \mathbb{P}'_n$$
$$\mathbb{P}_n\left(S_n < -R - T(g_n)\right) = \mathbb{P}_n\left(S_n < -R\right) + \mathbb{P}''_n$$

where

$$\mathbb{P}'_n = \begin{cases} \mathbb{P}_n\left(R - T(g_n) < S_n \le R\right) & \text{if } T(g_n) \ge 0 \\ -\mathbb{P}_n\left(R - T(g_n) > S_n \ge R\right) & \text{if } T(g_n) \le 0 \end{cases}$$

and analogously for $\mathbb{P}''$. Let $R_0 = R + \max|T|$. Note that $\mathbb{P}'_n$ and $\mathbb{P}''_n$ have the same sign and they are bounded, in absolute value, by $\mathbb{P}_n\left(|S_n| \le R_0\right)$. According to Proposition 4.4.4, the latter is bounded by $D_4 R_0 \log n/\sqrt{n}$ for every $n \ge 2$. Hence, we may find $D'_4 > 1$ such that $|\mathbb{P}'|$ and $|\mathbb{P}''_n|$ are both bounded by $D'_4 R n^{-1/3}$ for every $n \ge 1$. It follows that

$$\begin{aligned}\delta_{n+1}(\xi, R) &= |\mathbb{P}_n(S_{n+1} > R) - \mathbb{P}_n(S_{n+1} < -R)| \\ &\le \Big|\big[\nu(\{\kappa = 0\} - \nu(\{\kappa = 1\})\big]\big[\mathbb{P}_n\left(S_n > R\right) - \mathbb{P}_n\left(S_n < -R\right)\big]\Big| + D'_4 R n^{-1/3} \\ &= |\Delta|\, \delta_n(\xi, R) + D'_4 R n^{-1/3}\end{aligned}$$

for every $n \geq 1$. This yields

$$\delta_n(\xi, R) \leq |\Delta|^{n-1}\delta_1(\xi, R) + \sum_{j=1}^{n-1} |\Delta|^{n-1-j} D_4' R j^{-1/3}.$$

Recall that $|\Delta| < 1$ and $\delta_1(\xi, R) \leq 1$. Fix $D_5' > 1$ large enough that

$$|\Delta|^{(n-1)/2} \leq D_5' n^{-1/3} \text{ and } D_4' j^{-1/3} \leq D_5' |\Delta|^{(j+1-n)/2} n^{-1/3}$$

for $n \geq j \geq 1$. It follows that

$$\delta_n(\xi, R) \leq \sum_{j=0}^{n} |\Delta|^{(n-1-j)/2} D_5' R n^{-1/3} \leq D_5 R n^{-1/3},$$

with $D_5 = D_5' \sum_{i=-1}^{\infty} |\Delta|^{i/2}$. □

4.5 Degenerate triangular case

Take ν to be degenerate triangular. The setting and terminology are as in Section 3.3, except that now $\lambda_-(\nu) = \lambda_+(\nu)$, that is,

$$\int_G \log |\tau(g)| \, d\nu(g) = 0. \tag{4.5.1}$$

Every $g \in \operatorname{supp} \nu$ may be written in the form (3.1.1), and then its action on P is given by

$$z_v \mapsto \tau(g) \left[z_v + \theta(g) \right] \text{ with } \tau(g) = \frac{\alpha(g)}{\beta(g)} \text{ and } \theta(g) = \frac{\gamma(g)}{\alpha(g)}.$$

Thus the action of $g_{n-1} \cdots g_0$ on P is given by

$$z_v \mapsto \tau(g_{n-1}) \cdots \tau(g_0) z_v + \sum_{i=0}^{n-1} \tau(g_{n-1}) \cdots \tau(g_{i+1}) \tau(g_i) \theta(g_i). \tag{4.5.2}$$

We are going to distinguish two cases, depending on whether $|\tau|$ is constant on the support of ν or not. The first one turns out to be similar to the simply irreducible case, and is dealt with in Section 4.5.1. The second one, in Section 4.5.2, requires new methods, that we borrow from the theory of perpetuities. We refer the interested reader to [BDM16] and the references therein for more information on that theory.

4.5.1 Sub-case $|\tau| \equiv 1$

Write $\tau(g) = (-1)^{\kappa(g)}$, with $\kappa(g) \in \{0, 1\}$. Then (4.5.2) becomes

$$z_v \mapsto (-1)^{\kappa_n} \left(z_v + \sum_{i=0}^{n-1} (-1)^{\kappa_i} \theta(g_i) \right), \tag{4.5.3}$$

where $\kappa_i = \kappa(g_0) + \cdots + \kappa(g_{i-1})$ for every $i \geq 0$. This has the same form as (4.4.3), and that allows us to adapt the analysis in Section 4.4 to the present situation, as we are going to explain.

Let us deal first with a few special sub-cases. If θ is identically zero then every $g \in \operatorname{supp} \nu$ is conformal and we have already seen (Section 4.1) that Theorem C holds in that case. So, it is no restriction to assume that θ is not identically zero, and so $\varsigma^2 = \int_G \theta^2 \, d\nu$ is strictly positive

Next, let $\Delta = \nu(\{\kappa = 0\}) - \nu(\{\kappa = 1\})$, as defined in (4.4.5). Note that Δ coincides with $\int_G (1/\tau) \, d\nu$. If $\Delta = 1$ then $\kappa(g) = 0$ for every $g \in \operatorname{supp} \nu$, and (4.5.3) further reduces to

$$z_v \mapsto z_v + \sum_{i=0}^{n-1} \theta(g_i).$$

Since the random variables $\theta(g_i)$ are independent and identically distributed, we are back in the situation of Section 4.3, except for the fact that the expected value of θ need not be zero. The Berry-Esseen theorem (4.3.4) gives that

$$\left| \mu\left(\left\{ (g_j)_j \in G^{\mathbb{Z}} : \frac{1}{\varsigma\sqrt{n}} \sum_{i=0}^{n-1} \left[\theta(g_i) - \int_G \theta \, d\nu \right] \in [a, b] \right\} \right) - \mathcal{N}(a, b) \right| \leq \frac{\varrho}{\varsigma^3 \sqrt{n}}$$

for any $n \geq 1$ and $a < b$, where $\varrho = \int \theta^3 \, d\nu$. This implies that, for any $u \neq E$,

$$\begin{aligned} \mathcal{P}_\nu^n \delta_u\big(\{v \in P : -R \leq z_v \leq R\}\big) &= \mu\left(\left\{ (g_j)_j : z_u + \sum_{i=0}^{n-1} \theta(g_i) \in [-R, R] \right\} \right) \\ &\leq \mathcal{N}\left(\frac{-R - z_u - n\int_G \theta \, d\nu}{\varsigma\sqrt{n}}, \frac{R - z_u - n\int_G \theta \, d\nu}{\varsigma\sqrt{n}} \right) + \frac{\varrho}{\varsigma^3\sqrt{n}} \end{aligned}$$

Using that $\mathcal{N}(a, b) < (b - a)$ and then integrating with respect to u we find that

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \leq z_v \leq R\}\big) \leq \frac{2R}{\varsigma\sqrt{n}} + \frac{\varrho}{\varsigma^3\sqrt{n}}$$

for any $\xi \in \mathcal{M}(P)$, which implies (4.2.8).

If $\Delta = -1$ then $\alpha(g) = -\beta(g)$ for every $g \in \operatorname{supp} \nu$. It follows that every h in the support of the 2-convolution $\nu * \nu$ has the form

$$h = \begin{pmatrix} \hat{\alpha}(h) & \hat{\gamma}(h) \\ 0 & \hat{\beta}(h) \end{pmatrix}$$

with $\hat{\alpha}(h) = \hat{\beta}(h)$. This means that $\nu * \nu$ is in the conditions of the previous paragraph, and so Theorem C holds at $\nu * \nu$. By Remark 2.2.2, it follows that the theorem also holds at ν.

Now we only need to consider the case when $|\Delta| < 1$. Up to conjugating the measure ν (Remark 2.2.3) we may assume that

$$\int_G \theta \, d\nu = 0. \tag{4.5.4}$$

To see this, consider $t \in \mathbb{R}$ and $h : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x + ty, y)$. For g as in (3.1.1), the action of hgh^{-1} on P is given by

$$z_v \mapsto \hat{\tau}(hgh^{-1}) \left[z_v + \hat{\theta}(hgh^{-1}) \right] \text{ with } \begin{cases} \hat{\tau}(hgh^{-1}) = \tau(g) \text{ and} \\ \hat{\theta}(hgh^{-1}) = \theta(g) + (1/\tau(g) - 1)\, t. \end{cases}$$

Since $\int_G 1/\tau d\nu = \Delta \neq 1$, we may choose $t \in \mathbb{R}$ so that the expected value of $\hat{\theta}$ vanishes, as claimed.

Then the following statement is analogous to Proposition 4.4.1, with θ and ς in the place of T and σ:

Proposition 4.5.1. *There exists $D > 0$ such that*

$$\left| \mu \left(\left\{ (g_j)_j \in G^{\mathbb{Z}} : \frac{1}{\varsigma \sqrt{n}} \sum_{i=0}^{n-1} (-1)^{\kappa_i} \theta(g_i) \in [a, b] \right\} \right) - \mathcal{N}(a, b) \right| \leq D \frac{\log n}{\sqrt{n}}$$

for every $n \geq 1$ and any $a < b$.

Now the property (4.2.8) is contained in the following statement:

Proposition 4.5.2. *There exists $D_6 > 0$ such that for any $\xi \in \mathcal{M}(P)$*

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \leq z_v \leq R\}\big) \leq \frac{D_6 R \log n}{\sqrt{n}} \text{ for every } n \geq 1 \text{ and } R \geq 1.$$

Proof. Analogous to that of Proposition 4.4.4 (and Proposition 4.3.1), with Proposition 4.5.1 in the place of Proposition 4.4.1 (and the Barry-Esseen theorem). Note that $\log |z_v|$ is also replaced with z_v. □

4.5.2 Sub-case $|\tau|$ non-constant

This is inspired by the proof of Lemma 5.18 of Goldie, Maller [GM00], which is itself based on ideas of Grincevičjus [Gri74, Gri75]. By the Jensen inequality, the assumption that $|\tau|$ is not constant implies that $\sigma^2 = \int \log^2 |\tau| \, d\nu$ is strictly positive.

We start with some general facts from probability theory. Given any (real) random variable Z, let Z^s denote a symmetrized version of Z. This means that $Z^s = Z - Z'$ where Z' is some random variable independent of Z and with the same distribution function. For each $\lambda > 0$, define

$$Q(Z;\lambda) = \sup_{x\in\mathbb{R}} \mathbb{P}(x \le Z \le x+\lambda).$$

Lemma 1.11 in Petrov [Pet95] asserts that

$$Q(Z^s;\lambda) \le Q(Z;\lambda) \text{ for every } \lambda \in \mathbb{R}. \tag{4.5.5}$$

Lemma 4.5.3. *Let Z, Z' and T be independent random variables such that Z and Z' are identically distributed, and let $Z^s = Z - Z'$. Then*

$$\mathbb{P}(|Z+T| \le R/2)^2 \le \mathbb{P}(|Z^s| \le R) \text{ for every } R > 0.$$

Proof. The hypotheses imply that $Z+T$ and $Z'+T$ are independent and identically distributed. So,

$$\begin{aligned}\mathbb{P}(|Z+T| \le R/2)^2 &= \mathbb{P}(|Z+T| \le R/2)\mathbb{P}(|Z'+T| \le R/2)\\ &= \mathbb{P}(|Z+T| \le R/2, |Z'+T| \le R/2) \le \mathbb{P}(|Z-Z'| \le R)\end{aligned}$$

as claimed. □

Lemma 4.5.4. *There exists $A > 0$ such that if $X_1, \dots, X_n$ are independent identically distributed random variables and $S_n = X_1 + \cdots + X_n$ then*

$$\mathbb{P}(|S_n^s| \le R) \le \frac{A}{\sqrt{\sum_{k=1}^n \mathbb{P}\left(|X_k^s| \ge R\right)}} \text{ for any } R > 0.$$

Proof. This is an easy consequence of Theorem 2.15 in Petrov [Pet95]. Indeed, the theorem asserts that

$$Q(S_n;\lambda) \le \frac{A\lambda}{\sqrt{\sum_{k=1}^n \lambda_k^2 \mathbb{P}\left(|X_k^s| \ge \lambda_k/2\right)}}$$

for any $\lambda > 0$ and $\lambda_1, \dots, \lambda_k \in (0, \lambda]$. Taking $\lambda = \lambda_1 = \cdots = \lambda_k = 2R$,

$$\mathbb{P}(|S_n^s| \le R) \le Q(S_n^s; 2R) \le Q(S_n; 2R) \le \frac{A}{\sqrt{\sum_{k=1}^n \mathbb{P}\left(|X_k^s| \ge R\right)}}$$

where the middle inequality uses (4.5.5). □

Going back to the setting we are interested in, let us consider

$$\tau_n = \tau(g_n), q_n = \tau(g_n)\theta(g_n) \text{ and } S_n = \sum_{i=0}^{n-1} \tau_{n-1}\cdots\tau_{i+i}q_i. \tag{4.5.6}$$

Keep in mind that the random variables (τ_n, q_n), $n \geq 0$ are independent and identically distributed.

Let $\{\mathbb{P}(\cdot \mid \tau_0 = m) : m \in \mathbb{R}\}$ be a family of conditional probabilities of $\mathbb{P}$ given τ_0, that is, a Rokhlin disintegration (see [VO15, Chapter 5]) of $\mathbb{P}$ with respect to the partition $\{\tau_0^{-1}(m) : m \in \mathbb{R}\}$. Let $F(q, m)$ be the *conditional distribution* of q_0 given τ_0, that is,

$$F(q, m) = \mathbb{P}(q_0 \leq q \mid \tau_0 = m) \text{ for } q \in \mathbb{R}.$$

Clearly, $q \mapsto F(q, m)$ is non-decreasing and right-continuous for each $m \in \mathbb{R}$. Let $q \mapsto F^{-1}(q, m)$ be the right-continuous inverse.

Let $U_0, \ldots, U_j, \ldots$ be independent random variables uniformly distributed on $[0, 1]$ and such that U_j, $j \geq 0$ and (τ_i, q_i), $i \geq 0$ are independent. Uniform distribution means that $\mathbb{P}(U_j \leq y) = y$ for every $y \in (0, 1)$ and $j \geq 0$. Define $q_n' = F^{-1}(U_n, \tau_n)$ for each $n \geq 0$. Note that q_n and q_n' have the same conditional distribution given τ_n:

$$\begin{aligned}
\mathbb{P}\left(q_n' \leq q \mid \tau_n = m\right) &= \mathbb{P}\left(U_n \leq F(q, m) \mid \tau_n = m\right) \\
&= \mathbb{P}\left(U_n \leq F(q, m)\right) = F(q, m) \\
&= \mathbb{P}\left(q_n \leq q \mid \tau_n = m\right).
\end{aligned}$$

As an immediate consequence, (q_n, τ_n) and (q_n', τ_n) have the same distribution. In particular, the (q_n', τ_n), $n \geq 0$ are identically distributed.

Furthermore, q_n and q_n' are conditionally independent given τ_n: since U_n is independent of (q_n, τ_n),

$$\begin{aligned}
\mathbb{P}\left(q_n' \leq q', q_n \leq q \mid \tau_n = m\right) &= \mathbb{P}\left(U_n \leq F(q', m), q_n \leq q \mid \tau_n = m\right) \\
&= \mathbb{P}\left(U_n \leq F(q', m)\right) \mathbb{P}\left(q_n \leq q \mid \tau_n = m\right) \\
&= F(q', m) F(q, m) \\
&= \mathbb{P}\left(q_n' \leq q' \mid \tau_n = m\right) \mathbb{P}\left(q_n \leq q \mid \tau_n = m\right).
\end{aligned}$$

More generally, the variables $q_0, \ldots, q_{n-1}, q_0', \ldots, q_{n-1}'$ are conditionally independent given $\tau_0, \ldots, \tau_{n-1}$ for any $n \geq 1$. As an immediate consequence, (q_i, τ_i), $i \geq 0$ and (q_j', τ_j), $j \geq 0$ are independent.

It follows from these observations that $q_n^s = q_n - q_n'$ is a conditionally symmetrized version of q_n given τ_n, and then

$$S_n^s = \sum_{i=0}^{n-1} \tau_{n-1} \cdots \tau_{i+i} q_i^s$$

is a conditionally symmetrized version of S_n given $\tau_{n-1}, \ldots, \tau_0$. The random variables $\tau_{n-1} \cdots \tau_{i+i} q_i^s$, $0 \leq i \leq n-1$ are conditionally independent given $\tau_0, \ldots, \tau_{n-1}$:

$$\begin{aligned}
&\mathbb{P}\left(\tau_{n-1} \cdots \tau_{i+i} q_i^s \geq R_i, 0 \leq i \leq n-1 \mid \tau_j = m_j, 0 \leq j \leq n-1\right) \\
&\qquad = \prod_{i=0}^{n-1} \mathbb{P}\left(|\tau_{n-1} \cdots \tau_{i+i} q_i^s| \geq R_i \mid \tau_j = m_j, 0 \leq j \leq n-1\right).
\end{aligned}$$

for any $R_i, m_j \in \mathbb{R}$, because the q_i', $0 \le i \le n-1$ are conditionally independent given $\tau_0, \ldots, \tau_{n-1}$. So, applying Lemma 4.5.4 to the conditional distributions given $\tau_0, \ldots, \tau_{n-1}$,

$$\begin{aligned}&\mathbb{P}\left(|S_n^s| \le R \mid \tau_j = m_j, 0 \le j \le n-1\right)\\&\quad\le \frac{A}{\sqrt{\sum_{i=0}^{n-1} \mathbb{P}\left(|\tau_{n-1}\cdots\tau_{i+1} q_i^s| \ge R \mid \tau_j = m_j, 0 \le j \le n-1\right)}}.\end{aligned} \tag{4.5.7}$$

At this point we need to distinguish two sub-cases. Suppose first that q_i^s is not concentrated at zero, meaning that $\mathbb{P}(q_i^s = 0) < 1$ (this condition is independent of $i \ge 0$, of course). Then property (4.2.8) is contained in the following statement:

Proposition 4.5.5. *There exists $D_7 > 0$ such that, given any $\xi \in \mathcal{M}(P)$,*

$$\mathcal{P}_\nu^n \xi\big(\{v \in P : -R \le z_v \le R\}\big) \le D_7 \left(\frac{\log R}{n}\right)^{1/8} \quad \textit{for every } n \ge 1 \textit{ and } R \ge e.$$

Proof. Just as in the proof of Proposition 4.3.1, it suffices to consider the case when $\xi = \delta_u$ for some $u \in P \setminus \{E\}$.

The assumption that q_i^s is not concentrated at zero ensures that there exist $b > 0$ and $\delta > 0$ such that $\mathbb{P}(|q_i^s| \ge \delta) = b$. So, by independence,

$$\begin{aligned}&\mathbb{P}\left(|\tau_{n-1}\cdots\tau_{i+1} q_i^s| \ge R \mid \tau_j = m_j, 0 \le j \le n-1\right)\\&\qquad\qquad = \mathbb{P}(|q_i^s| \ge R/|m_{n-1}\cdots m_{i+1}|) \ge \mathbb{P}(|q_i^s| \ge \delta) = b\end{aligned}$$

whenever $|m_{n-1}\cdots m_{i+1}| \ge R/\delta$. Define,

$$\kappa_n(x, m_0, \ldots, m_{n-1}) = \#\left\{0 \le i \le n-1 : |m_{n-1}\cdots m_{i+1}| \ge x\right\}.$$

So, (4.5.7) entails

$$\mathbb{P}\left(|S_n^s| \le R \mid \tau_j = m_j, 0 \le j \le n-1\right) \le \frac{A}{\sqrt{b\kappa_n(R/\delta, m_0, \ldots, m_{n-1})}}. \tag{4.5.8}$$

Lemma 4.5.6. *There exists $D_8 > 0$ such that*

$$\mathbb{P}(|S_n^s| \le R) \le D_8 \left(\frac{\log R}{n}\right)^{1/4} \quad \textit{for every } n \ge 1 \textit{ and } R \ge e.$$

Proof. By assumption, the random variables $Y_j = \log|\tau_j|$, $j \ge 0$ are independent and identically distributed, with $\mathbb{E}(Y_j) = 0$ and $\sigma^2 = \mathbb{E}(Y_j^2) \in (0, \infty)$. Let $\kappa_n = \kappa(1, \tau_0, \ldots, \tau_{n-1})$. That is, κ_n is the number of integers $0 \le i \le n-1$ such that $Y_{n-1} + \cdots + Y_{i+1} > 0$. By the Lévy-Erdös-Kac arcsine law [EK47],

$$\lim_n \mathbb{P}\left(\frac{\kappa_n}{n} < \alpha\right) = \frac{2}{\pi} \arcsin \alpha^{1/2} \le \alpha^{1/2} \text{ for every } 0 \le \alpha \le 1.$$

The proof also gives that the convergence is uniform on $\alpha \in [0,1]$. So there exists $n_0 \geq 1$ such that

$$\mathbb{P}\left(\kappa_n < \alpha n\right) \leq \alpha^{1/2} \text{ for every } 0 \leq \alpha \leq 1 \text{ and } n \geq n_0. \tag{4.5.9}$$

Now define $\kappa_n(x) = \kappa(x, \tau_0, \ldots, \tau_{n-1}) =$ number of integers $0 \leq i \leq n-1$ such that $Y_{n-1} + \cdots + Y_{i+1} > \log x$. It is clear that, for any $x > 0$ and $l \geq 1$, $Y_{n-1} + \cdots + Y_{i+1} < 0$ if $Y_{n+l-1} + \cdots + Y_{i+1} < \log x$ and $Y_{n+l-1} + \cdots + Y_n \leq -\log x$. So, given any $w \in \mathbb{R}$,

$$\mathbb{P}\left(n - \kappa_n > w\right) \geq \mathbb{P}\left(n - \kappa_{n+l}(x) > w\right) \mathbb{P}(Y_{n+l-1} + \cdots + Y_n \leq -\log x). \tag{4.5.10}$$

Fix $x = R/\delta$. Since $\mathbb{E}(Y_j) = 0$, $\mathbb{E}(Y_j^2) > 0$ and $\rho = \mathbb{E}(Y_j^3)$ is finite, it follows from the Berry–Esseen theorem (4.3.4) that

$$\begin{aligned} \mathbb{P}\left(Y_{n+l-1} + \cdots + Y_n \leq -\log \frac{R}{\delta}\right) &\geq \mathcal{N}\left(-\infty, -\frac{\log(R/\delta)}{\sigma\sqrt{l}}\right) - \frac{\rho}{\sigma^3\sqrt{l}} \\ &\geq \frac{1}{2} - \frac{\log(R/\delta)}{\sigma\sqrt{l}} - \frac{\rho}{\sigma^3\sqrt{l}} \geq \frac{1}{3}, \end{aligned} \tag{4.5.11}$$

as long as we choose $l \geq 36(\sigma^2 \log(R/\delta) + \rho)^2 \sigma^{-6}$. Fix $D > 0$, depending on σ, ρ and δ, such that this last constraint is compatible with $l \leq D \log R$. Inserting (4.5.11) in (4.5.10),

$$\mathbb{P}\left(n - \kappa_n > w\right) \geq \frac{1}{3}\mathbb{P}\left(n - \kappa_{n+l}(R/\delta) > w\right) \text{ for every } w \in \mathbb{R} \text{ and } n \geq 1.$$

Taking $w = n - \alpha n$, we ge that $\mathbb{P}\left(\kappa_n < \alpha n\right) \geq \mathbb{P}\left(\kappa_{n+l}(R/\delta) < \alpha n\right)/3$. Combining this with (4.5.9), we find that

$$\mathbb{P}\left(\kappa_{n+l}(R/\delta) < \alpha n\right) \leq 3\alpha^{1/2} \text{ for every } n \geq n_0 \text{ and } 0 \leq \alpha \leq 1.$$

Then, integrating (4.5.8) with respect to $m_0, \ldots, m_{n-1}$,

$$\begin{aligned} \mathbb{P}\left(|S^s_{n+l}| \leq R\right) &\leq \mathbb{P}\left(\kappa_{n+l}(R/\delta) < \alpha n\right) 1 + \mathbb{P}\left(\kappa_{n+l}(R/\delta) \geq \alpha n\right) \frac{A}{(b\alpha n)^{1/2}} \\ &\leq 3\alpha^{1/2} + \frac{A}{(b\alpha n)^{1/2}} \text{ for every } 0 \leq \alpha \leq 1 \text{ and } n \geq n_0. \end{aligned}$$

Then, taking $\alpha = n^{-1/2}$,

$$\mathbb{P}\left(|S^s_{n+l}| \leq R\right) \leq \left(3 + \frac{A}{b^{1/2}}\right) \frac{1}{n^{1/4}} \leq \left(3 + \frac{A}{b^{1/2}}\right) \left(\frac{2n_0 l}{n+l}\right)^{1/4}.$$

So, taking $D_8 = \left(3 + A/b^{1/2}\right)(2n_0 D)^{1/4}$,

$$\mathbb{P}\left(|S^s_n| \leq R\right) \leq D_8 \left(\frac{\log R}{n}\right)^{1/4} \quad \text{for } n \geq n_0 + l.$$

The inequality remains valid for $n < n_0 + l$, because $\mathbb{P}\left(|S^s_n| \leq R\right) \leq 1$. $\square$

It is clear that the random variables S_n, S'_n and $T = \tau_{n-1} \cdots \tau_0 z_u$ are conditionally independent given $\tau_0, \dots, \tau_{n-1}$. Denote

$$Z_n(u) = \tau(g_{n-1}) \cdots \tau(g_0) z_u + \sum_{i=0}^{n-1} \tau(g_{n-1}) \cdots \tau(g_{i+1}) \tau(g_i) \theta(g_i). \quad (4.5.12)$$

for $u \in P \setminus \{E\}$ and $n \geq 0$. Then Lemma 4.5.3 gives that

$$\begin{aligned} \mathbb{P}\left(|Z_n(u)| \leq R \mid \tau_j = m_j, 0 \leq j \leq n-1\right)^2 & \\ &= \mathbb{P}\left(|S_n + T| \leq R \mid \tau_j = m_j, 0 \leq j \leq n-1\right)^2 \\ &\leq \mathbb{P}\left(|S_n^s| \leq 2R \mid \tau_j = m_j, 0 \leq j \leq n-1\right). \end{aligned}$$

Integrating with respect to $m_0, \dots, m_{n-1}$, and using the Jensen inequality and Lemma 4.5.6, we find that

$$\mathbb{P}\left(|Z_n(u)| \leq R\right)^2 \leq \mathbb{P}\left(|S_n^s| \leq 2R\right) \leq D_8 \left(\frac{\log(2R)}{n}\right)^{1/4}.$$

In view of (4.5.2), this proves the conclusion of the proposition for $\xi = \delta_u$. □

Finally, we deal with the special sub-case when $\mathbb{P}(q_i^s = 0) = 1$. Since q_i and q'_i are conditionally independent and identically distributed given τ_i, that can only happen if these two variables are concentrated on a single point for each value of τ_i, that is, if for each $m \in \mathbb{R}$ there is $q \in \mathbb{R}$ such that

$$\mathbb{P}(q_i = q \mid \tau_i = m) = \mathbb{P}(q'_i = q \mid \tau_i = m) = 1.$$

Now, this means that q_i is a function of τ_i. Recalling that $\tau_i = \tau(g_i)$ and $q_i = \tau(g_i)\theta(g_i)$, we see that this can only happen if $\theta = \varphi(\tau)$ for some real function φ.

Our assumption that there is a unique invariant line E means that the matrices $g \in \operatorname{supp} \nu$ have no common fixed point other than E, that is, there is no $z \in \mathbb{R}$ such that

$$\tau(g)\left[z + \theta(g)\right] = z \text{ for every } g \in \operatorname{supp} \nu.$$

Equivalently, $\tau\theta/(1-\tau)$ cannot be constant and finite on the support of ν.

Consider the 2-convolution $\tilde{\nu} = \nu * \nu$ and the corresponding functions

$$\tilde{\tau}(g_1 g_0) = \tau(g_1)\tau(g_0) \text{ and } \tilde{\theta}(g_1 g_0) = \theta(g_0) + \theta(g_1)/\tau(g_0).$$

If $\tilde{\theta}$ is a function of $\tilde{\tau}$ then $\tilde{\theta}(g_1 g_0) = \tilde{\theta}(g_0 g_1)$, as it is clear that $\tilde{\tau}(g_1 g_0) = \tilde{\tau}(g_0 g_1)$. In other words,

$$\theta(g_0) + \theta(g_1)/\tau(g_0) = \theta(g_1) + \theta(g_0)/\tau(g_1) \Leftrightarrow \frac{\tau(g_0)\theta(g_0)}{\tau(g_0) - 1} = \frac{\tau(g_1)\theta(g_1)}{\tau(g_1) - 1}$$

for every $g_0, g_1 \in \operatorname{supp} \nu$. This contradicts the observation in the previous paragraph. That contradiction proves that $\tilde{\theta}$ cannot be a function of $\tilde{\tau}$.

Then we can apply the previous argument to the convolution $\tilde{\nu} = \nu * \nu$, to conclude that it is a point of log-Hölder continuity for the Lyapunov exponents. By Remark 2.2.2 it follows that the same is true for ν.

The proof of Theorem C is now complete.

Chapter 5

A counter-example

Let $\nu = (\delta_g + \delta_h)/2$ be the average of the Dirac masses at

$$g = \begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}.$$

We are going to prove that

Proposition 5.0.1. *For any $\gamma > 1$, $\beta > 2$ and $C > 0$, there exist $\nu' = (\delta_{g'} + \delta_{h'})/2$ and $\nu'' = (\delta_{g''} + \delta_{h''})/2$ arbitrarily close to ν such that*

$$\left|\lambda_+(\nu') - \lambda_+(\nu'')\right| > C \exp\left(-\beta\left(\log\log\frac{1}{\delta_1(\nu', \nu'')}\right)^\gamma\right). \tag{5.0.1}$$

Given $\gamma, \beta > 0$, we say that a function $f : M \to \mathbb{R}$ on a metric space (M, d) is (γ, β)*-log-Hölder continuous* if there exists $C > 0$ such that

$$|f(x) - f(y)| \leq C \exp\left(-\beta\left(\log\log\frac{1}{d(x, y)}\right)^\gamma\right) \text{ for any } x, y \in M. \tag{5.0.2}$$

In particular, f is log-Hölder continuous if and only if it is $(1, \beta)$-log-Hölder continuous for some $\beta > 0$.

Proposition 5.0.1 implies that λ_+ fails to be (γ, β)-log-Hölder continuous on every neighborhood of ν, for any $\gamma > 1$ and $\beta > 2$. In particular, λ_+ is not locally Hölder continuous at ν. This shows that the assumption $\lambda_-(\nu) < \lambda_+(\nu)$ in Theorem B can not be removed. The proof of the proposition is a rather straightforward adaptation to our setup of a construction of Duarte, Klein, Santos [DKS], itself based on an example of Halperin, as presented by Simon, Taylor [ST85].

5.1 Schrödinger cocycles

Let $(\mathcal{X}, \mathcal{B}, \mu)$ be a probability space, $f : \mathcal{X} \to \mathcal{X}$ be an ergodic measure preserving transformation and $V : \mathcal{X} \to \mathbb{R}$ be a bounded measurable function (called *potential*). The associated *Schrödinger cocycle* is the parameterized family of linear cocycles

$$F_E : \mathcal{X} \times \mathbb{R}^2 \to \mathcal{X} \times \mathbb{R}^2, \quad F_E(x, v) = (f(x), A_E(x)v)$$

with

$$A_E(x) = \begin{pmatrix} V(x) - E & -1 \\ 1 & 0 \end{pmatrix} \text{ for } E \in \mathbb{R}.$$

Let $\lambda_\pm(E)$ denote the Lyapunov exponents of F_E, relative to the probability measure μ. Note that $\lambda_+(E) + \lambda_-(E) = 0$, since A_E takes values in the SL(2).

For each $x \in \mathcal{X}$, consider the *Schrödinger operator* $H : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ defined on the space $\ell^2(\mathbb{Z}) = \{(u(n))_{n\in\mathbb{Z}} : \sum_n \|u(n)\|^2 < \infty\}$ by

$$(H_x u)(n) = -u(n+1) - u(n-1) + V(f^n(x))u(n) \text{ for each } n \in \mathbb{Z}. \quad (5.1.1)$$

Note that H_x is linear and self-adjoint. By ergodicity, its spectral properties are independent of x on a full μ-measure set. For each $n \geq 1$, the *n-truncation* of H_x is the finite-rank linear operator

$$H_x^{(n)} = P_n \circ H_x \circ P_n^* : \ell^2(\{0, \dots, n-1\}) \to \ell^2(\{0, \dots, n-1\})$$

where $P_n : \ell^2(\mathbb{Z}) \to \ell^2(\{0, \dots, n-1\})$ is the canonical projection and the adjoint $P_n^* : \ell^2(\{0, \dots, n-1\}) \to \ell^2(\mathbb{Z})$ is the natural embedding. It is clear that $H_x^{(n)}$ is also self-adjoint. Let $\gamma_x^n(0) < \cdots < \gamma_x^n(n-1)$ be its eigenvalues.

By the Birkhoff ergodic theorem, for every $E \in \mathbb{R}$ there exists a number $N(E)$, called *integrated density of states*, such that

$$N(E) = \lim_{n\to\infty} \frac{1}{n} \#\{0 \leq j \leq n : \ \gamma_x^n(j) < E\} \quad (5.1.2)$$

for μ-almost every $x \in \mathcal{X}$ and every $E \in \mathbb{R}$. See Damanik [Dam17]. The integrated density of states and the Lyapunov exponents are related by the *Thouless formula*,

$$\lambda_+(E) = \int_{\mathbb{R}} \log |E - E'| \, dN(E'), \quad (5.1.3)$$

which means that $\lambda_+(E)$ is the Hilbert transform of $N(E)$.

Goldstein, Schlag [GS01, Lemma 10.3] proved that, given any $\gamma > 1$ and $\beta \geq 1$, the Lyapunov exponent $\lambda_+(E)$ is (γ, β)-log-Hölder continuous if and only if the integrated density of states $N(E)$ is (γ, β)-log-Hölder continuous. Craig, Simon [CS83] proved that the integrated density of states is always log-Hölder continuous. This has no immediate implications on the regularity of the Lyapunov exponents because the result of Goldstein, Schlag just mentioned does not extend to $\gamma = 1$.

5.2 Embedding in a Schrödinger cocycle

Going back to the setting of Proposition 5.0.1, let $f : \Sigma \to \Sigma$ be the shift map on $\Sigma = \{g, h\}^{\mathbb{Z}}$ and $\mu = \nu^{\mathbb{Z}}$ be the Bernoulli measure defined by ν. Define

$$\phi_1 = \phi_2 = \phi_3 = \begin{pmatrix} -2 & -1 \\ 1 & 0 \end{pmatrix} \text{ and } \psi_1 = \psi_2 = \psi_3 = \begin{pmatrix} -1/2 & -1 \\ 1 & 0 \end{pmatrix}.$$

Observe that $g = \phi_3\psi_2\phi_1$ and $h = \psi_3\phi_2\psi_1$.

Let $f_P : \Sigma_P \to \Sigma_P$ be the subshift of finite type defined on the alphabet $\mathcal{A} = \{\phi_1, \phi_2, \phi_2, \psi_1, \psi_2, \psi_3\}$ by the aperiodic stochastic matrix

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 \end{pmatrix},$$

and let μ_P be the corresponding Markov measure on Σ_P. Keep in mind that (f_P, μ_P) is mixing and so (f_P^n, μ_P) is ergodic for every $n \geq 1$.

Clearly, the elements of Σ_P are the sequences $(y_n)_n \in \mathcal{A}^{\mathbb{Z}}$ such that each block $(y_{3k}, y_{3k+1}, y_{3k+2})$, $k \in \mathbb{Z}$ is equal to either (ϕ_1, ψ_2, ϕ_3) or (ψ_1, ϕ_2, ψ_3). The map $h : \Sigma_P \to \Sigma$ that replaces each block $(y_{3k}, y_{3k+1}, y_{3k+2})$ with the corresponding product $y_{3k+2}y_{3k+1}y_{3k}$ is a bijection from Σ_P to Σ, conjugating the third iterate $f_P^3 : \Sigma_P \to \Sigma_P$ to the shift map $f : \Sigma \to \Sigma$. Moreover, h maps the Markov measure μ_P to the Bernoulli measure $\mu = \nu^{\mathbb{Z}}$ on Σ.

Consider the linear cocycles

$$F : \Sigma \times \mathbb{R}^2 \to \Sigma \times \mathbb{R}^2, \quad (x, v) \mapsto (f(x), A(x)v) \text{ and}$$
$$F : \Sigma_P \times \mathbb{R}^2 \to \Sigma_P \times \mathbb{R}^2, \quad (y, v) \mapsto (f_P(y), A_P(y)v),$$

where $A : \Sigma \to \mathrm{SL}(2)$ and $A_P : \Sigma_P \to \mathrm{SL}(2)$ are the locally constant functions defined by

$$A((x_n)_n) = x_0 \text{ and } A_P((y_n)_n) = y_0.$$

The map $H = h \times \mathrm{id} : \Sigma_P \times \mathbb{R}^2 \to \Sigma \times \mathbb{R}^2$ conjugates F_P^3 to F.

For each $E \in \mathbb{R}$ and $i = 1, 2, 3$, define

$$\phi_i(E) = \begin{pmatrix} -2 - E & -1 \\ 1 & 0 \end{pmatrix} \text{ and } \psi_i(E) = \begin{pmatrix} -1/2 - E & -1 \\ 1 & 0 \end{pmatrix}.$$

and $\nu(E) = (\delta_{g(E)} + \delta_{h(E)})/2$, where

$$g(E) = \phi_3(E)\psi_2(E)\phi_1(E) = \begin{pmatrix} 2 - 4E - (9/2)E^2 - E^3 & -(5/2)E - E^2 \\ (5/2)E + E^2 & 1/2 + E \end{pmatrix}$$
$$h(E) = \psi_3(E)\phi_2(E)\psi_1(E) = \begin{pmatrix} 1/2 - (1/4)E - 3E^2 - E^3 & -(5/2)E - E^2 \\ (5/2)E + E^2 & 2 + E \end{pmatrix}.$$

It follows from (2.1.1) that there exists $D > 0$ such that

$$\delta_1(\nu(E'), \nu(E'')) \leq D|E' - E''| \text{ for every } E', E'' \text{ close to zero.} \tag{5.2.1}$$

Next, define

$$F(E) : \Sigma \times \mathbb{R}^2 \to \Sigma \times \mathbb{R}^2, \quad (x, v) \mapsto (f(x), A(E, x)v) \text{ and}$$
$$F_P(E) : \Sigma_P \times \mathbb{R}^2 \to \Sigma_P \times \mathbb{R}^2, \quad (y, v) \mapsto (f_P(y), A_P(E, y)v),$$

with

$$A\,(E, (x_n)_n) = x_0(E) \text{ and } A_P\,(E, (y_n)_n) = y_0(E).$$

$F(E)$ corresponds to the random multiplication of matrices associated to the probability measure $\nu(E)$, whereas $F_P(E)$ is the Schrödinger cocycle associated to f_P, μ_P and the potential

$$V\,((y_n)_n) = \begin{cases} -2 & \text{if } y_0 \in \{\phi_1, \phi_2, \phi_3\} \\ -1/2 & \text{if } y_0 \in \{\psi_1, \psi_2, \psi_3\} \end{cases}$$

Let $\lambda_\pm(F(E))$ denote the Lyapunov exponents of $F(E)$ with respect to the measure μ, and $\lambda_\pm(F_P(E))$ denote the Lyapunov exponents of $F_P(E)$ with respect to μ_P. Then $\lambda_\pm(F(0), \mu) = \lambda_\pm(\nu) = 0$ and

$$\lambda_\pm(F_P(E)) = \frac{1}{3}\lambda_\pm(F_P(E)^3) = \lambda_\pm(F(E)) = \lambda_\pm(\nu(E)) \text{ for every } E \in \mathbb{R}, \tag{5.2.2}$$

because $H = h \times \mathrm{id}$ conjugates $F_P(E)^3$ to $F(E)$. We are going to prove:

Proposition 5.2.1. *For any $\beta > 2$ and $C > 0$ there exist E', E'' arbitrarily close to $E = 0$ such that*

$$|N(E') - N(E'')| \geq C \left(\log \frac{1}{|E' - E''|}\right)^{-\beta} \tag{5.2.3}$$

Proposition 5.0.1 is a consequence. Indeed, let $\gamma > 1$ and $\beta > 2$ be fixed. Proposition 5.2.1 means that $E \mapsto N(E)$ is not $(1, \beta)$-log-Hölder continuous, and so it is also not (γ, β)-log-Hölder continuous, near $E = 0$. Then, by the result of Goldstein, Schlag quoted in the previous section, $E \mapsto \lambda_\pm(F_P(E))$ is not (γ, β)-log-Hölder continuous near $E = 0$. Using (5.2.2), it follows that $E \mapsto \lambda_+(\nu(E))$ is also not (γ, β)-log-Hölder continuous near $E = 0$. By (5.2.1), this gives Proposition 5.0.1 with ν' and ν'' of the form $\nu' = \nu(E')$ and $\nu'' = \nu(E'')$.

5.3 Proof of Proposition 5.2.1

See Simon, Taylor [ST85, Lemma A.3.2] for a proof of the following fact:

Lemma 5.3.1. *Let A be a self-adjoint operator and $\{f_j\}_{j=1}^k$ be an orthonormal family such that, for some $\epsilon > 0$ and $E_0 \in \mathbb{R}$,*

$$\|(A - E_0)f_i\| \le \epsilon \text{ and } f_i \cdot Af_i = Af_i \cdot Af_j = 0 \text{ if } i \ne j. \tag{5.3.1}$$

Then A has at least k eigenvalues (counted with multiplicity) in $[E_0-\epsilon, E_0+\epsilon]$.

By construction, the product $y_{3k}y_{3k+1}y_{3k+2}$ is equal to either g or h, for any $y \in \Sigma_P$ and $k \in \mathbb{Z}$. Consequently, given any $y \in \Sigma_P$ and $n \ge 1$,

$$y_{3n-1} \cdots y_0 = \begin{pmatrix} 2^{\sum_{k=0}^{n-1} X_k} & 0 \\ 0 & 2^{-\sum_{k=0}^{n-1} X_k} \end{pmatrix}$$

with

$$X_k = \begin{cases} 1 & \text{if } y_{3k}y_{3k+1}y_{3k+2} = g \\ -1 & \text{if } y_{3k}y_{3k+1}y_{3k+2} = h. \end{cases}$$

Lemma 5.3.2. *There is $a > 0$ such that*

$$\mu_P\left(\left\{y \in \Sigma_P : \sum_{k=0}^{n-1} X_k \ge a\sqrt{n}\right\}\right) > \frac{1}{10} \text{ for every } n \ge 16.$$

Proof. Since the variables X_k are independent and identically distributed, with $\mathbb{E}(X_k) = 0$ and $\mathbb{E}(X_k^2) = 1 = \mathbb{E}(|X_k|^3)$, it follows from the Berry-Esseen theorem that

$$\mu_P\left(\left\{y \in \Sigma_P : \sum_{k=0}^{n-1} X_k \ge a\sqrt{n}\right\}\right) \ge \mathcal{N}(a, \infty) - \frac{1}{\sqrt{n}}$$

for every $a \in \mathbb{R}$ and $n \ge 1$. Fix $a > 0$ such that $\mathcal{N}(a, \infty) > 4/5$. □

Let $a > 0$ and $n \ge 16$ be fixed from now on. At the very end we let $n \to \infty$. Define

$$\Gamma_n(a) = \left\{y \in \Sigma_P : \sum_{k=0}^{n-1} X_k \ge a\sqrt{n} \text{ and } \sum_{k=n}^{2n-1} X_k \le -a\sqrt{n}\right\}.$$

Since the variables $X_0, \ldots, X_{n-1}, -X_n, \ldots, -X_{2n-1}$ are independent and identically distributed,

$$\mu_P(\Gamma_n(a)) = \mu_P\left(\left\{y \in \Sigma_P : \sum_{k=0}^{n-1} X_k \ge a\sqrt{n}\right\}\right)^2 > \frac{1}{100}.$$

Lemma 5.3.3. *Given $y \in \Gamma_n(a)$, there exists $(\xi_j)_j \in \ell^2(\mathbb{Z})$ such that*

1. *$\xi_j = 0$ for every $j \notin [0, 6n]$ and $\|\xi\| \ge 1$;*

2. *$(H_y\xi)_j = 0$ for every $j \notin \{-1, 6n, 6n+1\}$ and $\|H_y\xi\| \leq 3 \cdot 2^{-a\sqrt{n}}$.*

Proof. Consider the sequence $(\xi_j)_j \in \ell^2(\mathbb{Z})$ defined as follows:

(a) $\xi_j = 0$ for $j < 0$;

(b) $\xi_0 = 2^{-X_0 - \cdots - X_{n-1}}$;

(c) $\xi_{j+1} = V(f_P^j(y))\xi_j - \xi_{j-1}$ for $0 \leq j \leq 6n-1$

(d) $\xi_j = 0$ for $j > 6n$.

The first part of condition (1) is obvious from the definition. To prove the second part, note that the rule (c) may be written as

$$\begin{pmatrix} \xi_{j+1} \\ \xi_j \end{pmatrix} = \begin{pmatrix} V(f_P^j(y)) & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \xi_j \\ \xi_{j-1} \end{pmatrix} = y_j \begin{pmatrix} \xi_j \\ \xi_{j-1} \end{pmatrix}$$

for $0 \leq j \leq 3n-1$. In particular,

$$\begin{aligned} \begin{pmatrix} \xi_{3n} \\ \xi_{3n-1} \end{pmatrix} &= y_{3n-1} \cdots y_0 \begin{pmatrix} \xi_0 \\ \xi_{-1} \end{pmatrix} \\ &= \begin{pmatrix} 2^{\sum_{k=0}^{n-1} X_k} & 0 \\ 0 & 2^{-\sum_{k=0}^{n-1} X_k} \end{pmatrix} \begin{pmatrix} 2^{-\sum_{k=0}^{n-1} X_k} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

This gives $\xi_{3n} = 1$ and so $\|\xi\| \geq 1$. This completes the proof of (1).

Analogously,

$$\begin{aligned} \begin{pmatrix} \xi_{6n} \\ \xi_{6n-1} \end{pmatrix} &= y_{6n-1} \cdots y_{3n} \begin{pmatrix} \xi_{3n} \\ \xi_{3n-1} \end{pmatrix} \\ &= \begin{pmatrix} 2^{\sum_{k=n}^{2n-1} X_k} & 0 \\ 0 & 2^{-\sum_{k=n}^{2n-1} X_k} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2^{\sum_{k=n}^{2n-1} X_k} \\ 0 \end{pmatrix} \end{aligned}$$

which implies that $\xi_{6n-1} = 0$ and $\xi_{6n} = 2^{\sum_{k=n}^{2n-1} X_k}$. Then, by the definition of the Schrödinger operator,

$$\begin{aligned} (H_y\xi)_j &= 0 \text{ for } j < -1 \\ (H_y\xi)_{-1} &= -\xi_0 = -2^{-\sum_{k=0}^{n-1} X_k} \\ (H_y\xi)_j &= -\xi_{j+1} - \xi_{j-1} + V(f_P^j(y))\xi_j = 0 \text{ for } 0 \leq j \leq 6n-1 \\ (H_y\xi)_{6n} &= V(f_P^{6n}(y))\xi_{6n} \\ (H_y\xi)_{6n+1} &= -\xi_{6n} = -2^{\sum_{k=n}^{2n-1} X_k} \\ (H_y\xi)_j &= 0 \text{ for } j > 6n+1. \end{aligned}$$

This proves the first part of condition (2). Moreover, using the definition of $\Gamma_n(a)$ and the fact that $|V| \le 2$,

$$\begin{aligned}\|H_y\xi\|^2 &= \|(H_y\xi)_{-1}\|^2 + \|(H_y\xi)_{6n}\|^2 + \|(H_y\xi)_{6n+1}\|^2 \\ &\le 2^{-2a\sqrt{n}} + V(f_P^{6n}(y))^2 2^{-2a\sqrt{n}} + 2^{-2a\sqrt{n}} \le 6 \cdot 2^{-2a\sqrt{n}}\end{aligned}$$

which implies the remaining claim. □

For every $y \in \Sigma_P$ and $q \ge 1$, define

$$l_q(y) = \#\left\{0 \le i \le q-1 : f_P^{6(n+1)i+1}(y) \in \Gamma_n(a)\right\}.$$

Lemma 5.3.4. *Let $L = 6(n+1)q$ for some $q \ge 1$. For any $y \in \Sigma_P$, the L-truncation $H_y^{(L)}$ has at least $l_q(y)$ eigenvalues (counted with multiplicity) in $[-3 \cdot 2^{-a\sqrt{n}}, 3 \cdot 2^{-a\sqrt{n}}]$.*

Proof. For each $0 \le i \le q-1$ such that $f_P^{6(n+1)i+1}(y) \in \Gamma_n(a)$, let be $\xi^i \in \ell^2(\mathbb{Z})$ as in Lemma 5.3.3 with y replaced with $f_P^{3(n+1)i+1}(y)$. Since ξ^i is supported in $[6(n+1)i+1, 6(n+1)i+6n+1] \subset [0, L-1]$, it has exactly the same non-zero coefficients as $P_L\xi^i \in \ell^2(0,\ldots,L-1)$. Moreover, $H_y\xi^i$ is supported in $\{6(n+1)i, 6(n+1)i+6n+1, 6(n+1)i+6n+2\}$ and so it has the same non-zero coefficients as $P_L H_y \xi = H_y^{(L)} P_L \xi^i$. In particular,

$$\|P_L\xi^i\| = \|\xi^i\| \ge 1 \text{ and } \|H_y^{(L)} P_L \xi^i\| = \|H_y\xi^i\| \le 3 \cdot 2^{-a\sqrt{n}}.$$

Define $\psi^i = P_L\xi^i / \|P_L\xi^i\| \in \ell^2(0,\ldots,L-1)$. It follows from the previous observations that

(i) $\|\psi^i\| = 1$;

(ii) ψ^i is supported in $[6(n+1)j+1, 6(n+1)j+6n+1]$;

(iii) $\|H_y^{(L)}\psi^i\| = \|H_y\xi^i\|/\|\xi^i\| \le \|H_y\xi^i\| \le 3 \cdot 2^{-a\sqrt{n}}$

(iv) $H_y^{(L)}\psi^i$ is supported in $\{6(n+1)i, 6(n+1)i+6n+1, 6(n+1)i+6n+2\}$.

The supports of ψ^i and $H_y^{(L)}\psi^i$ are contained in $J_i = [6(n+1)i, 6(n+1)(i+1)-1]$ and these intervals J_i are pairwise disjoint. It is easy to deduce that the family $\{\psi^i\}$ satisfies the hypotheses of Lemma 5.3.1 for $A = H_y^{(L)}$, $E_0 = 0$ and $\epsilon = 3 \cdot 2^{-a\sqrt{n}}$: properties (i) and (ii) imply that the family $\{\psi^i\}$ is orthonormal; (ii) and (iv) ensure that $\psi^i \cdot H_y^{(L)}\psi^j = H_y^{(L)}\psi^i \cdot H_y^{(L)}\psi^j = 0$ if $i \ne j$; and (iii) gives the remaining hypothesis. Now, since $\{\psi^i\}$ contains $l_q(y)$ elements, the conclusion follows from Lemma 5.3.1. □

Proof of Proposition 5.2.1. Fix $y \in \Sigma_P$ such that the integrated density of states is well defined and

$$\lim_q \frac{l_q}{q} = \mu_P\big(\Gamma_n(a)\big) > \frac{1}{100}$$

(the latter holds μ-almost everywhere, because $(f_P^{6(n+1)}, \mu_P)$ is ergodic. Then, by Lemma 5.3.4,

$$N(3 \cdot 2^{-a\sqrt{n}}) - N(-3 \cdot 2^{-a\sqrt{n}}) \geq \lim_q \frac{l_q}{L} = \frac{1}{6(n+1)} \lim_q \frac{l_q}{q} > \frac{1}{600(n+1)}.$$

Let $E' = 3 \cdot 2^{-a\sqrt{n}}$ and $E'' = -3 \cdot 2^{-a\sqrt{n}}$. For any $\beta > 2$,

$$\left(\log \frac{1}{|E' - E''|}\right)^{-\beta} = \frac{1}{\big((a \log 2)\sqrt{n} - \log 6\big)^{\beta}} \ll \frac{1}{600(n+1)}$$

if n is large. Moreover, E' and E'' go to to zero when $n \to \infty$. □

Bibliography

[AEV] A. Avila, A. Eskin, and M. Viana. Continuity of Lyapunov exponents for products of random matrices. In preparation.

[AV10] A. Avila and M. Viana. Extremal Lyapunov exponents: an invariance principle and applications. *Invent. Math.*, 181(1):115–189, 2010.

[BDM16] D. Buraczewski, E. Damek, and T. Mikosch. *Stochastic models with power-law tails.* Springer Series in Operations Research and Financial Engineering. Springer, 2016. The equation $X = AX + B$.

[BGMV03] 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.

[BV17] C. Bocker and M. Viana. Continuity of Lyapunov exponents for random two-dimensional matrices. *Ergodic Theory Dynam. Systems*, 37:1413–1442, 2017.

[CS83] W. Craig and B. Simon. Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices. *Comm. Math. Phys.*, 90:207–218, 1983.

[Dam17] D. Damanik. Schrödinger operators with dynamically defined potentials. *Ergodic Theory Dynam. Systems*, 37:1681–1764, 2017.

[DE86] A. Douady and J. C. Earle. Conformally natural extension of homeomorphisms of the circle. *Acta Math.*, 157:23–48, 1986.

[DK16] P. Duarte and S. Klein. *Lyapunov exponents of linear cocycles*, volume 3 of *Atlantis Studies in Dynamical Systems.* Atlantis Press, 2016. Continuity via large deviations.

[DKS] P. Duarte, S. Klein, and M. Santos. A random Schrödinger cocycle with non Hölder Lyapunov exponent. Preprint 2017.

[Dur10] R. Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, fourth edition, 2010.

[EK47] P. Erdös and M. Kac. On the number of positive sums of independent random variables. *Bull. Amer. Math. Soc.*, 53:1011–1020, 1947.

[FK60] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.

[GM89] I. Ya. Gol'dsheid and G. A. Margulis. Lyapunov indices of a product of random matrices. *Uspekhi Mat. Nauk.*, 44:13–60, 1989.

[GM00] C. Goldie and R. Maller. Stability of perpetuities. *Ann. Probab.*, 28:1195–1218, 2000.

[Gri74] A. K. Grincevičjus. The continuity of the distribution of a certain sum of dependent variables that is connected with independent walks on lines. *Teor. Verojatnost. i Primenen.*, 19:163–168, 1974.

[Gri75] A. K. Grincevičjus. Limit theorem for products of random linear transformations of the line. *Litovsk. Mat. Sb.*, 15:61–77, 241, 1975.

[GS01] M. Goldstein and W. Schlag. Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. *Annals of Math.*, 154:155–203, 2001.

[Pag89] 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.

[Pet87] V. V. Petrov. Inequalities for the maximum of partial sums of random variables and the law of the iterated logarithm. In *Probability theory and mathematical statistics, Vol. II (Vilnius, 1985)*, pages 461–465. VNU Sci. Press, 1987.

[Pet95] V. V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability.* The Clarendon Press, Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications.

[Sch81] E. Schneider. On the speed of convergence in the random central limit theorem for φ-mixing processes. *Z. Wahrsch. Verw. Gebiete*, 58:125–138, 1981.

[ST85] B. Simon and M. Taylor. Harmonic analysis on $\mathrm{SL}(2, \mathbf{R})$ and smoothness of the density of states in the one-dimensional Anderson model. *Comm. Math. Phys.*, 101:1–19, 1985.

[Via] M. Viana. (Dis)continuity of Lyapunov exponents. In preparation.

[Via14] M. Viana. *Lectures on Lyapunov exponents*, volume 145 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2014.

[Vil09] C. Villani. *Optimal transport, old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, 2009.

[VO15] M. Viana and K. Oliveira. *Foundations of Ergodic Theory*. Cambridge University Press, 2015.