Upper Bound on the Regularity of the Lyapunov Exponent for Random Products of Matrices

Jamerson Bezerra¹, Pedro Duarte²

¹ Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland. E-mail: jdbezerra@mat.unik.pl
² FCUL-Faculdade de Ciências da Universidade de Lisboa, Campo grande, 1749-016 Lisbon, Portugal. E-mail: pedromiguel.duarte@gmail.com

Received: 19 December 2022 / Accepted: 10 July 2023
Published online: 8 August 2023 – © The Author(s) 2023

Abstract: We prove that if $\mu$ is a finitely supported measure on $SL_2(\mathbb{R})$ with positive Lyapunov exponent but not uniformly hyperbolic, then the Lyapunov exponent function is not $\alpha$-Hölder around $\mu$ for any $\alpha$ exceeding the Shannon entropy of $\mu$ over the Lyapunov exponent of $\mu$.

1. Introduction

The law of large numbers, stating that on average an i.i.d. process is close to its theoretical mean, often used to describe the typical statistical behavior of a random sample, is the basis for understanding general additive processes with applications in various branches of mathematics such as probability, combinatorics or ergodic theory.

The multiplicative version of the law of large numbers for products of random matrices is the classical theorem of Furstenberg and Kesten [1], which asserts that with probability 1 the logarithmic growth rate of products of random matrices equals its mean growth rate. More formally, a special case of this theorem states that for an i.i.d. sequence of random matrices $L_1, L_2, \ldots$, with common law given by a compactly supported probability measure $\mu$ on $GL_d(\mathbb{R})$, the following asymptotic equality holds almost surely

$$
\lim_{n \to \infty} \frac{1}{n} \log \|L_n \cdots L_1\| = \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[\log \|L_n \cdots L_1\|],
$$

where the right-hand-side, denoted by $L(\mu)$, is the so called Lyapunov exponent of the law $\mu$. The investigation of how the Lyapunov exponent changes as a function of the underlying measure $\mu$ lies at the core of the multiplicative ergodic theory, with many fundamental contributions during the last 60 years.

Jamerson Bezerra and Pedro Duarte contributed equally to this work.
The continuity of the Lyapunov exponent $L(\mu)$ as a function of the measure $\mu$, with respect to the weak* topology, was established by Furstenberg and Kifer [2] under a generic irreducibility assumption. A measure $\mu$ is called irreducible if there exists no proper subspace of $\mathbb{R}^d$ which is $\mu$-invariant, i.e., invariant under all matrices in the support of $\mu$. Otherwise $\mu$ is called reducible, and any proper $\mu$-invariant subspace $S \subset \mathbb{R}^d$ determines the Lyapunov exponent $L(\mu|_S)$ corresponding to the logarithmic growth rate of the norms $\|(L_n|_S) \ldots (L_1|_S)\|$. The continuity of Furstenberg and Kifer actually holds under the weaker quasi-irreducibility assumption. A measure $\mu$ is called quasi-irreducible if $L(\mu|_S) = L(\mu)$ for every proper $\mu$-invariant subspace $S \subset \mathbb{R}^d$. See [3, Theorem 1.46].

In [4], Bocker and Viana proved that for measures supported in $GL_2(\mathbb{R})$ the Lyapunov exponent $\mu \mapsto L(\mu)$ is continuous with respect the weak* topology and the Hausdorff distance between their supports. Avila, Eskin and Viana announced that the same result holds for measures supported in $GL_d(\mathbb{R})$, any $d \geq 2$. See the remark after Theorem 10.1 in [5].

It is then natural to raise the question about the precise modulus of continuity of this map.

A lower bound for this regularity was provided by Le Page in [6]. The Lyapunov exponent is locally Hölder continuous over an open and dense set of compactly supported measures on $GL_d(\mathbb{R})$, namely the set of quasi-irreducible measures $\mu$ with a gap between the first and second Lyapunov exponents. See also [7, Theorem 1]. Recall that a function $E \mapsto f(E)$ is said to be Hölder with exponent $\alpha$, or $\alpha$-Hölder, if there exists a constant $C < \infty$ such that for all $E, E'$,

$$|f(E) - f(E')| \leq C |E - E'|^\alpha.$$ 

A function which is Hölder in a neighborhood of each point of its domain is called locally Hölder. Alternatively, we say that a function $E \mapsto f(E)$ is point-wisely $\alpha$-Hölder if for every $E_0$ there exists a constant $C < \infty$ and a neighborhood of $E_0$ where for all $E$,

$$|f(E) - f(E_0)| \leq C |E - E_0|^\alpha.$$ 

Notice that point-wise Hölder is weaker than locally Hölder. In fact the modulus of continuity around a point of a point-wisely Hölder function can be arbitrary bad.

In E. Tall and M. Viana [8] proved that for random $GL_2(\mathbb{R})$ cocycles the Lyapunov exponents are always point-wisely log-Hölder, and even point-wisely Hölder when the Lyapunov exponents are distinct.

In the same direction, the quasi-irreducibility hypothesis was discarded in [9], where it was established that for finitely supported measures in $GL_2(\mathbb{R})$ with distinct Lyapunov exponents, the function $\mu \mapsto L(\mu)$ is either locally Hölder or else locally weak-Hölder. Given positive constants $\alpha, \beta \leq 1$, a function $E \mapsto f(E)$ is said to be $(\alpha, \beta)$-weak Hölder if there exists a constant $C < \infty$ such that for all $E, E'$,

$$|f(E) - f(E')| \leq C e^{-\alpha (\log |E - E'|)^\beta}.$$ 

Notice that $(\alpha, 1)$-weak Hölder is equivalent to $\alpha$-Hölder.

In the reverse direction, an example due to Halperin [10, Appendix 3A] provides an upper bound on this regularity. The example consists of the following 1-parameter family of measures on $SL_2(\mathbb{R})$, $\mu_{a,b,E} := \frac{1}{2} \delta_A + \frac{1}{2} \delta_B$, where

$$A_E = \begin{pmatrix} a - E - 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad B_E = \begin{pmatrix} b - E & 0 \\ 1 & 0 \end{pmatrix}.$$
It follows from [10, Theorem A.3.1] that the function $E \mapsto L(\mu_{a,b}, E)$ can not be $\alpha$-Hölder continuous for any $\alpha > \frac{2 \log 2}{\text{arccosh}(1+|a-b|/2)}$ (see also Proposition 4.1). On the other hand, it is not difficult to see that the measures $\mu_{a,b}, E$ satisfy the assumptions of Le Page’s theorem, which implies that the function $E \mapsto L(\mu_{a,b}, E)$ is indeed Hölder continuous, but with a very small Hölder exponent $\alpha$ when $a-b$ is large.

In the same spirit, in [11], the authors provide the following example where the Lyapunov exponent is not even weak-Hölder continuous. They consider the measure

$$\mu := \frac{1}{2} \delta_A + \frac{1}{2} \delta_B, \quad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} e & 0 \\ 0 & e^{-1} \end{pmatrix}$$

and prove that there exists a curve $\tilde{\mu}_t = \frac{1}{2} \delta_{A_t} + \frac{1}{2} \delta_{B_t}$ through $\tilde{\mu}_0 = \mu$ such that $t \mapsto L_1(\tilde{\mu}_t)$ is not weak-Hölder around $t = 0$. Notice that $\mu$ is not quasi-irreducible and $L(\mu) = 0$ so that $\mu$ does not satisfy any of the assumptions of Le Page’s theorem.

In contrast with this low regularity, a classical theorem of Ruelle proves the analyticity of the Lyapunov exponent for uniformly hyperbolic measures with 1-dimensional unstable direction, see [12, Theorem 3.1]. A compactly supported measure $\mu$ on $GL_d(\mathbb{R})$ is said to be uniformly hyperbolic if the linear cocycle generated by $\mu$ is uniformly hyperbolic (see Sect. 3.1).

From now on we focus on the class of finitely supported measures in $SL_2(\mathbb{R})$, where the lack of regularity of the Lyapunov exponent can only occur outside of the class of uniformly hyperbolic measures. In [13], Avila, Bochi and Yoccoz gave a characterization of the uniformly hyperbolic cocycles generated by a finitely supported measure in $SL_2(\mathbb{R})$ in terms of existence of an invariant multicone. With this characterization they prove that the complement of the closure of the uniformly hyperbolic measures is the set of elliptic measures, meaning the finitely supported measures such that the semigroup $\Gamma_\mu$ generated by the support of $\mu$ contains a elliptic element, i.e., a matrix conjugated to a rotation.

Given a hyperbolic matrix $A \in SL_2(\mathbb{R})$ we denote by $\hat{s}(A)$, respectively $\hat{u}(A)$, the stable direction, respectively the unstable direction of $A$ in the projective space $\mathbb{P}^1$. We say that $\mu$ has a heteroclinic tangency if there are matrices $A, B, C \in \Gamma_\mu$ such that $A$ and $B$ are hyperbolic and $C \hat{u}(B) = \hat{s}(A)$. In this case we also say that $(B, C, A)$ is a tangency for $\mu$. If moreover $A = B$, we say that $\mu$ has a homoclinic tangency. Heteroclinic tangencies are referred to as heteroclinic connections in [13]. If $\mu$ is not uniformly hyperbolic but $L(\mu) > 0$, i.e., if $\mu$ is non-uniformly hyperbolic, then $\Gamma_\mu$ contains hyperbolic matrices. By Theorem 4.1 of [13], in this case the semigroup $\Gamma_\mu$ contains either a heteroclinic tangency or else a non hyperbolic matrix, i.e., an elliptic or parabolic matrix. In each of these two cases we can produce heteroclinic tangencies with an arbitrary small perturbation. See Proposition 7.8. Hence measures with heteroclinic tangencies are dense in the class of non-uniformly hyperbolic measures.

1.1. Results. Let $H(\mu)$ be the Shannon’s entropy (see Sect. 3.3) of the finitely supported measure $\mu$.

**Theorem A.** Let $\mu$ be a finitely supported measure on $SL_2(\mathbb{R})$. Assume that $L(\mu) > 0$, $\mu$ is irreducible and that $\mu$ has a heteroclinic tangency. Then, there exists an analytic

\[1\] In this work Avila, Bochi and Yoccoz characterize the boundary of uniformly hyperbolic cocycles. By [13, Remark 4.2] heteroclinic connections can never be homoclinic connections for cocycles at the boundary of the uniformly hyperbolic ones.
one parameter family of finitely supported measures $\{\mu_E\}_E$ such that $\mu_0 = \mu$ and for any $\alpha > H(\mu)/L(\mu)$, the function $E \mapsto L(\mu_E)$ is not locally $\alpha$-Hölder at any neighborhood of $E = 0$.

**Remark 1.** Our result implies a similar conclusion as in Halperin/Simon-Taylor example with a less sharper threshold. For simplicity we consider the parameters $a = 0$ with energy $E = 0$. In this example $H(\mu_{0,0}) = \log 2$ while

$$L(\mu_{0,0}) \leq \frac{1}{2} \log \|B_0\| = \frac{1}{2} \log \sqrt{\frac{2 + b^2 + |b|\sqrt{b^2 + 4}}{2}} < \frac{1}{2} \arccosh \left(1 + \frac{|b|}{2}\right).$$

The last two quantities are asymptotically equivalent, which implies that

$$\frac{H(\mu_{0,b,0})}{\frac{1}{2} \log \|B_0\|} \sim \frac{2 \log 2}{\arccosh(1 + |b|/2)} \quad \text{as } b \to \infty.$$  

Set

$$\alpha_\mu := \sup \{\alpha > 0 : \text{L is locally } \alpha\text{-Hölder around } \mu\}.$$  

**Corollary A.** Let $\mu$ be a finitely supported measure on $SL_2(\mathbb{R})$ with $L(\mu) > 0$. Then, either $\mu$ is uniformly hyperbolic and $L$ is locally analytic around $\mu$, or else

$$\alpha_\mu \leq \frac{H(\mu)}{L(\mu)}.$$  

As a consequence of the proof of A we have the following application in mathematical physics (for precise definitions see Sect. 4).

**Corollary B.** Consider the Anderson model of the discrete Schrödinger operators associated with a finitely supported measure $\mu$. Let $\alpha > \frac{H(\mu)}{L(\mu)}$ and $E_0$ be an energy in the spectrum. Then, the integrated density of states function $E \mapsto N(E)$ and the Lyapunov exponent function $E \mapsto L(E)$ are not $\alpha$-Hölder continuous at any neighborhood of $E_0$.

1.2. Relations with other dimensions. See Sect. 3.3 for a precise description of the objects treated in this subsection.

The study of formulas relating (some type of) dimension, entropy and Lyapunov exponent has a vast history with many contributions in different settings (see for instance [14, 15] for diffeomorphisms of a compact manifold and [16] for self affine measures).

For $SL_2(\mathbb{R})$ supported measures $\mu$ with $L(\mu) > 0$, Ledrappier in [17], proved that we have a dimension type formula for any (forward) stationary measures $\eta$ associated with $\mu$, namely

$$\dim \eta = \min \left\{1, \frac{h_F(\eta)}{2L(\mu)}\right\}. \quad (1)$$
Dim is a different notion of dimension from \( \dim \) given in Sect. 3.3 (See [3, Remark 2.34]), but in the case that \( \eta \) is exact dimensional they coincide. The exactness of the dimension of the stationary measures was established by Hochman and Solomyak in [18] assuming additionally that \( \mu \) is irreducible. In particular, if \( \eta^+ \) and \( \eta^- \) denote respectively the forward and backward stationary measures then

\[
\dim \eta^\pm = \min \left\{ 1, \frac{h_F(\eta^\pm)}{2L(\mu)} \right\} \leq \min \left\{ 1, \frac{H(\mu)}{2L(\mu)} \right\}.
\]

Moreover, in [18] they provided, among other things, conditions to obtain Ledrappier-Young type formulas relating the dimension of the stationary measure, the entropy and the Lyapunov exponents, i.e.,

\[
\dim \eta^\pm = \min \left\{ 1, \frac{H(\mu)}{2L(\mu)} \right\},
\]

where \( h_F(\eta^\pm) \) is the Furstenberg entropy of \( \eta^\pm \). In light of the above discussion we leave the following questions.

**Question 1.** Assume that \( \alpha > \dim \eta^+ + \dim \eta^- \). Under the assumptions of Theorem A, is it true that the Lyapunov exponent is not \( \alpha \)-Hölder continuous in any neighborhood of \( \mu \)?

**Question 2.** Is \( \frac{H(\mu)}{L(\mu)} \) a sharp bound for the regularity? In other words, is there an example where \( \alpha_\mu = \frac{H(\mu)}{L(\mu)} \)?

Halperin’s example above does not answer this question.

**Question 3.** In the case that \( \frac{H(\mu)}{L(\mu)} \geq 1 \), is it true that the Lyapunov exponent is Lipschitz continuous function around \( \mu \)?

**Question 4.** Is it possible to express the lower bound for the regularity in terms of some of the previous measurements?

### 1.3. Sketch of the proof and organization.

In Mathematical Physics the Thouless formula (5) relates the Lyapunov exponent of a Schrödinger cocycle with the integrated density of sates (IDS) of the corresponding Schrödinger operator. It follows from this identity (5) that the Lyapunov exponent and the IDS, as functions of the energy, share the same modulus of continuity. See Proposition 4.1. The IDS is a spectral quantity that measures the asymptotic distribution of the eigenvalues of truncation matrices of the Schrödinger operator as the size of the truncation tends to infinity. The strategy to break the Hölder regularity of the IDS in Halperin’s example is to establish around a certain energy a very large concentration of eigenvalues of the Schrödinger truncated matrices which implies a disproportionately large leap of the IDS around that energy, see [10, Appendix 3]. Then, as explained above, the loss of Hölder regularity passes from the IDS to the Lyapunov exponent.

Let \( \mu \) be an irreducible and finitely supported measure on \( SL_2(\mathbb{R}) \) with positive Lyapunov exponent and \( A : \Omega \to SL_2(\mathbb{R}) \) be the associated locally constant cocycle. In order to use the strategy described above, in Sect. 5 we embed the cocycle \( A \) into a family of locally constant Schrödinger cocycles over a Markov shift.
We call matching to a configuration where the horizontal direction \( e_1 = (1, 0) \) is mapped in \( n \) iterations to the vertical direction \( e_2 = (0, 1) \) in a way that that \( e_1 \) is greatly expanded in the first half iterations followed by a similar contraction in the second half iterations. The number \( n \) is referred to as the size of the matching. A matching of size \( n \) at some energy \( E_0 \) determines an almost eigenvector for an \( n \times n \) truncated Schrödinger matrix, which then implies a true nearby eigenvalue \( E_0^* \approx E_0 \) of the same matrix. Hence matchings of size \( n \) for energies in some small interval \( I \) can be used to count eigenvalues of a truncated Schrödinger operator of size \( n \).

By Proposition 7.14, a heteroclinic tangency of the cocycle \( A = A(0) \) implies many nearby matchings of any chosen large size \( n \), spreading through a small interval of length \( \sim e^{-cn} \). Because these matchings are still not enough to break the Hölder regularity in the stated form, we prove in Proposition 7.15 that a single tangency will cause many more tangencies to occur at nearby energies, which are in some sense typical. Propositions 7.14 and 7.15 were designed to be used recursively in the sense that the output of the second feeds the input of the first. They could be used recursively to characterize the fractal structure of matchings and tangencies, a path we do not explore in this work. We do use them in a single cycle to gather the matchings, of some appropriate size, associated to a typical nearby heteroclinic tangency. The matchings coming from a typical tangency are now enough to break the Hölder regularity in the stated form.

Proposition 7.11 plays a key role in the proof of Theorem A, to estimate the number of matchings and tangencies from Propositions 7.14 and 7.15. On the other hand the proof of Proposition 7.11 relies on a characterization of the projective random walk distribution in Proposition 7.10 and a few Linear Algebra facts on the geometry of the projective action in Appendix A. See propositions A.3, A.6 and Lemma A.9.

**Organization.** This work is organized as follows. Section 2 contains the general definitions that will be used throughout the paper. In Sect. 3 we define and state some properties of locally constant linear cocycles and Furstenberg measures. We discuss general spectral properties of Schrödinger operators in Sects. 4 and 5 we show how to embed a general locally constant cocycle into a Schrödinger family over a Markov shift. In Sect. 6 we obtain a lower bound for the oscillation of the integrated density of states in terms of counting matchings. Section 7 contains the core technical results of the work, namely propositions 7.11, 7.14 and 7.15. Section 8 provides lower bounds for the measure of the set of matchings. In Sect. 9 we give the proof of the results. The Appendix A contains the linear algebra tools needed in this work and Appendix B describes some of the formulas for derivatives of projective actions.

**Logical structure** Figure 1 describes the logical structure of the proof of Theorem A.

### 2. Basic Definitions and General Concepts

In this subsection we establish some of the general notation used throughout this work.

#### 2.1. Preliminary definitions and notations.

- We denote by \( GL_d(\mathbb{R}) \) and \( SL_d(\mathbb{R}) \) respectively the group of \( d \times d \) invertible matrices and its subgroup of matrices with determinant one. Given a \( d \times d \) square matrix \( H \) we denote its spectrum by \( \text{Spec}(H) \) and by \( |\text{Spec}(H)| \) the number of
elements in Spec(H) counted with multiplicity. Unless otherwise stated, \( \| H \| \) refers to the operator norm of the matrix \( H \).

- The projective space of \( \mathbb{R}^2 \), consisting of all lines in \( \mathbb{R}^2 \), is denoted by \( P^1 \). Its points are denoted by \( \hat{v} \in P^1 \). After introducing a projective point \( \hat{v} \), by convention the letter \( v \) will stand for any unit vector aligned with the line \( \hat{v} \). A natural distance in \( P^1 \) is given by \( d(\hat{v}, \hat{w}) := |v \wedge w| = \sin(\angle(\hat{v}, \hat{w})) \).

Each \( A \in SL_2(\mathbb{R}) \) induces a projective automorphism \( \hat{A} : P^1 \to P^1 \), where \( \hat{A} \hat{v} := \hat{A} v \) is the line determined by the unit vector \( A v / \| A v \| \). For the sake of notational simplicity we often write \( A \hat{v} \) instead of \( \hat{A} \hat{v} \).

- We use the standard classification for \( SL_2(\mathbb{R}) \) matrices as elliptic, parabolic or hyperbolic meaning respectively that the absolute value of the trace is smaller than, equal, or greater than two.

- Let \( X \) be a compact metric space. The space of all Borel probability measures on \( X \) is denoted by \( \mathcal{P}(X) \). This is a convex and compact set with respect to the weak* topology. Given a sequence of measures \( \eta_n \in \mathcal{P}(X) \), we say that \( \eta_n \) converges weak*
to $\eta$ in $\mathcal{P}(X)$, and write $\eta_n \rightharpoonup \eta$, if for every continuous function $\varphi \in C^0(X)$,
\[
\int \varphi \, d\eta = \lim_{n \to \infty} \int \varphi \, d\eta_n.
\]

- Given two probability measures $\mu_1, \mu_2 \in \mathcal{P}(SL_2(\mathbb{R}))$, the convolution between $\mu_1$ and $\mu_2$ is the measure

\[
\mu_1 \ast \mu_2 := \int_{SL_2(\mathbb{R})} g \ast \mu_2 \, d\mu_1(g).
\]

The $n$-th convolution power, $\mu^{*n}$, of a measure $\mu \in SL_2(\mathbb{R})$ is defined inductively by
\[
\mu^{*n} := \mu^{*(n-1)} \ast \mu.
\]

- Let $\Lambda$ be a finite set and $\Sigma = \Lambda^\mathbb{Z}$. Given $k \in \mathbb{Z}$ and a finite word $a = (a_0, a_1, \ldots, a_{m-1}) \in \Lambda^m$ the set

\[
[k; a] := \{ \xi \in \Sigma : \xi_{j+k} = a_j, \forall j = 0, 1, \ldots, m-1 \}
\]

is called the cylinder of $\Sigma$ determined by the word $a$ and the position $k$. The integer $m$ is referred to as the length of the cylinder.

- Given a compact metric space $(X, d)$, $0 < \theta < 1$ and a continuous function $\varphi \in C^0(X)$, the $\theta$-Hölder constant of $\varphi$ is defined by

\[
v_\theta(\varphi) := \sup_{\substack{x, y \in X \\neq y}} \frac{|\varphi(x) - \varphi(y)|}{d(x, y)^\theta}.
\]

The space of $\theta$-Hölder continuous functions on $X$ is

\[
C^\theta(X) := \left\{ \varphi \in C^0(X) : v_\theta(\varphi) < \infty \right\}
\]

which endowed with the norm

\[
\|\varphi\|_\theta := \|\varphi\|_\infty + v_\theta(\varphi)
\]

becomes a Banach algebra.

- Given sequences of real numbers $(a_n)$ and $(b_n)$ with $a_n, b_n > 0$ we write
  - $a_n = O(b_n)$ if there exists an absolute constant $C > 0$ and $n_0 \in \mathbb{N}$ such that $a_n \leq C b_n$ for every $n \geq n_0$;
  - $a_n \lesssim b_n$ or $b_n \gtrsim a_n$ if $a_n = O(b_n)$;
  - $a_n \sim b_n$ if $\lim_{n \to \infty} a_n / b_n = 1$.
  - For $\gamma, \tau > 0$ we write $\gamma \gg \tau$ to indicate that $\tau$ is much smaller than $\gamma$.

- Given some interval $J \subset \mathbb{R}$ or $J \subseteq P^1$, we denote by $|J|$ the length (size) of $J$. Given a positive number $t$, we denote by $t J$ the interval with the same center as $J$ and size $t |J|$.
2.2. Linear cocycles. Let $X$ be a compact metric space, with Borel $\sigma$-algebra $\mathcal{B}$. Consider a homeomorphism $T: X \to X$ which preserves a probability measure $\xi$ defined on $\mathcal{B}$ and such that the system $(T, \xi)$ is ergodic. The triple $(X, T, \xi)$ is referred to as the base dynamics.

Any continuous map $A: X \to SL_2(\mathbb{R})$ defines a linear cocycle, over the base dynamics $(X, T, \xi)$, $F_A: X \times \mathbb{R}^2 \to X \times \mathbb{R}^2$ given by $F_A(x, v) := (T x, A(x) v)$. By linearity of the fiber action we can define the projectivization of $F_A$ as the map $\hat{F}_A: X \times \mathbb{P}^1 \to X \times \mathbb{P}^1$ given by $\hat{F}_A(x, \hat{v}) := (T x, \hat{A}(x) \hat{v})$. We also use the term linear cocycle referring to the map $A: X \to SL_2(\mathbb{R})$ when the base dynamics is fixed.

Note that for each $n \in \mathbb{Z}$, the $n$-th iteration of the linear cocycle $F_A$ sends a point $(x, v) \in X \times \mathbb{R}^2$ to $(T^n x, A^n(x)(v))$, where

$$A^n(x) = \begin{cases} A(T^{n-1} x) \ldots A(T x) A(x) & \text{if } n > 0 \\ I & \text{if } n = 0 \\ A(T^n x)^{-1} \ldots A(T^{-2} x)^{-1} A(T^{-1} x)^{-1} & \text{if } n < 0. \end{cases}$$

The Lyapunov exponent of the cocycle $F_A$ can be defined as the limit

$$L(A) = \lim_{n \to \infty} \frac{1}{n} \log \| A^n(x) \|,$$

which exists and is constant for $\xi$-a.e. $x \in X$ as a consequence of Kingman’s sub-additive ergodic theorem. Notice that the Lyapunov exponent depends on the base dynamics despite the fact that the notation $L(A)$ does not refer to $(X, T, \xi)$. The underlying base dynamics should always be clear from the context.

2.3. Transition kernel and stationary measures. Let $X$ be a compact metric space. We call transition kernel to any continuous map $K: X \to \mathcal{P}(X)$. Any transition kernel $K$ induces a linear operator $K: C^0(X) \to C^0(X)$

$$(K \varphi)(x) := \int_X \varphi dK_x,$$

acting on the space $C^0(X)$ of continuous functions $\varphi: X \to \mathbb{R}$. This is called the Markov operator associated to the transition kernel $K$. The adjoint of $K$, $K^*$, in the space of probability measures $\mathcal{P}(X)$ is given by

$$K^* \xi := \int K_x d\xi(x).$$

We say that a probability measure $\xi_0$ is stationary for $K$ if $\xi_0$ is a fixed point of $K^*$, i.e., if for every $\varphi \in C^0(X)$,

$$\int_X \varphi d\xi_0 = \int_X \left( \int_X \varphi(y) dK_x(y) \right) d\xi_0(x).$$

Consider the process $e_n: X^\mathbb{N} \to X$, $e_n(\omega) := \omega_n$. Given a probability measure $\xi \in \mathcal{P}(X)$ there exists a unique measure $\tilde{\xi}$ in the space $X^\mathbb{N}$ such that:

(a) $\tilde{\xi}(e_0^{-1}(E)) = \xi(E), \forall E \in \mathcal{B}$;
(b) $\tilde{\xi}(e_n^{-1}(E) | e_{n-1} = x) = K_x(E), \forall E \in \mathcal{B}, \forall x \in X$. 

We say that the measure \( \tilde{\xi} \) is the **Kolmogorov extension** of the pair \((K, \xi)\). The following statements are equivalent:

1. \( \xi \) is \( K \)-stationary,
2. \( \tilde{\xi} \) is invariant under the shift map \( T : X^\mathbb{N} \to X^\mathbb{N}, \quad T(x_n)_{n \in \mathbb{N}} := (x_{n+1})_{n \in \mathbb{N}}, \)
3. \( e_n : X^\mathbb{N} \to X \) is a stationary Markov process with transition kernel \( K \) and common law \( \xi \).

When these conditions hold we refer to \((K, \xi)\) as a **Markov system**. In this case the Kolmogorov extension \( \tilde{\xi} \) admits a natural extension to \( X^\mathbb{Z} \), still denoted by \( \tilde{\xi} \), for which the two sided process \( e_n : X^\mathbb{Z} \to X, \quad e_n(\omega) := \omega_n, \) is a stationary Markov process. Moreover \( \tilde{\xi} \) is invariant under the two sided shift map \( T : X^\mathbb{Z} \to X^\mathbb{Z}, \)

\[
T(\ldots, x_{-1}, x_0, x_1, x_2, \ldots) := (\ldots, x_0, x_1, x_2, \ldots),
\]

where the bold term in the above expression indicates the 0-th position of the sequence.

The dynamical system \((T, \tilde{\xi})\) is then called the Markov shift over \( X \) induced by the pair \((K, \xi)\).

We say that a Markov system \((K, \xi)\) is **strongly mixing** if

\[
\lim_{n \to \infty} \| K^n \varphi - \int \varphi \, d\xi \|_\infty = 0
\]

with uniform convergence over bounded sets of \( C^0(X) \). It is important to observe that if a Markov system \((K, \xi)\) is strongly mixing then the Markov shift \((T, \tilde{\xi})\) is mixing. See [19, Proposition 5.1].

### 3. Random Product of Matrices

In this section we describe the base dynamics associated with random i.i.d. products of matrices generated by a probability measure on \( SL_2(\mathbb{R}) \).

In the subsequent sections \( \mu \) is a probability measure on \( SL_2(\mathbb{R}) \) with finite support given by \( \text{supp} \, \mu = \{A_1, \ldots, A_\kappa\} \subset SL_2(\mathbb{R}) \). We write

\[
\mu = \sum_{i=1}^\kappa \mu_i \delta_{A_i},
\]

where the components \( \mu_i := \mu([A_i]) > 0 \).

**Remark 2.** The positivity requirements \( \mu_i > 0 \) avoids discontinuities as in Kifer counter-example. See [20] or [4, Remark 7.5].
3.1. General locally constant cocycles.

Locally constant cocycles. Let \( \Omega = \{1, \ldots, \kappa\}^\mathbb{Z} \) be the space of sequences in the symbols \( \{1, \ldots, \kappa\}, \bar{\mu} = (\mu_1, \ldots, \mu_\kappa)^{\mathbb{Z}} \) be the Bernoulli product measure on \( \Omega \) and consider \( \sigma : \Omega \to \Omega \) the shift map. Note that the system \((\sigma, \bar{\mu})\) is ergodic. We say that the triple \((\Omega, \sigma, \bar{\mu})\) is the base dynamics determined by \( \mu \).

It is important to point out that the base dynamics \((\Omega, \sigma, \bar{\mu})\) does not depend on the \( \kappa \)-tuple \((A_1, \ldots, A_\kappa) \in (SL_2(\mathbb{R}))^\kappa \) but only on the values \( \mu_i = \mu\{A_i\} \).

Consider the map \( A : \Omega \to SL_2(\mathbb{R}) \) given by

\[ A(\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots) := A_{\omega_0}. \]

Notice that for each sequence \( \omega \in \Omega \), \( A(\omega) \) only depends on the 0-th coordinate of the sequence \( \omega \). Such maps are known in the literature as locally constant linear cocycles. This is an agreed abuse of the term since for the standard topology in \( \Omega \), locally constant observables include a broader class of functions. Since the base dynamics is fixed, \((A_1, \ldots, A_\kappa)\) determines the Lyapunov exponent \( L(A) \) and for that reason some times we write \( L(\mu) = L(A_1, \ldots, A_\kappa) \) to emphasize this dependence. This definition of Lyapunov exponent of \( A \) agrees with the one given in the introduction for the distribution law \( \mu \), so that \( L(\mu) = L(A) = L(A_1, \ldots, A_\kappa) \).

Uniformly hyperbolic cocycles. The measure \( \mu \), or equivalently, the locally constant cocycle \( A : \Omega \to SL_2(\mathbb{R}) \) is said to be uniformly hyperbolic if there exist \( C > 0 \) and \( \gamma > 0 \) such that for every \( n \geq 1 \) and \( \omega \in \Omega \),

\[ \|A^n(\omega)\| \geq Ce^{\gamma n}. \]  

(2)

It is known [5] that this is equivalent to the existence of two \( A \)-invariant continuous sections \( F^u, F^s : \Omega \to \mathbb{P}^1 \) such that for every \( \omega \) \( F^u(\omega) \oplus F^s(\omega) = \mathbb{R}^2 \) and there exist \( C > 0 \) and \( \gamma > 0 \) such that

\[ \|A^n(\omega)|_{F^s(\omega)}\| \leq Ce^{-\gamma n} \quad \text{and} \quad \|A^{-n}(\omega)|_{F^u(\omega)}\| \leq Ce^{-\gamma n}. \]

Forward and backward stationary measures Consider the transition kernels \( Q_+ : \mathbb{P}^1 \to \mathcal{P}(\mathbb{P}^1) \) and \( Q_- : \mathbb{P}^1 \to \mathcal{P}(\mathbb{P}^1) \) defined, respectively, by

\[ Q_+(\hat{v}) := \sum_{i=1}^{\kappa} \mu_i \delta_{A_i \hat{v}} \quad \text{and} \quad Q_- (\hat{v}) := \sum_{i=1}^{\kappa} \mu_i \delta_{A_{i-1}^{-1} \hat{v}}. \]

Definition 1. A measure \( \eta \in \mathcal{P}(\mathbb{P}^1) \) is called forward, resp. backward, stationary for \( \mu \) if \( Q_+^* \eta = \eta \), resp. \( Q_-^* \eta = \eta \), i.e., if \( \eta \) is stationary for \( Q_+ \), resp. for \( Q_- \).

Notice that a backward stationary measure for \( \mu \) is a forward stationary measure for the reverse measure \( \mu^{-1} := \sum_{i=1}^{\kappa} \mu_i \delta_{A_{i-1}} \).
3.2. Irreducible cocycles. Throughout this section, unless otherwise explicitly said, we assume that the probability measure \( \mu = \sum \mu_i \delta_{A_i} \) has positive Lyapunov exponent and is quasi-irreducible.

It follows that the Markov operator \( Q_+ : C^0(\mathbb{P}^1) \to C^0(\mathbb{P}^1) \) defined by

\[
(Q_+ \varphi)(\hat{v}) := \sum_{i=1}^{\kappa} \mu_i \varphi(A_i \hat{v}),
\]

preserves the space of \( \theta \)-Hölder continuous functions \( C^\theta(\mathbb{P}^1) \), for some \( \theta > 0 \), and \( Q_+|_{C^\theta(\mathbb{P}^1)} : C^\theta(\mathbb{P}^1) \to C^\theta(\mathbb{P}^1) \) is a quasi-compact and simple operator, i.e., it has a simple largest eigenvalue, namely 1 associated to the constant functions, and all other elements in the spectrum have absolute value strictly less than 1.

**Proposition 3.1.** There exist a unique forward stationary measure \( \eta^+ \) and a unique backward stationary measure \( \eta^- \) for \( \mu \).

*Proof.* See [21, Proposition 4.2]. \( \square \)

The operator \( Q_+ \) contracts the Hölder seminorm.

**Proposition 3.2.** There exist positive constants \( 0 < \theta < 1, C \) and \( c \) such that

\[
v^\theta(Q^+_n \varphi) \leq C e^{-cn} v^\theta(\varphi) \quad \forall n \in \mathbb{N}.
\]

for every \( \varphi \in C^\theta(\mathbb{P}^1) \).

*Proof.* See [21, Propositions 4.1 and 4.2]. \( \square \)

Another consequence of the quasi-compactness is that the locally constant linear cocycle \( A : \Omega \to SL_2(\mathbb{R}) \), associated with \( \mu \) and defined on the product space \( \Omega = \{1, \ldots, \kappa\}^\mathbb{Z} \), satisfies uniform large deviation estimates of exponential type in a neighborhood of \( A \).

**Proposition 3.3.** There exist constants \( \delta > 0, C > 0, \tau > 0 \) and \( \varepsilon_0 > 0 \) such that for every \( \varepsilon \in (0, \varepsilon_0) \), for all locally constant \( B : \Omega \to SL_2(\mathbb{R}) \) with \( \|B - A\|_\infty < \delta \), every \( \hat{v} \in \mathbb{P}^1 \) and \( n \in \mathbb{N} \),

\[
\tilde{\mu} \left( \left\{ \omega \in \Omega : \frac{1}{n} \log \|B^n(\omega) v\| - L(B) \geq \varepsilon \right\} \right) \leq C e^{-\tau \varepsilon^2 n},
\]

and

\[
\tilde{\mu} \left( \left\{ \omega \in \Omega : \frac{1}{n} \log \|B^n\| - L(B) \geq \varepsilon \right\} \right) \leq C e^{-\tau \varepsilon^2 n}.
\]

*Proof.* See Theorem 4.1 and its proof in [21]. \( \square \)

Let \( 0 < \theta < 1 \) be the constant in Proposition 3.2.

**Proposition 3.4.** There exist constants \( C > 0 \) and \( c > 0 \) such that for any interval \( I \subset \mathbb{P}^1 \), every \( \hat{v} \in \mathbb{P}^1 \) and \( n \geq 1 \) we have,

\[
\eta^\pm(I/2) - C e^{-cn} \sum_{|I|^\theta} \leq \tilde{\mu} \left( [A^\pm n(\cdot) \hat{v} \in I] \right) \leq \eta^\pm(2I) + C e^{-cn} \sum_{|I|^\theta}.
\]
Proof. Set $\varepsilon = |I|/4$ and consider piece-wise linear functions $f^\pm_\varepsilon \in C^0(\mathbb{P}^1)$ such that $v_\theta(f^\pm_\varepsilon) = \varepsilon^{-\theta}$, $0 \leq f^-_\varepsilon \leq \chi_I \leq f^+_\varepsilon \leq 1$, $f^-_\varepsilon = 1$ on $I/2$ and $f^+_\varepsilon = 0$ out of $2I$. By Proposition 3.2 we have that

$$Q_n^+(f^\pm_\varepsilon) - \int f^\pm_\varepsilon d\eta^+ \leq C e^{-cn} v_\theta(f^-_\varepsilon) = C e^{-cn} \varepsilon^{-\theta}.$$ 

Therefore,

$$\tilde{\mu} \left( \left[ A^n(\cdot) \hat{v} \in I \right] \right) = Q_n^+(\chi_I) \geq Q_n^+(f^-_\varepsilon) \geq \int f^-_\varepsilon d\eta^+ - C e^{-cn} \varepsilon^{-\theta} \geq \eta^+(I/2) - 4^\theta C e^{-cn} |I|^{-\theta},$$

and

$$\tilde{\mu} \left( \left[ A^n(\cdot) \hat{v} \in I \right] \right) = Q_n^+(\chi_I) \leq Q_n^+(f^+_\varepsilon) \leq \int f^+_\varepsilon d\eta^+ + C e^{-cn} \varepsilon^{-\theta} \leq \eta^+(2I) + 4^\theta C e^{-cn} |I|^{-\theta}.$$ 

The argument for $A^{-n}$ is analogous. \qed

3.3. Entropy and dimensions.

Entropies. For a finitely supported measure $\mu \in \mathcal{P}(SL_2(\mathbb{R}))$ the Shannon’s entropy defined by

$$H(\mu) := - \sum_{g \in \text{supp}(\mu)} \mu(\{g\})\log \mu(\{g\}) = - \sum_{i=1}^K \mu_i \log \mu_i.$$ 

A measurement of how far the semigroup generated by the supp$(\mu)$ is from being free is given by

$$h_{WR}(\mu) := \lim_{n \to \infty} \frac{1}{n} H(\mu^n) = \inf_{n \in \mathbb{N}} \frac{1}{n} H(\mu^n),$$

which is usually called the random walk entropy of $\mu$. It holds that $h_{RW}(\mu) \leq H(\mu)$ and the equality is equivalent to the semigroup generated by supp$(\mu)$ being free. This is the typical case.

The Furstenberg’s entropy, also known as Boundary entropy, is defined by

$$h_F(\eta) := \int \int \log \frac{dg \eta}{d\eta}(v)\ d\eta(v)\ d\mu(g).$$

We always have that,

$$h_F(\eta) \leq h_{RW}(\mu) \leq H(\mu).$$

See [3, Theorem 2.31] for details.
Dimension. Let $\eta$ be a probability measure on $\mathbb{R}$ (or $\mathbb{P}^1$). For any $t \in \mathbb{R}$, the limits
\[
\overline{\dim}(\eta, t) = \limsup_{\delta \to 0} \frac{\log \eta([t-\delta, t+\delta])}{\log \delta},
\]
\[
\underline{\dim}(\eta, t) = \liminf_{\delta \to 0} \frac{\log \eta([t-\delta, t+\delta])}{\log \delta},
\]
are called, respectively, the upper local dimension and lower local dimension of $\eta$ at the point $t \in \mathbb{R}$. We say that $\eta$ is exact dimensional if there exists a real number $\alpha \geq 0$ such that $\overline{\dim}(\eta, t) = \underline{\dim}(\eta, t) = \alpha$, for $\eta$-a.e. $t \in \mathbb{R}$. In this case, the number $\alpha$ is the dimension of the probability measure $\eta$ and is denoted just by $\dim \eta$.

As mentioned in the introduction the stationary measures of an irreducible cocycle with positive Lyapunov exponent are always exact dimensional.

3.3.1. Entropy deviations. Consider $p_n : \Omega \to \mathbb{R}$, $p_n(\omega) := \prod_{j=0}^{n-1} p_{\omega_j}$, the function $\varphi : \Omega \to \mathbb{R}, \varphi(\omega) := -\log p_{\omega_0}$, and notice that
\[
\int \varphi \, d\tilde{\mu} = H(\mu) \quad \text{and} \quad (S_n \varphi)(\omega) := \sum_{j=0}^{n-1} \varphi(\sigma^j \omega) = -\log p_n(\omega).
\]

Proposition 3.5. Assuming $L(\mu) > 0$ let $h := \max_{1 \leq j \leq \kappa} -\log p_j > 0$. For every $n \in \mathbb{N}$ and $\beta > 0$,
\[
\tilde{\mu}\left( \left\{ \omega \in \Omega : \left| \frac{1}{n} \log p_n(\omega) + H(\mu) \right| > \beta \right\} \right) \leq 2 \exp\left( -n \frac{2 \beta^2}{h^2} \right).
\]

Proof. The large deviation set in the statement is
\[
\Delta_n := \{ \omega \in \Omega : |(S_n \varphi)(\omega) - \mathbb{E}(S_n \varphi)| > n \beta \}
\]
and by Hoeffding’s inequality [22, Theorem 2]
\[
\tilde{\mu}(\Delta_n) \leq 2 \exp\left( -2n^2 \frac{\beta^2}{n h^2} \right) = 2 \exp\left( -n \frac{2 \beta^2}{h^2} \right).
\]

4. Schrödinger Cocycles

In this section we present some background in the theory of Schrödinger cocycles. The advantage in dealing with this family is the intrinsic relation with the spectral theory of (discrete) Schrödinger operators which allow us, among other things, to analyze the behaviour of the Lyapunov exponent in terms of properties of the spectrum of these operators.
4.1. Schrödinger operators and cocycles. Consider the base dynamics \((X, T, \xi)\), where \(T : X \to X\) is a homeomorphism on the compact metric space \(X\) and \(\xi\) is a probability measure on \(X\) such that the system \((T, \xi)\) is ergodic. Fix a continuous function \(\phi : X \to \mathbb{R}\).

For each \(x \in X\), the (discrete) Schrödinger operator at \(x\) is the self-adjoint bounded linear operator \(H_x : l^2(\mathbb{Z}) \to l^2(\mathbb{Z})\) defined, for \(u = (u_n)_{n \in \mathbb{Z}} \in l^2(\mathbb{Z})\) by

\[
(H_x u)_n := -u_{n+1} - u_{n-1} + \phi(T^n x) u_n
\]

or in short notation

\[
H_x u := -\Delta u + \phi_x u,
\]

where \(\Delta\) is the Laplace operator and \(\phi_x\) is the multiplication by \((\phi(T^n x))_{n \in \mathbb{Z}}\).

It is convenient to express the operator \(H_x\) as a matrix in the canonical basis \((e_i)_{i \in \mathbb{Z}}\) of \(l^2(\mathbb{Z})\), where \((e_i)_n = \delta_{i,n}\).

\[
H_x = \begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots & \\
\ldots & \phi(T^{-1} x) & -1 & 0 & \ldots & 0 & 0 & \ldots \\
\ldots & -1 & \phi(x) & -1 & \ldots & 0 & 0 & \ldots \\
\ldots & 0 & -1 & \phi(T x) & \ldots & 0 & 0 & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\ldots & 0 & 0 & 0 & \ldots & \phi(T^{n-2} x) & -1 & \ldots \\
\ldots & 0 & 0 & 0 & \ldots & -1 & \phi(T^{n-1} x) & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\end{pmatrix}
\]

A matrix with this structure where all entries outside the three main diagonals vanish is usually called tridiagonal matrix.

Assume that there exists a sequence \(u = (u_n)_{n \in \mathbb{Z}}\), not necessarily in \(l^2(\mathbb{Z})\), which satisfies the eigenvalue equation for some \(E \in \mathbb{R}\), i.e.,

\[
H_x u = E u. \tag{3}
\]

Using the definition of \(H_x\), Eq. (3) gives us a second order recurrence equation which can be written in matrix form as

\[
\begin{pmatrix}
\phi(T^{n-1} x) - E & -1 \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
u_{n-1} \\
u_{n-2}
\end{pmatrix}
= 
\begin{pmatrix}
u_n \\
u_{n-1}
\end{pmatrix}.
\]

This implies that

\[
\begin{pmatrix}
\phi(T^{n-1} x) - E & -1 \\
1 & 0
\end{pmatrix}
\cdots
\begin{pmatrix}
\phi(x) - 1 \\
1 & 0
\end{pmatrix}
\begin{pmatrix}
u_0 \\
u_{-1}
\end{pmatrix}
= 
\begin{pmatrix}
u_n \\
u_{n-1}
\end{pmatrix}. \tag{4}
\]

Hence, if we define the family of cocycles \(A_E : X \to SL_2(\mathbb{R})\)

\[
A_E(x) := \begin{pmatrix}
\phi(x) - E & -1 \\
1 & 0
\end{pmatrix},
\]

\(^2 l^2(\mathbb{Z})\) denotes the set of square-summable sequences \((u_n)_{n \in \mathbb{Z}}\).
then equation (4) can be rewritten as

\[ A^n_E(x) \begin{pmatrix} u_0 \\ u_{-1} \end{pmatrix} = \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix}. \]

In other words, any (formal) eigenvector \( u = (u_n) \) of the Schrödinger operator \( H_x \) associated with an eigenvalue \( E \) is completely determined by the orbit of the cocycle \( A_E \) starting at \( (u_0, u_{-1}) \in \mathbb{R}^2 \). This is one of the first indications of the close relationship between the action of the cocycle \( A_E \) and the properties of the spectrum of \( H_x \).

The cocycles \( A_E : X \to SL_2(\mathbb{R}) \) are called Schrödinger cocycles with potential \( \phi : X \to \mathbb{R} \), generated by the dynamical system \((X, T, \xi)\).

4.2. Integrated density of states and Thouless formula. For each \( n \in \mathbb{N} \) and for each \( x \in X \), \( H^n_x \in M_n(\mathbb{R}) \) denotes the truncated Schrödinger operator defined by

\[
H^n_x = \begin{pmatrix}
\phi(x) & -1 & 0 & \ldots & 0 & 0 \\
-1 & \phi(Tx) & -1 & \ldots & 0 & 0 \\
0 & -1 & \phi(T^2x) & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & \phi(T^{n-2}x) & -1 \\
0 & 0 & 0 & \ldots & -1 & \phi(T^{n-1}x)
\end{pmatrix}.
\]

For any interval \( I \subset \mathbb{R} \) denote by \( |\text{Spec}(H^n_x) \cap I| \) the number of eigenvalues of \( H^n_x \) in \( I \) counted with multiplicity. With this notation, set for each \( x \in X \)

\[ N_{n,x}(t) := \frac{1}{n} |\text{Spec}(H^n_x) \cap (-\infty, t]|. \]

So, by definition \( N_{n,x}(t) \) is a distribution function of a probability measure supported in the spectrum of \( H^n_x \).

It is known [23, Subsections 3.2 and 3.3] that for each \( t \in \mathbb{R} \) the limit

\[ N(t) = \lim_{n \to \infty} N_{n,x}(t), \]

exists and by ergodicity of the base dynamics \((X, T, \xi)\) is constant for \( \xi\)-a.e. \( x \in X \). The function \( N : \mathbb{R} \to [0, \infty) \) is called the integrated density of states.

The following equation, known as the Thouless formula, relates the Lyapunov exponent of a Schrödinger cocycle with the integrated density of states.

\[ L(A_E) = \int_{-\infty}^{\infty} \log |E - t| \, dN(t), \quad \forall E \in \mathbb{R}. \] (5)

See [23, Theorem 3.16]. Integrating by parts the Riemann-Stieltjes integral on the right-hand side of equation (5), we see that this equation expresses \( L(A_E) \) as the Hilbert transform of \( N(t) \). This fact implies, by the work of Goldstein and Schlag, see [24, Lemma 10.3], that the Lyapunov exponent and the integrated density of states must share all ‘sufficiently nice’ modulus of continuity. These nice moduli of continuity include the Hölder and weak-Hölder regularities. In particular we have:

**Proposition 4.1.** \( N(E) \) is not \( \beta \)-Hölder if and only if \( E \mapsto L(A_E) \) is not \( \beta \)-Hölder.
4.3. Temple’s lemma. The Thouless formula allows us to shift the analysis of the regularity from the Lyapunov exponent to the integrated density of states, and more specifically to the counting of eigenvalues of the truncated Schrödinger operators $H_n^n$. An important tool is the next linear algebra fact, known as Temple’s lemma, which allows us to count eigenvalues by counting instead orthonormal almost eigenvectors.

**Lemma 4.2.** (Temple’s lemma) Let $(V, \langle \cdot, \cdot \rangle)$ be a finite dimensional Hilbert space and let $H : V \to V$ be a self-adjoint linear operator on $V$. Given $\delta > 0$ and $\lambda_0 \in \mathbb{R}$, assume that there exists an orthonormal set $\{u_1, \ldots, u_k\} \subset V$ such that
1. $\langle Hu_i, u_j \rangle = \langle Hu_i, Hu_j \rangle = 0$ if $i \neq j$,
2. $\|Hu_i - \lambda_0 u_i\| \leq \delta$ for every $i$.

Then $|\text{Spec}(H) \cap (\lambda_0 - \delta, \lambda_0 + \delta)| \geq k$.

*Proof.* See [10, Lemma A.3.2]. \( \square \)

We will say that $u \in V \setminus \{0\}$ is a $\delta$-almost eigenvector associated with an almost eigenvalue $\lambda_0$ if condition 2 above is satisfied.

5. Embedding Cocycles Into Schrödinger Families

Let $\mu$, as in the previous section, be a probability measure on $SL_2(\mathbb{R})$ supported in $\{A_1, \ldots, A_\kappa\}$.

We use the notation $S(t) \in SL_2(\mathbb{R})$ to denote the Schrödinger matrix

$$S(t) = \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}$$

The following lemma is the ground basis of the entire section. The fact that we can decompose any given $SL_2(\mathbb{R})$ matrix as a product of four Schrödinger matrices provides a way to embed our random cocycle in a Schrödinger cocycle over a Markov shift.

**Lemma 5.1.** For every $B \in SL_2(\mathbb{R})$, there exists real numbers $t_0, t_1, t_2$ and $t_3$ such that $B = S(t_3) S(t_2) S(t_1) S(t_0)$.

*Proof.* Consider first the map $\mathbb{R}^3 \ni (t_1, t_2, t_3) \mapsto S(t_3) S(t_2) S(t_1) \in SL_2(\mathbb{R})$. A direct calculation shows that the range of this map is the set $SL_2(\mathbb{R}) \setminus M$ where

$$M := \left\{ \begin{pmatrix} a & \lambda \\ -\lambda^{-1} & 0 \end{pmatrix} : \lambda \neq 0, 1, \text{ and } a \in \mathbb{R} \right\}.$$ 

So, the range of the map $\mathbb{R}^3 \ni (t_1, t_2, t_3) \mapsto S(0) S(t_3) S(t_2) S(t_1) \in SL_2(\mathbb{R})$ is the set $SL_2(\mathbb{R}) \setminus S(0) M$ where

$$S(0) M = \left\{ \begin{pmatrix} \lambda^{-1} & 0 \\ a & \lambda \end{pmatrix} : \lambda \neq 0, 1 \text{ and } a \in \mathbb{R} \right\}.$$ 

Another simple calculation shows that if

$$(t_1, t_2, t_3, t_4) = (1, 1 - \lambda^{-1}, -\lambda, \lambda^{-2} - \lambda^{-1} - a\lambda^{-1})$$
then

$$S(t_3) S(t_2) S(t_1) = \begin{pmatrix} \lambda^{-1} & 0 \\ a & \lambda \end{pmatrix}.$$

Hence every matrix in $SL_2(\mathbb{R})$ is a product of four Schrödinger matrices. \( \square \)
5.1. Construction of the embedding. For each $i = 1, \ldots, \kappa$, by Lemma 5.1, there exists $t^i = (t^i_0, \ldots, t^i_3) \in \mathbb{R}^4$ such that

$$A_i = S(t^i_3) S(t^i_2) S(t^i_1) S(t^i_0).$$

(6)

Consider the set $\Lambda = \{1, \ldots, \kappa\} \times \{0, 1, 2, 3\}$. We define the following transition kernel $K : \Lambda \to \mathcal{P}(\Lambda)$, for each element $(i, j) \in \Lambda$,

$$K_{(i,j)} := \begin{cases} \delta_{(i,j+1)} & \text{if } j \in \{0, 1, 2\} \\ \sum_{k=1}^\kappa \mu_k \delta_{(k,0)} & \text{if } j = 3, \end{cases}$$

where $\mu_k = \mu(A_k)$, for any $k \in \{1, \ldots, \kappa\}$ and $\delta_{(k,l)}$ denotes the Dirac measure supported in $(k, l)$. Note that the measure

$$\nu = \frac{1}{4} \sum_{j=0}^3 \sum_{i=1}^\kappa \mu_i \delta_{(i,j)}.$$

defines a $K$-stationary measure on $\Lambda$. Let $\tilde{\nu}$ be the Kolmogorov extension of $(K, \nu)$ on the product space $\Sigma = \Lambda^\mathbb{Z}$. This defines the base dynamics $(\Sigma, \sigma, \tilde{\nu})$, where $\sigma : \Sigma \to \Sigma$ is the shift map and supp $\tilde{\nu}$ is the set of $K$-admissible sequences.

5.2. Conjugating the embedded and original cocycle. Consider the real function $\phi : \Sigma \to \mathbb{R}$ defined by

$$\phi(\xi) := t^i_{j_0}, \quad \text{where } \xi = ((i_n, j_n))_n,$$

where the numbers $t^i_{j_0}$ were defined in (6). We can express the family of Schrödinger cocycles, $A_E : \Sigma \to SL_2(\mathbb{R})$, with potential $\phi$, generated by the Markov shift $(\Sigma, \sigma, \tilde{\nu})$, by

$$A_E(\xi) = S(\phi(\xi) - E),$$

for every $E \in \mathbb{R}$ and $\xi \in \Sigma$. It is important to notice that iterating the cocycle $A_0$ four times we recover the locally constant cocycle $A : \Omega \to SL_2(\mathbb{R})$. More precisely, for each element $\xi = ((i_n, j_n))_n \in \Sigma$, with $j_0 = 0$, consider the sequence $\omega = (i_{4n})_n \in \Omega$. By (6) we have that

$$A^4_0(\xi) = A_0(\sigma^3(\xi)) A_0(\sigma^2(\xi)) A_0(\sigma(\xi)) A_0(\xi) = S(t^i_{j_3}) S(t^i_{j_2}) S(t^i_{j_1}) S(t^i_{j_0}) = A_0(\omega).$$

In this case, we say that $A_0 : \Sigma \to SL_2(\mathbb{R})$ is the embedding of the cocycle $A : \Omega \to SL_2(\mathbb{R})$ into the Schrödinger family $\{A_E : \Sigma \to SL_2(\mathbb{R})\}_{E \in \mathbb{R}}$ over $(\Sigma, \sigma, \tilde{\nu})$.

For each $j \in \{0, 1, 2, 3\}$, set $\Sigma_j := \{(i_n, j_n)_n \in \Sigma : j_0 = j\}$. Note that

$$\Sigma = \bigcup_{j=0}^3 \Sigma_j.$$
is a partition of the set $\Sigma$ and for each $j \in \{0, 1, 2, 3\}$, $\sigma(\Sigma_j) = \Sigma_{j+1 \mod 4}$. In particular, for every $j = 0, 1, 2, 3$, $\Sigma_j$ is $\sigma^4$-invariant. Denote by $\pi : \Sigma \to \Omega$ the natural projection mapping $\Sigma \ni (i_n, f_n) \mapsto (i_{4n})_n \in \Omega$.

Using the notation above we see that $(\Omega, \sigma, \tilde{\mu})$ is a factor of $(\Sigma, \sigma, \tilde{\nu})$ in the following sense.

**Lemma 5.2.** The map $\pi : \Sigma \to \Omega$ is surjective, $\pi_* \tilde{\nu} = \tilde{\mu}$ and $\sigma \circ \pi = \pi \circ \sigma$. Moreover, for each $j \in \{0, 1, 2, 3\}$, $\pi|_{\Sigma_j}$ conjugates $(\Sigma_j, \sigma^4, \tilde{\nu})$ to $(\Omega, \sigma, \tilde{\mu})$, where $4\tilde{\nu}$ is the normalization of $\tilde{\nu}$ on $\Sigma_j$.

For the linear cocycle we have:

**Lemma 5.3.** For every $j = 0, 1, 2, 3$ the linear cocycle

$$\Sigma_j \times \mathbb{R}^2 \ni (\zeta, v) \mapsto (\sigma^4(\zeta), A^4_0(\zeta)v) \in \Sigma_j \times \mathbb{R}^2$$

is conjugated to the linear cocycle

$$\Omega \times \mathbb{R}^2 \ni (\omega, v) \mapsto (\sigma(\omega), A(\omega)v) \in \Omega \times \mathbb{R}^2.$$ 

In particular, taking $j = 0$ we have that $F^4_{A_0} : \Sigma_0 \times \mathbb{R}^2 \to \Sigma_0 \times \mathbb{R}^2$ is conjugated to $F_A : \Omega \times \mathbb{R}^2 \to \Omega \times \mathbb{R}^2$. The same considerations hold for the projectivized cocycles.

As consequence of the previous lemmas we have

**Lemma 5.4.** $L(\mu) = L(A) = 4 L(A_0)$.

Using the conjugation in Lemma 5.2 we build the one parameter family of cocycles $A_{(E)} : \Omega \to SL_2(\mathbb{R})$,

$$A_{(E)}(\omega) := A^4_E(\zeta),$$

where $\zeta = (\pi|_{\Sigma_0})^{-1}(\omega)$. The cocycles of this family are locally constant and determined by the probability measures $\mu_E$ on $SL_2(\mathbb{R})$ defined by

$$\mu_E = \sum_{i=1}^{\kappa} \mu_i \delta_{A_{(E)}(i)},$$

where $\tilde{i}$ is any sequence $\omega \in \Omega$ such that $\omega_0 = i$, for every $i = 1, \ldots, \kappa$. This family is the smooth curve of measures through $\mu$ whose existence is claimed in Theorem A. It depends analytically on $E$ in the sense that the function $E \mapsto \int \varphi d\mu_E = \sum_{i=1}^{\kappa} \mu_i \varphi(A_{(E)}(\tilde{i}))$ is analytic for every analytic function $\varphi(A)$ on $SL_2(\mathbb{R})$. In particular the curve $E \mapsto \mu_E$ is continuous with respect to the weak* topology.

**Corollary 5.5.** For every $E \in \mathbb{R}$, $L(\mu_E) = L(A_{(E)}) = 4 L(A_E)$. 
6. Oscillations of the IDS

Consider the family of Schrödinger cocycles \( A_E : \Sigma \to SL_2(\mathbb{R}) \) with potential \( \phi : \Sigma \to \mathbb{R} \) over the basis dynamics \((\Sigma, \sigma, \tilde{v})\), as in the Sect. 5. The purpose of this section is to get a lower bound on the oscillations of the finite scale IDS \( N_{n, \zeta} \) in terms of counting certain configurations along the orbit of \( \zeta \), referred to as \( \delta \)-matchings.

Let \( \{e_1, e_2\} \) be the canonical basis of \( \mathbb{R}^2 \). Given \( \delta > 0 \) and \( k \in \mathbb{N} \), we say that \( \zeta \in \Sigma \) has a \( \delta \)-matching of size \( k \) at \( E \), or a \( (\delta, k, E) \)-matching, if

\[
A^k_E(\zeta) \hat{e}_1 = \hat{e}_2 \quad \text{and} \quad \tau_k(\zeta, E) := \max_{0 \leq j \leq k-1} \left| A^j_E(\zeta) e_1 \right| < \delta.
\]

For each \( \zeta \in \Sigma \) and \( k \in \mathbb{N} \), we consider the truncated Schrödinger operator \( H^k_\zeta : \mathbb{R}^k \to \mathbb{R}^k \) defined in Sect. 4.2, which can be described, for \( u \in \mathbb{R}^k \), by

\[
H^k_\zeta u := \left( (H^k_\zeta u)_0, \ldots, (H^k_\zeta u)_{k-1} \right),
\]

where

\[
(H^k_\zeta u)_j := \begin{cases} 
-u_1 + \phi(\zeta) u_0, & \text{if } j = 0 \\
-u_{j+1} - u_{j-1} + \phi(\sigma^j(\zeta)) u_j, & \text{if } j \neq 0, k - 1 \\
-u_{k-2} + \phi(\sigma^{k-1}(\zeta)) u_{k-1}, & \text{if } j = k - 1.
\end{cases}
\]

Let \( E \in \mathbb{R} \) and \((v_0, v_{-1}) \in \mathbb{R}^2 \). Define the sequence \((v_j)_{j \in \mathbb{Z}}\) by the following equation

\[
A_E(\sigma^j(\zeta)) \begin{pmatrix} v_j \\ v_{j-1} \end{pmatrix} = \begin{pmatrix} v_{j+1} \\ v_j \end{pmatrix},
\]

which is equivalent to say that for every \( j \in \mathbb{Z} \),

\[
-v_{j-1} - v_{j+1} + \phi(\sigma^j(\zeta)) v_j = E v_j. \tag{7}
\]

Let \( e_0, \ldots, e_{k-1} \) be the canonical basis of \( \mathbb{R}^k \).

**Lemma 6.1.** Given \((v_j)_{j \in \mathbb{Z}}\) solution of (7), the vector \( v^* = (v_0, \ldots, v_{k-1}) \in \mathbb{R}^k \) satisfies

\[
H^k_\zeta v^* - Ev^* = v_{-1}e_0 + v_k e_{k-1}.
\]

Moreover, \( A^k_E(\zeta) \hat{e}_1 = \hat{e}_2 \) if and only if there exists a solution \((v_j)_{j \in \mathbb{Z}}\) of (7) such that \( v^* \) is an eigenvector of \( H^k_\zeta \) with the eigenvalue \( E \).

**Proof.** The first statement follows from (7). For the second part observe that \( A^k_E(\zeta) \hat{e}_1 = \hat{e}_2 \) if and only if there is a solution of (7) such that \( v_{-1} = v_k = 0 \). \( \square \)
Consider a large integer $N = m (k+2)$ and split the interval $[0, N-1]$ into $m$ disjoint slots of length $k$, namely $S_j := [j (k+2), j (k+2) + k - 1]$ for $j = 0, 1, \ldots, m - 1$. The integers $j (k+2) - 1$ and $j (k+2) + k + 1$ are referred to as the boundary points of the slot $S_j$. Notice that $\bigcup_{j=0}^{m-1} S_j$ has $m k$ elements which exclude the boundary points of the slots. We say that $\zeta \in \Sigma$ has a $(\delta, k, E)$-matching in the slot $S_j$ if $\sigma^j (k+2) (\zeta)$ has a $(\delta, k, E)$-matching. Next lemma says that when the sequence $\zeta$ has a $(\delta, k, E)$-matching in the slot $S_j$ we can construct a $\delta$-almost eigenvector for $H_\xi^N$ which is supported in that slot $S_j$. Moreover, because consecutive slots share no boundary points in common, if $\zeta$ admits several $(\delta, k)$-matchings in different slots then the corresponding $\delta$-almost eigenvectors are pairwise orthogonal.

Lemma 6.2. Given $\zeta \in \Sigma$ and $j_0 \in \mathbb{N}$ such that the sequence $\sigma^{j_0 (k+2)} (\zeta)$ has a $(2^{-1/2} \delta, k, E)$-matching consider the vector $v^* = (v_0, \ldots, v_{k-1}) \in \mathbb{R}^k$ with components determined by
\[
\begin{pmatrix} v_j \\ v_{j-1} \end{pmatrix} = A^i_E (\sigma^{j_0 (k+2)} \zeta) \begin{pmatrix} v_0 \\ 0 \end{pmatrix}
\]
where $v_0$ is fixed so that $\max_{0 \leq j \leq k-1} |v_j| = 1$. Then the vector $v_{j_0, k} (\zeta) \in \mathbb{R}^N$, with all coordinates zero except those in the slot $S_{j_0}$ which coincide with the respective coordinates of $v^*$, satisfies
\[
\| H_\xi^N v_{j_0, k} (\zeta) - E v_{j_0, k} (\zeta) \| < \delta.
\]
In other words, $v_{j_0, k} (\zeta)$ is an $\delta$-almost eigenvector of $H_\xi^N$ in the sense of Lemma 4.2.

Proof. For the sake of simplicity let $j_0 = 0$ so that $\zeta = \sigma^{j_0 (k+2)} (\zeta)$ is the sequence with a $(\delta/\sqrt{2}, k, E)$-matching. By definition of $v_{j_0, k} (\zeta)$ and Lemma 6.1 we have that
\[
(H_\xi^N - E) v_{j_0, k} (\zeta) = - v_{k-1} e_k.
\]
Therefore,
\[
\| H_\xi^N v_{j_0, k} (\zeta) - E v_{j_0, k} (\zeta) \| \leq |v_{k-1}| \leq |v_0| \| A^i_E (\zeta) e_1 \| \leq \sqrt{2} \tau_k (\xi, E) < \delta
\]
because $v_{k-1}$ is one of the components of $v_0 A^k_E e_1$ and
\[
\frac{|v_0|}{\sqrt{2}} \max_{0 \leq j \leq k-1} \| A^i_E (\zeta) e_1 \| \leq v_0 \max_{0 \leq j \leq k-1} |\langle A^i_E (\zeta) e_1, e_1 \rangle| = \max_{0 \leq j \leq k-1} |v_j| = 1.
\]
\[\square\]

From the point of view of Mathematical Physics, a $\delta$-matching determines a $\delta$-almost eigenvector of the Schrödinger operator.

Dynamically, these configurations correspond to stable-unstable matchings in the following sense: let $k = k_1 + k_2$ be some partition of $k$ such that both factors in the decomposition $A^k_E (\zeta) = A^{k_2}_E (\sigma^{k_1} (\zeta)) A^{k_1}_E (\zeta)$ are very hyperbolic with nearly horizontal unstable direction and almost vertical stable one. If $k_1, k_2$ are large then $A^{k_1}_E (\zeta) e_1$ is a good approximation of the Oseledets unstable direction $E^u (\sigma^{k_1} \zeta)$ at the point $\sigma^{k_1} (\zeta)$, while $A^{-k_2}_E (\zeta) e_2$ is a good approximation of the stable direction $E^s (\sigma^{k_1} (\zeta))$ at the same
point. The condition $A^k(\xi) \hat{e}_1 = \hat{e}_2$ is equivalent to the matching $A_k^1(\xi) \hat{e}_1 = A_k^2(\xi) \hat{e}_2$ between these two approximate stable and unstable directions at the middle point. This nearly stable-unstable matching also explains why $\tau_k(\xi, E)$ should be very small.

The oscillation of the non-decreasing function $N$ and its finite scale analogue $\hat{N}_{N, \xi}$ on some interval $I = [\alpha, \beta]$ are denoted by

$$\Delta_I N := N(\beta) - N(\alpha), \text{ resp. } \Delta_I \hat{N}_{N, \xi} := \hat{N}_{N, \xi}(\beta) - \hat{N}_{N, \xi}(\alpha).$$

Denote by $\Sigma(\delta, k, I)$ the subset of $\Sigma$ formed by $(\delta, k, E)$-matching sequences $\xi \in \Sigma$ with $E \in I$.

**Lemma 6.3.** For any interval $I \subset \mathbb{R}$ and $\xi \in \Sigma$,

$$\Delta_{I_\delta} \hat{N}_{N, \xi} \geq \frac{1}{N} \sum_{j=0}^{m-1} \chi_{\Sigma(\delta, k, I)}(\sigma^j(\xi))$$

where $I_\delta := I + [-\delta, \delta]$ is the $\delta$-neighborhood of $I$.

**Proof.** Let $m \in \mathbb{N}$ and set

$$Z_{m, k}(\xi) := \left\{ 0 \leq j \leq m - 1 : \sigma^j(\xi) \in \Sigma(\delta, k, I) \right\}.$$

The set of vectors $\{v_{j, k}(\xi) \in \mathbb{R}^N : j \in Z_{m, k}(\xi)\}$ is orthonormal and by Lemma 6.2 these are $\delta$-almost eigenvectors. By Lemma 4.2 there is the same amount of eigenvalues of $H_{\xi}^N$ in $I_\delta$ (counted with multiplicity). Whence,

$$\sum_{j=0}^{m-1} \chi_{\Sigma(\delta, k, I)}(\sigma^j(\xi)) = |Z_{m, k}(\xi)| \leq |\text{Spec}(H_{\xi}^N) \cap I_\delta| = N \Delta_{I_\delta} \hat{N}_{N, \xi}.$$

Applying Birkhoff’s ergodic theorem sending $m \to \infty$ in the previous lemma we have the following corollary.

**Corollary 6.4.** For any interval $I \subseteq \mathbb{R}$,

$$\Delta_{I_\delta} N \geq \frac{1}{k+2} \tilde{\nu}(\Sigma(\delta, k, I)).$$

### 7. Variation with Respect to the Energy

This is the main technical section of the work.
7.1. Trace property. The main purpose of this subsection is to prove that if \( R_0 = A_0^{4n_0}(\zeta_0) \) is elliptic then as we move the parameter \( E \) the rotation angle of \( R_E \) varies with non-zero speed around \( E = 0 \). This will be a consequence of the following proposition, which is a general fact about Schrödinger matrices. Recall that

\[
S(t) = \begin{pmatrix} t & -1 \\ 1 & 0 \end{pmatrix}
\]

denotes a Schrödinger type matrix. For a vector \( x = (x_1, \ldots, x_n) \in \mathbb{R}^n \), write

\[
S^n(x - E) := S(x_n - E) \ldots S(x_1 - E).
\]

Lemma 7.1. If \( E \in \mathbb{C} \setminus \mathbb{R} \) then the matrix \( S^n(x - E) \) is hyperbolic.

Proof. See [25, Lemma 2.4]. For the sake of completeness we provide a proof of this fact. By induction the entries in the main diagonal of \( S^n(x) \) are polynomials in the variables \( x_1, \ldots, x_n \) of degrees \( n \) and \( n - 2 \), respectively, whose monomials have degrees with same parity as \( n \), while the entries on the second diagonal are polynomials of degrees \( n - 1 \), whose monomials have degrees with same parity as \( n - 1 \). It follows that for all \( x \in \mathbb{R}^n \),

\[
\text{tr}(S^n(-x)) = (-1)^n \text{tr}(S^n(x)).
\]

In particular \( \text{tr}(S^n(x - E)) = \pm \text{tr}(S^n(E - x)) \) and we only need to consider the case \( \text{Im} E < 0 \). In this case the open set \( U := \{ z \in \mathbb{C} : \text{Im} z > 0 \} \) determines an open cone in \( \mathbb{P}(\mathbb{C}^2) \equiv \mathbb{C} \cup \{ \infty \} \). The projective action of the matrices \( S^n(x - E) \) with \( x \in \mathbb{R}^n \) and \( E \in U \) sends \( U \) inside of \( U \). In fact, if \( n = 1 \) and \( z \in U \setminus \{ 0 \} \) (possibly \( z = \infty \)) then \( -z^{-1} \in \overline{U} \) and since \( \text{Im} E < 0 \),

\[
\text{Im}(S(x_1 - E) \cdot z) = \text{Im} \left( x_1 - E - \frac{1}{z} \right) \geq -\text{Im}(E) > 0,
\]

for every \( x_1 \in \mathbb{R} \). Otherwise if \( z = 0 \), then \( S(x_1 - E) \cdot z = \infty \) and the statement follows iterating once again. The existence of this invariant cone implies that \( S^n(x - E) \) is hyperbolic. Similarly, if \( \text{Im} E > 0 \) we consider the open set \( U^- := \{ z \in \mathbb{C} : \text{Im} z < 0 \} \) and prove that, under the projective action, \( S^n(x - E) \) sends \( U^- \) inside of \( U^- \). \( \square \)

Proposition 7.2. For any \( n \in \mathbb{N} \), if \( |\text{tr}(S^n(x))| < 2 \), then

\[
\frac{d}{dE} \left. \text{tr} \left( S^n(x - E) \right) \right|_{E=0} \neq 0.
\]

Proof. Define the analytic function \( \psi : \mathbb{C} \rightarrow \mathbb{C} \) given by

\[
\psi(E) := \text{tr} \left( S^n(x - E) \right).
\]

Observe that \( \psi \) is real in the sense that \( \psi(E) \in \mathbb{R} \) for every \( E \in \mathbb{R} \). By assumption \( |\psi(0)| < 2 \) and so there exists a radius \( r_0 > 0 \) such that for every \( E \) in the disk centered in \( 0 \) and radius \( r_0 \), \( \mathbb{D}_{r_0}(0) \), we have that \( |\psi(E)| < 2 \).

Assume by contradiction that \( \psi'(0) = 0 \). By analyticity of \( \psi \), we can write

\[
\psi(E) = \psi(0) + E^k \Psi(E),
\]
in a neighborhood of 0, where \( k \geq 2 \) and \( \Psi(0) \neq 0 \). In particular, there exists \( E^* \in \mathbb{D}_{\delta_0}(0) \setminus \mathbb{R} \) such that \( \psi(E^*) \in \mathbb{R} \). As a consequence, we conclude that if \( \lambda \) and \( 1/\lambda \) are the eigenvalues of \( S^n(x - E^*) \), then \( \lambda + \lambda^{-1} \in \mathbb{R} \). But, that can only happen if either \( |\lambda| = 1 \) or else \( |\lambda| \neq 1 \) and \( \lambda \) itself is real. The former can not happen since by Lemma 7.1 the matrix \( S^n(x - E^*) \) is hyperbolic.

The latter implies that

\[
|\text{tr} \left( S^n(x - E^*) \right)| = \left| \lambda + \frac{1}{\lambda} \right| > 2.
\]

This contradicts the fact that \( |\psi(E^*)| < 2 \) and proves the result. \( \square \)

**Lemma 7.3.** If \( A_0^m(\zeta_0) \) is an elliptic element for some \( \zeta_0 \in \Sigma \), then

\[
\frac{d}{dE} \text{tr}(A_0^m(\zeta_0)) \bigg|_{E=0} \neq 0.
\]

**Proof.** Direct consequence of Proposition 7.2. \( \square \)

**Proposition 7.4.** Given a Schrödinger cocycle \( A_E : X \to SL_2(\mathbb{R}) \) with continuous potential \( \phi : X \to \mathbb{R} \) and generated by the dynamical system \( (X, T, \xi) \), for all \( n \in \mathbb{N} \), \( \rho \in (-2, 2) \) and \( x \in X \), the polynomial \( f_\rho : \mathbb{R} \to \mathbb{R} \), \( f_\rho(E) := \text{tr}(A_E^n(x)) - \rho \), has \( n \) distinct real roots.

**Proof.** This polynomial can not have a real root \( E_0 \) with multiplicity \( \geq 2 \) because this would imply that \( f_\rho(E_0) = f_\rho'(E_0) = 0 \), contradicting the conclusion of Proposition 7.3. To see that it can not have complex non-real roots, assume that there exists \( E_0 \in \mathbb{C} \setminus \mathbb{R} \) such that \( f_\rho(E_0) = 0 \). By Lemma 7.1, the matrix \( S^n(x - E_0) \) is hyperbolic. Denoting by \( \lambda \) and \( \lambda^{-1} \) the eigenvalues of \( S^n(x - E_0) \) we have

\[
\rho = \text{tr}(S^n(x - E_0)) = \lambda + \lambda^{-1}
\]

which implies that \( |\lambda| = 1 \). Therefore the matrix \( S^n(x - E_0) \) can not be hyperbolic. This contradiction proves that \( f_\rho(E) \) can not have complex non-real roots. \( \square \)

**Corollary 7.5.** In the previous context, \( f : \mathbb{R} \to \mathbb{R} \), \( f(E) := \text{tr}(A_E^n(x)) \), is a Morse function, with \( f(E) \geq 2 \) at local maxima, and \( f(E) \leq -2 \) at local minima.

**Proof.** Since \( f \) has \( n \) different real roots, \( f' \) has \( n - 1 \) different real roots and for any pair \( a < b \) of roots of \( f \), there exists a unique \( c \in (a, b) \) root of \( f' \). Moreover, by Proposition 7.4, if \( f''(c) < 0 \), then \( f(c) \geq 2 \) and similarly, \( f''(c) > 0 \) implies \( f(c) \leq -2 \). \( \square \)

### 7.2. Density of tangencies.

In this subsection we prove that cocycles with heteroclinic tangencies are dense outside the class of uniformly hyperbolic cocycles.

Let \( A_0^{4\delta_0}(\zeta_0) \) be an elliptic matrix and \( \delta_0 > 0 \) be such that \( R_E = A_E^{4\delta_0}(\zeta_0) \) is elliptic for every \( |E| \leq \delta_0 \).

**Lemma 7.6.** There exist \( c > 0 \) such that for every \( m \geq 1 \), every \( E \in [-\delta_0, \delta_0] \) and every \( \hat{v} \in \mathbb{P}^1 \) we have

\[
\left| \frac{d}{dE} R_E^m \hat{v} \right| \geq m c.
\]
Proof. Take $E_0 \in [-\delta_0, \delta_0]$ and an inner product in $\mathbb{R}^2$ for which $R_{E_0}$ is a rotation. Then by Proposition B.1 and Lemma B.3

$$\frac{1}{m} \left| \frac{d}{dE} R_{E_0}^m \hat{v} \right|_{E = E_0} = \frac{1}{m} \sum_{j=1}^{m} R_{E_0} v_{j-1} \wedge \dot{R}_{E_0} v_{j-1}$$

is bounded away from 0. Notice that by compactness of $[-\delta_0, \delta_0]$, all norms associated with inner products that turn the matrices $R_{E}$ into rotations are uniformly equivalent. □

Proposition 7.7. Given $\zeta \in \Sigma$, $E \in [-\delta_0, \delta_0]$ and $\ell \in \mathbb{N}$, if $A_{E_0}^\ell(\zeta)$ is parabolic then there exist $E$ arbitrary close to $E_0$ such that $A_E^\ell(\zeta)$ is irrational elliptic.

Proof. Consider the function $f : \mathbb{R} \to \mathbb{R}$, $f(E) := \text{tr}(A_E^\ell(\zeta))$. Assume $A_{E_0}^\ell(\zeta)$ parabolic, i.e., $f(E_0) = \pm 2$. When $f'(E_0) \neq 0$, all matrices $A_E^\ell(\zeta)$ are elliptic in a 1-sided neighborhood of $E_0$. On the other hand, if $f'(E_0) = 0$ by Corollary 7.5 all matrices $A_E^\ell(\zeta)$ are elliptic in a 2-sided neighborhood of $E_0$. □

Proposition 7.8. Assume $L(\mu_{E_0}) > 0$ and $A_{E_0}^{4\ell}$ is not uniformly hyperbolic, then there exist $E$ arbitrary close to $E_0$ at which $\mu_E$ admits heteroclinic tangencies.

Proof. By [13, Theorem 4.1], either $\mu_{E_0}$ has an heteroclinic tangency, or else the semigroup generated by $\text{supp} \mu_{E_0}$ contains a parabolic or an elliptic matrix. Since $L(\mu_{E_0}) > 0$, supp($\mu_{E_0}$) admits hyperbolic matrices $A_{E_0}$ and $B_{E_0}$. By Proposition 7.7 we can assume that $C_{E_0} := A_{E_0}^{4\ell}(\zeta)$ is an irrational elliptic rotation, which implies that the distance $d(C_{E_0}^mu(B_{E_0}), s(A_{E_0}))$ gets arbitrary small for some large $m$. On the other hand, the curves $E \mapsto u(B_E), s(A_E)$ are smooth, see Proposition B.4, while by Lemma 7.6 the projective curve $E \mapsto C_{E_0}^mu(B_E)$ has large speed when $m$ is large. Hence the equation $C_{E_0}^mu(B_E) = s(A_E)$ has infinitely many solutions with $E$ arbitrary close to $E_0$. See Figure 2. □
7.3. Projective random walk distribution. In this subsection we establish some estimates on the distribution of the projective random walk, needed to prove Proposition 7.11.

**Proposition 7.9.** Assume that $L(\mu) > 0$ and $\mu$ is irreducible. There exist $C > 0$ and $t \in (0, 1)$ such that

$$\sup_{\widehat{x} \in \mathbb{P}^1} \int_{\mathbb{P}^1} \frac{1}{d(\widehat{x}, \widehat{y})^t} \, d\eta(\widehat{y}) \leq C.$$ 

In particular, $\eta$ is $t$-Hölder, i.e., for every $\widehat{x} \in \mathbb{P}^1$ and $r > 0$

$$\eta(B(\widehat{x}, r)) \leq Cr^t.$$

**Proof.** Since $\mu$ is an irreducible measure on $SL_2(\mathbb{R})$ and $L(\mu) > 0$ we have that $\mu$ is strongly irreducible (see [26, Theorem 6.1]). Thus the proof follows from [27] or [28, Theorem 13.1].

The first item of the next proposition corresponds to (13.8) of Proposition 13.3 in [28].

**Proposition 7.10.** Assume that $L(\mu E_0) > 0$ and $\mu E_0$ irreducible. Given $\beta > 0$, there exist constants $C, c_1, c_2 > 0$ and $k_0 \in \mathbb{N}$ such that for every $\ell, n \in \mathbb{N}$, $n \geq k_0 \ell$ and directions $\hat{v}, \hat{w} \in \mathbb{P}^1$, the sets

(1) $\{\omega \in \Omega : \exists E, |E - E_0| \leq e^{-c_1 n}, A^n_{(E)}(\omega) \hat{v} \in B(\hat{w}, e^{-\beta}\ell)\}$;

(2) $\{\omega \in \Omega : \exists E, |E - E_0| \leq e^{-c_1 n}, A^{-n}_{(E)}(\omega) \hat{w} \in B(\hat{v}, e^{-\beta}\ell)\}$;

(3) $\{(\omega, \hat{w}) \in \Omega^2 : \exists E, |E - E_0| \leq e^{-c_1 n}, d(A^n_{(E)}(\omega) \hat{v}, A^{-n}_{(E)}(\hat{w}) \hat{w}) \leq e^{-\beta}\ell\}$

have probability $\leq Ce^{-c_2\ell}$.

**Proof.** By Lemma B.2, there exist constants $C^*, c^*_1, c^*_2 > 0$ such that

$$d\left(A^n_{(E_0)}(\omega) \hat{v}^+, A^n_{(E)}(\omega) \hat{v}^+\right) \leq C^* e^{-c^*_1 n},$$

(8)

for every $E$ with $|E - E_0| \leq e^{-c^*_2 n}$. By Proposition 3.4 and Proposition 7.9, we have that

$$\tilde{\mu}\left(\exists E, |E - E_0| \leq e^{-c_1 n}, A^n_{(E)}(\cdot) \hat{v} \in B(\hat{w}, e^{-\beta}\ell)\right)$$

$$\lesssim \eta^*\left(B(\hat{w}, e^{-\beta}\ell)\right) + C \frac{e^{-cn}}{e^{-c_1\beta\ell}}$$

$$\lesssim e^{-t\beta\ell} + C \frac{e^{-cn}}{e^{-c_1\beta\ell}},$$

where $n \geq k_0 \ell$ with $k_0 > \beta(t + \theta)/c$ and $c_2 := t\beta$. The argument to estimate the probability in 2) is entirely analogous, making use of $\eta^*$ instead of $\eta^+$.

We now study the probability of the set in 3). We extend the Markov operators $Q_{\pm}$ to the product space $\mathbb{P}^1 \times \mathbb{P}^1$ defining a new operator $Q : C^\theta(\mathbb{P}^1 \times \mathbb{P}^1) \to C^\theta(\mathbb{P}^1 \times \mathbb{P}^1)$ by

$$(Q\varphi)(\hat{x}, \hat{y}) := \sum_{i,j=1}^k \mu_i \mu_j \varphi(A_{i,E_0} \hat{x}, A_{j,E_0}^{-1} \hat{y}).$$
This operator is also a quasi-compact operator and there exists constants \( c, C > 0 \) such that for every observable \( \varphi \in C^\theta(\mathbb{P}^1 \times \mathbb{P}^1) \) we have

\[
\nu_\theta(Q^n \varphi) \lesssim e^{-cn} \nu_\theta(\varphi).
\]

For each \( r > 0 \), let \( \Delta_r := \{(\hat{x}, \hat{y}) \in \mathbb{P}^1 \times \mathbb{P}^1 : d(\hat{x}, \hat{y}) \leq r \} \) and \( \rho_r : [0, +\infty[ \to [0, 1] \) be a piece-wise linear function supported in \( [0, 3r] \) such that \( \rho_r(t) = 1 \) for \( t \in [0, 2r] \). Define the \( \theta \)-Hölder observable \( \psi_r(\hat{x}, \hat{y}) := \rho_r(d(\hat{x}, \hat{y})) \), with \( \nu_\theta(\psi_r) = (2r)^{-\theta} \) and \( \chi_{\Delta_r} \leq \psi_r \). Writing \( r = 2e^{-c_1n} + e^{-\beta \ell} \), we can use Markov's inequality and Proposition 7.9 to conclude that

\[
\hat{\mu} \times \hat{\mu} \left( \left\{ E, |E - E_0| \leq e^{-c_1n}, d \left( A_{\theta(E)}^n(\cdot) \hat{v}^+, A_{\theta(E)}^{-n}(\cdot) \hat{v}^- \right) \leq e^{-\beta \ell} \right\} \right) \\
\leq \hat{\mu} \times \hat{\mu} \left( \left\{ (\omega, \tilde{\omega}) \in \Omega^2 : d \left( A_{\theta(E_0)}^n(\omega) \hat{v}^+, A_{\theta(E_0)}^{-n}(\tilde{\omega}) \hat{v}^- \right) \leq 2e^{-c_1n} + e^{-\beta \ell} \right\} \right) \\
= Q^n(\chi_{\Delta_r})(\hat{v}^+, \hat{v}^-) \leq Q^n(\psi_r)(\hat{v}^+, \hat{v}^-) \\
\leq \left| Q^n(\psi_r) - \int_{\mathbb{P}^1 \times \mathbb{P}^1} \psi_r \, d(\eta_{E_0}^+ \times \eta_{E_0}^-) \right| + \int_{\mathbb{P}^1 \times \mathbb{P}^1} \psi_r \, d(\eta_{E_0}^+ \times \eta_{E_0}^-) \\
\lesssim e^{-cn} \nu_\theta(\psi_r) + (\eta_{E_0}^+ \times \eta_{E_0}^-)(\Delta_{3r}) \\
\lesssim \frac{e^{-cn}}{e^{-\beta \ell}} + 3^t (2e^{-c_1n} + e^{-\beta \ell}) \int_{\mathbb{P}^1 \times \mathbb{P}^1} \frac{1}{d(\hat{x}, \hat{y})^t} \, d(\eta_{E_0}^+ \times \eta_{E_0}^-)(\hat{x}, \hat{y}) \\
\lesssim \frac{e^{-cn}}{e^{-\beta \ell}} + 3^t (2e^{-c_1n} + e^{-\beta \ell}) \sup_{\hat{y} \in \mathbb{P}^1} \int_{\mathbb{P}^1 \times \mathbb{P}^1} \frac{1}{d(\hat{x}, \hat{y})^t} \, d(\eta_{E_0}^+)(\hat{x}) \\
\lesssim \frac{e^{-cn}}{e^{-\beta \ell}} + (2e^{-c_1n} + e^{-\beta \ell})^t \lesssim e^{-c_2t}.
\]

In the two last inequalities we have used Proposition 7.9 and that we can increase \( k_0 \) so that \( k_0 > \frac{\beta(1+\theta)}{c} \) and decrease \( c_2 \) so that \( c_2 \leq \beta \theta t \). This completes the proof of the Proposition.

7.4. Variation of the 'hyperbolic' elements. In this subsection we establish one of the core proposition for the proof of the Theorem A, providing plenty of good hyperbolic words. We will be using the notation introduced in the Sect. 5.2.

From now on we also use the following notation: given \( A \in SL_2(\mathbb{R}) \) with \( ||A|| > 1 \), denote by \( \hat{v}_1(A), \hat{v}_2(A), \hat{v}_1^*(A) \) and \( \hat{v}_2^*(A) \) the unique projective points such that taking unit vectors \( v_i \in \hat{v}_i(A) \) and \( v_j^* \in \hat{v}_j(A) \), with \( i, j = 1, 2 \), \( \{v_1, v_2\} \) and \( \{v_1^*, v_2^*\} \) are singular basis of \( A \) characterized by the relations \( A v_1 = ||A|| v_1^* \) and \( A v_2 = ||A||^{-1} v_2^* \).

Take \( \delta_1 = \delta_1(E_0) > 0 \) as in the Proposition 3.3, in the sense that the large deviations hold uniformly for all cocycles \( A_{\theta(E)} \) with \( |E - E_0| \leq \delta_1 \) and also so that

\[
\lambda := \min_{|E - E_0| \leq \delta_1} L(\mu_{\theta(E)}) > 0.
\]

Proposition 7.11. Assume \( L(\mu_{E_0}) > 0 \) and \( \mu_{E_0} \) irreducible. Given \( \beta > 0 \) there exist constants \( \tau > 0 \) and \( N_0 \in \mathbb{N} \) such that for every \( n \geq N_0 \) and every \( \hat{v}, \hat{w} \in \mathbb{P}^1 \), the set \( G_n(\hat{v}, \hat{w}, \beta, \tau, E_0) \) of all \( \omega \in \Omega \) satisfying for all \( |E - E_0| \leq e^{-\tau n^{1/4}} \):

\[
(1) \left\| A_{\theta(E)}^n(\omega) v \right\| \gtrsim e^{(\lambda - \beta)n} \text{ and } \left\| A_{\theta(E)}^{-n}(\sigma^n \omega) w \right\| \gtrsim e^{(\lambda - \beta)n};
\]
(2) $A^n_{(E)}(\omega)$ is hyperbolic and $\lambda(A^n_{(E)}(\omega)) \geq e^{(\lambda - \beta)n}$;
(3) $d(\hat{v}^*_1(A^n_{(E)}(\omega)), \hat{v}_2(A^n_{(E)}(\omega))) \geq e^{-\beta n^{1/8}}$.

has measure $\tilde{\mu}(\mathcal{G}_n(\hat{v}, \hat{w}, \beta, \tau, E_0)) > 1 - \beta$.

**Proof.** Split $n$ into blocks of size $n_0 \approx n^{1/4}$. For the sake of simplicity we assume that $n = m n_0$ with $n_0 = n^{1/4}$ and $E_0 = 0$. Consider the set $B_{n_0}$ of $\omega \in \Omega$ such that there exists $0 \leq j \leq m - 1$ with

$$\left\| A^{n_0}_{(E)}(\sigma j n_0 \omega) \right\| < e^{(\lambda - \frac{\beta}{\tau_0})n_0} \lor \left\| A^{n_0}_{(E)}(\sigma^j n_0 \omega) \right\| > e^{(\lambda + \frac{\beta}{\tau_0})n_0}.$$  

Write $B_{n_0}^* := B_{n_0} \cup B_{2n_0}$, where $B_{2n_0}$ is similarly defined. By large deviations, Proposition 3.3, there exists a constant $\tau_1 > 0$ such that for all large enough $n$, $\tilde{\mu}(B_{n_0}^*) \leq 2n^{3/4} e^{-\tau_1 n^{1/4}}$. By finite scale continuity, there exists $\tau > 0$ such that for all $|E| \leq e^{-\tau n^{1/4}} = e^{-\tau n_0}$ and $\omega \notin B_{n_0}^*$,

$$e^{(\lambda - \frac{\beta}{\tau_0})n_0} \leq \left\| A^{n_0}_{(E)}(\sigma^j n_0 \omega) \right\| \leq e^{(\lambda + \frac{\beta}{\tau_0})n_0} \quad \forall 0 \leq j < m$$

and

$$e^{2(\lambda - \frac{\beta}{\tau_0})n_0} \leq \left\| A^{2n_0}_{(E)}(\sigma^j n_0 \omega) \right\| \leq e^{2(\lambda + \frac{\beta}{\tau_0})n_0} \quad \forall 0 \leq j < m - 1.$$

Consider $(\hat{v}, \hat{w}) \in \mathbb{P}^1 \times \mathbb{P}^1$. We will apply Lemma A.9 with the data

- $A_j = A^{n_0}_{(E)}(\sigma^j n_0 \omega)$, $j = 0, \ldots, m - 1$;
- $\hat{v} = \hat{v}$ and $\hat{w} = \hat{w}$;
- $t = \beta n_0^{1/2}$, $\gamma := \frac{4}{5} \beta n_0$ and $\lambda := (\lambda - \frac{\beta}{\tau_0})n_0$.

Notice that if $\omega \notin B_{n_0}^*$ the assumptions (a)-(c) of Lemma A.9 are automatically satisfied.

Consider $C, c_1, c_2 > 0$ and $k_0$ given by Proposition 7.10 applied with $n = n_0$ and $\ell = n_0^{1/2}$. Denote by $C_{n_0}(\hat{v}, \hat{w})$ the set of sequences $\omega \in \Omega$ such that for every $|E| \leq e^{-\tau n_0}$ ($\tau > c_1$).

(a) $\min \left\{ d\left( A^{n_0}_{(E)}(\sigma^{(m-1)n_0} \omega), \hat{v}, \hat{w} \right), d\left( \hat{v}, A^{n_0}_{(E)}(\sigma^{(m)n_0} \omega) \hat{w} \right) \right\} \geq e^{-\beta n_0^{1/2}}$;

(b) $\min \left\{ d\left( A^{n_0}_{(E)}(\omega), \hat{v}, \hat{w} \right), d\left( \hat{v}, A^{n_0}_{(E)}(\sigma^{n_0} \omega) \hat{w} \right) \right\} \geq e^{-\beta n_0^{1/2}}$;

(c) $d\left( A^{n_0}_{(E)}(\sigma^{(m-1)n_0} \omega), \hat{v}, A^{n_0}_{(E)}(\sigma^{(m)n_0} \omega) \hat{w} \right) \geq e^{-\beta n_0^{1/2}}$.

If $n_0/\ell = n_0^{1/2} \geq k_0$, then by Proposition 7.10 the set $C_{n_0}^* := C_{n_0}(\hat{v}, \hat{w}) \setminus B_{n_0}^*$ satisfies

$$\tilde{\mu}(\Omega \setminus C_{n_0}^*) \leq Ce^{-c_2 n_0^{1/2}} = Ce^{-c_2 n_0^{1/8}}.$$

If $\omega \in C_{n_0}^*$ the above conditions (a–c) ensure that the hypothesis (d–f) of Lemma A.9 holds. Therefore items 1, 2, and 3 are direct consequence of Lemma A.9. This concludes the proof of the proposition. \qed

**Proposition 7.12.** If the cocycle $A_{(E_0)}$ is not irreducible with $L(A_{(E_0)}) > 0$ then there exists $\delta > 0$ such that for all $0 < |E - E_0| \leq \delta$, the cocycle $A_{(E)}$ is irreducible.
Proof. The cocycle $A(E_0)$ has either one or two invariant lines, i.e., invariant under all matrices in the support of $\mu_E$. Since $A(E_0)$ is not uniformly hyperbolic there exist hyperbolic periodic points $\omega_1$ and $\omega_2$, with periods $n_1$ and $n_2$, respectively, such that $\hat{u}(A_{(E_0)}^{n_1}(\omega_1)) = \hat{s}(A_{(E_0)}^{n_2}(\omega_2))$ or $\hat{s}(A_{(E_0)}^{n_1}(\omega_1)) = \hat{u}(A_{(E_0)}^{n_2}(\omega_2))$, for otherwise a simple argument implies that the reducible cocycle is uniformly hyperbolic, see inequality (2). By Proposition B.4 the directions $\hat{u}(A_{(E)}^{n_i}(\omega_i))$ and $\hat{s}(A_{(E)}^{n_i}(\omega_i))$ move in opposite directions with the parameter $E$. Hence in any case, for $E \neq E_0$ close enough to $E_0$, together the two matrices $A_{(E)}^{n_i}(\omega_i)$, $i = 1, 2$, have four distinct invariant directions. This implies that the cocycle $A(E)$ is irreducible. □

7.5. Variation of the heteroclinic tangencies. In this subsection we establish the core Propositions 7.14 and 7.15 for the proof of Theorem A which allows us drive matchings and typical tangencies from an existing tangency.

Consider a family of cocycles $A_{(E)} : \Omega \to SL_2(\mathbb{R})$ as introduced in Sect. 5.2. This family has a heteroclinic tangency at $E$ if and only if there exist periodic orbits $\omega_0, \omega_1 \in \Omega$ with periods $\ell_0, \ell_1 \geq 1$ such that $A_{(E)}^{\ell_0}((\omega_0))$ and $A_{(E)}^{\ell_1}((\omega_1))$ are hyperbolic matrices, and there exists a heteroclinic orbit $\omega \in W^u_{\text{loc}}(\omega_0) \cap \sigma^{-\ell_0}W^s_{\text{loc}}(\omega_1)$ such that

$$A_{(E)}^{\ell_1}(\omega) \hat{u}(A_{(E)}^{\ell_0}(\omega_0)) = \hat{s}(A_{(E)}^{\ell_1}(\omega_1)).$$

In this case we say that $(B_E, C_E, A_E)$ is a tangency for $A(E)$ where $A_E = A_{(E)}^{\ell_1}(\omega)$, $B_E = A_{(E)}^{\ell_0}(\omega_0)$ and $C_E = A_{(E)}^{\ell_1}(\omega)$ are respectively the target, the source and the transition matrix of this heteroclinic tangency. See Figure 3. The size of the tangency is by definition the size of the full word $B_E C_E A_E$ determined by the tangency.

Before entering in the main technical results of this section we state a version of Lemma B.3 suitable for our purposes. We identify the derivative of projective curves such as $E \mapsto A^n_{(E)}(\omega) \hat{v}$ with its scalar value.

Lemma 7.13. There exists $c_* > 0$ such that for all $n \geq 2$, $\omega \in \Omega$, $E \in \mathbb{R}$ and $\hat{v} \in \mathbb{P}^1$, we have

$$\frac{d}{dE} A_{(E)}^{-n}(\omega) \hat{v} < -c_* < 0 < c_* < \frac{d}{dE} A_{(E)}^{n}(\omega) \hat{v}.$$

Proof. Recall that for each $\omega \in \Omega$, $A_{(E)}^{n}(\omega) = A_{E}^{n}(\xi)$ for some $\xi \in \Sigma$ and that the cocycle $A_{E} : \Sigma \to SL_2(\mathbb{R})$ is a Schrödinger cocycle. Therefore, this lemma is a direct consequence of Lemma B.3. □

Definition 2. Given $\gamma$, $t$, $\rho > 0$, we say that a tangency $(B_{E_0}, C_{E_0}, A_{E_0})$ for a cocycle $A_{(E_0)}$ is $(\gamma, \rho, t)$-controlled if the following conditions are satisfied:

1. $\min\{|\lambda(A_{E_0})|, |\lambda(B_{E_0})|\} \geq e^\gamma$;
2. $\|C_{E_0}\| \leq e^\rho$;
3. $\min\{d(\hat{v}_1(B_{E_0})), d(\hat{v}_2(B_{E_0})), d(\hat{v}_1(B_{E_0})), d(\hat{v}_2(B_{E_0}))\} \geq e^{-t}$.

Proposition 7.14. There exists $c_* > 0$ such that for every $\beta > 0$ and $R > 0$ we can find $y_0$ with the following property: for every $\gamma \geq y_0$, if $(B_{E_0}, C_{E_0}, A_{E_0})$ is a $(\gamma^{1/2}, \gamma^{1/7})$-controlled tangency for $A_{E_0}$, then defining

$$I := [E_0 - 2c_*^{-1}(1 + \beta) R e^{-2\gamma(1-\beta)}, E_0 + 2c_*^{-1}(1 + \beta) R e^{-2\gamma(1-\beta)}],$$

for every pair of smooth curves $\hat{v}^+, \hat{v}^- : I \to \mathbb{P}^1$ satisfying
Fig. 3. Unfolding the heteroclinic tangency. Vertical lines represent $\hat{v}^1$.

A1. $d(\hat{v}^+(E_0), \hat{s}(B_0)) > R^{-1}$ and $d(\hat{v}^-(E_0), \hat{u}(A_0)) > R^{-1}$;
A2. $\frac{d}{dE} \hat{v}^+(E) \geq 0$ and $\frac{d}{dE} \hat{v}^-(E) \leq 0$, for every $E \in I$;

the equation

$$A_E C_E B_E \hat{v}^+(E) = \hat{v}^-(E),$$

has at least one solution $E^* \in I$.

Proof. We assume for the sake of simplicity $E_0 = 0$. First observe that using triangular inequality, condition A1, Proposition A.3 and the given control of the tangency, there exists $K_0 > 0$ such that

$$d(\hat{v}^+(0), \hat{v}_2(B_0)) \geq d(\hat{v}^+(0), \hat{s}(B_0)) - d(\hat{s}(B_0), \hat{v}_2(B_0))$$

$$\geq R^{-1} - \frac{K_0}{d(\hat{v}^+_1(B_0), \hat{v}_2(B_0)) \|B_0\|^2}$$

$$\geq R^{-1} \left(1 - K_0 R e^{-2\gamma(1-\frac{1}{2\gamma^{6/7}})}\right)$$

and

$$d(\hat{v}^-(0), \hat{v}_1^+(A_0)) \geq d(\hat{v}^-(0), \hat{s}(B_0)) - d(\hat{s}(B_0), \hat{v}_2(B_0))$$

$$\geq R^{-1} - \frac{K_0}{d(\hat{v}^+_1(A_0), \hat{v}_2(B_0)) \|A_0\|^2}$$

$$\geq R^{-1} \left(1 - K_0 R e^{-2\gamma(1-\frac{1}{2\gamma^{6/7}})}\right).$$

Now using the previous inequalities jointly with item (b) of Lemma A.2 and the control of the transition matrix,

$$d(C_0 B_0 \hat{v}^+(0), C_0 \hat{u}(B_0)) \leq \|C_0\|^2 d(B_0 \hat{v}^+(0), \hat{u}(B_0))$$

$$\leq \|C_0\|^2 \frac{1}{d(\hat{v}^+(0), \hat{v}_2(B_0)) \|B_0\|^2}$$

$$\leq R \left(1 - K_0 R e^{-2\gamma(1-\frac{1}{2\gamma^{6/7}})}\right)^{-1} e^{-2\gamma(1-\gamma^{-1/2})}$$
and

\[ d(A_0^{-1} \hat{\nu}^-(0), \hat{s}(A_0)) \leq \frac{1}{d(\hat{\nu}^-(0), \hat{s}(A_0))} \leq R \left( 1 - K_0 R e^{-2y(1 - \frac{1}{2\gamma \beta})} \right)^{-1} e^{-2\gamma}. \]

Taking

\[ \gamma_0 := \max \left\{ \beta^{-2}, (2\beta)^{-7/6}, \frac{1}{2(1 - \beta)} \log \left( \frac{K_0(1 + \beta)}{\beta} \right) \right\}, \]

we conclude that for every \( \gamma \geq \gamma_0 \)

\[ d(C_0 B_0 \hat{\nu}^+(0), A_0^{-1} \hat{\nu}^-(0)) \leq 2(1 + \beta) Re^{-2y(1 - \beta)}. \] (9)

Choose appropriate projective coordinates in such way to preserve the natural orientation. Consider the functions \( f_+, f_- : I \rightarrow \mathbb{P}^1 \) given by

\[ f_+(E) = C_E B_E \hat{\nu}^+(E) \quad \text{and} \quad f_-(E) = A_E^{-1} \hat{\nu}^-(E). \]

By condition A2 and Lemma 7.13 we have that there exists \( c_* > 0 \) such that \( f'_+(E) < -c_* < 0 < f'_-(E) \) for every \( E \in I \). Moreover, by inequality (9), \( d(f_+(0), f_-(0)) \leq 2(1 + \beta) Re^{-2y(1 - \beta)}. \) Therefore, there exists \( |E_*| \leq 2(1 + \beta)c_*^{-1} Re^{-2y(1 - \beta)} \) such that \( f_+(E_*) = f_-(E_*) \), i.e.,

\[ A_{E_*} C_{E_*} B_{E_*} \hat{\nu}^+(E_*) = \hat{\nu}^-(E_*). \]

We finish this section showing that if the cocycle \( A_{E_0} \) has a tangency we can perturb the parameter to produce plenty of new tangencies which are typical with respect to the Lyapunov exponent and the Shannon entropy in a finite scale. Recall the notation of Sect. 3.3.

**Proposition 7.15.** Assume the cocycle \( A_{(E_0)} \) has a heteroclinic tangency and is irreducible. Given \( \beta > 0 \), there exist constants \( C_1^*, C_2^*, c_1^*, c_2^* > 0 \), a sequence \( (l_k)_k \subset \mathbb{N} \), \( l_k \rightarrow \infty \), and \( k_0 \in \mathbb{N} \) such that for every \( k \geq k_0 \) we can find a set \( \mathcal{X}_k(\beta) \subset \Omega \) with \( \hat{\mu}(\mathcal{X}_k(\beta)) \geq C_1^* e^{-c_1^*/l_k^{1/3}} \) with the following property: for every \( \omega \in \mathcal{X}(\beta) \) there exists \( E_k = E_k(\omega) \) with \( |E_k - E_0| \leq C_2^* e^{-c_2^*/l_k^{1/3}} \) such that \( A_{(E_k)} \) has a tangency \( (P_{E_k}, T_{E_k}, S_{E_k}) \), of size \( l_k \), satisfying

1. \( (P_{E_k}, T_{E_k}, S_{E_k}) \) is \( (\gamma_k, \gamma_k^{1/2}, \gamma_k^{1/7}) \)-controlled with \( \gamma_k = \frac{1}{2} (\lambda - 3\beta) \);
2. \( p_{l_k}^{(0)}(\omega) := \prod_{j=0}^{l_k-1} p_{\omega_j} \geq e^{-(H(\mu) + \beta)l_k}. \)
Proof. To lighten notations assume $E_0 = 0$. Fix $β > 0$ and let $(B_0, C_0, A_0)$ be a tangency for $A_{(E_0)}$. Take integers $p_k, q_k \geq 1$ such that
\[
\left| \frac{p_k}{q_k} - \frac{\log \lambda(B_0)}{\log \lambda(A_0)} \right| < \frac{1}{q_k^2},
\]
or equivalently
\[
\lambda(B_0)^{q_k} \lambda(A_0)^{-q_k} < \lambda(A_0)^{p_k} < \lambda(B_0)^{q_k} \lambda(A_0)^{-q_k},
\]
and write $λ_k := λ(A_0)^{p_k} \sim λ(B_0)^{q_k}$. Consider for each $k \geq 1$ the new tangency $(B_0^{q_k}, C_0, A_0^{p_k})$ of size $m_k$, also for $A_{(E_0)}$. We claim that this tangency is $(γ, γ^{1/2}, γ^{1/7})$-controlled with $γ := (1 - β) \log λ_k$ and $k$ sufficiently large. Indeed, by inequality (10),
\[
\min[λ(A_0^{p_k}), λ(B_0^{q_k})] \geq e^γ,
\]
for every $k$ sufficiently large. Furthermore, the upper bound for $||C_0||$ and the lower bound for the distances $d(\hat{v}_1^*(A_0^{p_k}), \hat{v}_2(A_0^{p_k}))$ and $d(\hat{v}_1^*(B_0^{q_k}), \hat{v}_2(B_0^{q_k}))$ can be taken independently of $k$ and so the conditions of Definition 2 are automatic satisfied for every $k$ large.

For each $R > 0$, consider the projective intervals $J^s := (\hat{s}(B_0), R^{-1})$ and $J^u := (\hat{u}(A_0), R^{-1})$ as well as
\[
I_k := [E_0 - 2c_1^{-1}(1 + β)Rλ_k^{−2(1-2β)} + 2c_1^{-1}(1 + β)Rλ_k^{−2(1-2β)}].
\]
Denote by $τ_k$ the finite word of size $m_k$ determined by the tangency, i.e., for every $ω ∈ [0; τ_k], A_{(0)}^{m_k}(ω) = A_0^{p_k}C_0B_0^{q_k}$.

Since $A_0$ is strongly irreducible, the forward and backward stationary measures $η^+$ and $η^−$ are non-atomic. Hence, we can choose $R$ sufficiently large so that
\[
η^+_0(6J^s) \leq 1/4 \quad \text{and} \quad η^-_0(6J^u) \leq 1/4.
\]

Given $\hat{v}, \hat{w} ∈ \mathbb{P}^1$, consider the set $G_n := G_n(\hat{v}, \hat{w}, β, τ, 0)$ given by Proposition 7.11. For each $n ≥ 1$, define
\[
G^u_n := \left\{ ω ∈ G_n : \hat{u}(A_{(0)}^n(ω)) \notin 2J^s \right\}
\]
and
\[
G^s_n := \left\{ ω ∈ G_n : \hat{s}(A_{(0)}^n(ω)) \notin 2J^u \right\}.
\]
Notice that by item 3. of Proposition 7.11
\[
\left\{ ω ∈ G_n : A_{(0)}^n(ω) \hat{v} \notin 3J^s \right\} ⊂ G^u_n \quad \text{and} \quad \left\{ ω ∈ G_n : A_{(0)}^{-n}(ω) \hat{w} \notin 3J^s \right\} ⊂ G^s_n.
\]
Thus, by inequality (11), Propositions 3.4 and 7.11 we have that
\[
\hat{μ}(G^u_n) ≥ \hat{μ}(G_n) - \hat{μ}([A_{(0)}^n(\cdot) \hat{v} ∈ 3J^s]) ≥ 1 - β - 2η^+_0(6J^s) ≥ \frac{1}{2} - β,
\]
and similarly,
\[
\hat{μ}(G^s_n) ≥ \hat{μ}(G_n) - \hat{μ}([A_{(0)}^{-n}(\cdot) \hat{w} ∈ 3J^u]) ≥ 1 - β - 2η^-_0(6J^u) ≥ \frac{1}{2} - β.
\]
We define the set $T_k$ of tangencies by

$$T_k := G_{m_k^3}^u \cap [m_k^3; \tau_k] \cap \sigma^{-d_k} \left( G_{m_k^3}^s \right),$$

where $d_k := m_k^3 + m_k$. Take $\omega \in T_k$ and define the functions $\hat{v}^+, \hat{v}^- : I_k \to \mathbb{P}^1$

$$\hat{v}^+(E) := \hat{u}(\mathcal{A}^{m_k^3}_{(E)}(\omega)) \quad \text{and} \quad \hat{v}^-(E) := \hat{s}(\mathcal{A}^{m_k^3}_{(E)}(\sigma^{d_k}\omega)).$$

Notice that by definition of $T_k$,

$$\hat{v}^+(0) \notin 2J^s \supset B(\hat{s}(B_0), R^{-1}) \quad \text{and} \quad \hat{v}^-(0) \notin 2J^u \supset B(\hat{u}(A_0), R^{-1}).$$

Moreover, by the Lemma B.4,

$$\frac{d}{dE} \hat{v}^+(E) \geq 0 \quad \text{and} \quad \frac{d}{dE} \hat{v}^-(E) \leq 0.$$

Thus, we can apply Proposition 7.14, to guarantee that there exists $E_k = E_k(\omega) \in I_k$, satisfying

$$A_{E_k}^{p_k} C_{E_k} B_{E_k}^{q_k} \hat{u}(\mathcal{A}^{m_k^3}_{(E_k)}(\omega)) = \hat{s}(\mathcal{A}^{m_k^3}_{(E_k)}(\sigma^{d_k}\omega)).$$

Set $l_k := 2m_k^3 + m_k$ and consider the set

$$\mathcal{F}_k(\beta) := \left\{ \omega \in \Omega : p_l(\omega) \geq e^{-(H(\mu) + \beta)l_k} \right\}.$$

By Proposition 3.5 with $n = l_k$ and $\varepsilon = \beta$,

$$\tilde{\mu}(\mathcal{F}_k(\beta)) \geq 1 - 2e^{-\frac{1}{16}l_k \beta^2}, \quad (14)$$

where $h$ is a positive constant depending only on $\mu$.

For each $\omega \in T_k \cap \mathcal{F}_k(\beta)$, define

$$P_k := \mathcal{A}_{(E_k)}^{m_k^3}(\omega), \quad T_k := A_{E_k}^{p_k} C_{E_k} B_{E_k}^{q_k} \quad \text{and} \quad S_k := \mathcal{A}_{(E_k)}^{m_k^3}(\sigma^{d_k}\omega),$$

and observe that by Proposition 7.11, $P_k$ and $S_k$ are hyperbolic and if $\gamma_k := (\lambda - 3\beta)^{l_k^2}$, then $(P_k, T_k, S_k)$ is a $(\gamma_k, \gamma_k^{1/2}, \gamma_k^{1/7})$-controlled tangency for the cocycle $A_{(E_k)}$ of size $l_k$. Moreover,

$$\lambda(P_k) \geq e^{(\lambda - \beta)m_k^3} > e^{(\lambda - 2\beta)\frac{l_k^2}{2}} \quad \text{and} \quad \lambda(S_k) \geq e^{(\lambda - \beta)m_k^3} > e^{(\lambda - 3\beta)\frac{l_k^2}{2}}.$$

which proves item 1. Item 2 holds because $\omega \in \mathcal{F}_k(\beta)$.

To finish the proposition notice that by inequalities (12), (13) and (14)

$$\tilde{\mu}(T_k \cap \mathcal{F}_k(\beta)) \geq \tilde{\mu}(T_k) - \tilde{\mu}(\Omega \setminus \mathcal{F}_k(\beta)) \geq (1/2 - \beta)^2 \tilde{\mu}([0; \tau_k]) - 2e^{-\frac{1}{16}l_k \beta^2} \geq (1/2 - \beta)^2 e^{-c m_k}$$

for some constant $c > 0$. Taking $X_k(\beta) := T_k \cap \mathcal{F}_k(\beta)$ completes the proof. \qed
8. Counting Matchings

The purpose of this section is to give a lower bound for the \( \nu \)-measure of the set of sequences for which we have a \((\delta, k, l)\)-matching for some small interval of energies \( l \). Throughout this subsection we assume that the cocycle \( A_0 \) has a heteroclinic tangency and is irreducible. We keep the notation used in the Proposition 7.15.

Recall that for a suitable \( \delta_1 > 0 \) we use the notation

\[
\lambda = \min_{|E| \leq \delta_1} L(\mu_E) > 0.
\]

8.1. Subset of matchings. Take \( \beta > 0 \) and let \( \mathbb{N}' \) be the set of sizes \( l \in \mathbb{N} \) of the heteroclinic tangencies of \( A_{E_l}, (P_{E_l}, T_{E_l}, S_{E_l}) \), given by Proposition 7.15, applied with \( E_0 = 0 \), where \( E_l = E_l(\omega) \) for some \( \omega \in \Omega \) and \( |E_l| \leq C_2 e^{-c_2^1 l^{1/3}} \) is such that \( A_{(E_l)}^l(\omega) = S_{E_l} T_{E_l} P_{E_l} \). Denote by \( \tau_l \in \Omega \) the finite word of length \( l \) associated with the block \( S_{E_l} T_{E_l} P_{E_l} \), for a given size \( l \in \mathbb{N}' \). By item 3 of Proposition 7.15

\[
\mu(\{0; \tau_l\}) = \mathfrak{p}_l(\omega) \geq e^{-(H(\mu)+\beta)l}.
\] (15)

To apply Proposition 7.14 with the tangency \((P_{E_l}, T_{E_l}, S_{E_l})\) we consider the balls \( J^\lambda_l \) and \( J^\mu_l \) in \( \mathbb{P}^1 \) centered respectively in \( \hat{s}(P_{E_l}) \) and \( \hat{u}(S_{E_l}) \) with radius \( R^{-1} > 0 \). Consider the interval

\[
I_l := [E_l - C e^{-l(\lambda-\beta)}, E_l + C e^{-l(\lambda-\beta)}]
\]

provided by this proposition with \( C := 2c_2^{-1}(1+\beta)R \). Choosing \( \gamma := \frac{l}{2}(\lambda - \beta) \) by the said proposition the heteroclinic tangency \((P_{E_l}, T_{E_l}, S_{E_l})\) is \((\gamma, \gamma^{1/2}, \gamma^{1/7})\)-controlled so that the initial assumptions of Proposition 7.14 are automatically satisfied.

Fix \( \tau > 0 \) given by Proposition 7.11 and set \( G_{l,\tau} := G_{l,\tau}(\hat{e}_1, \hat{e}_2, \beta, \tau, 0) \). Define the sets

\[
\Theta^\mu_l := \left\{ \omega \in G_{l,\tau} : A_{E_l}^{(l)}(\omega) \hat{e}_1 \notin 2J^\lambda_l \right\}
\]

and

\[
\Theta^\tau_l := \left\{ \omega \in G_{l,\tau} : A_{E_l}^{(l)}(\omega) \hat{e}_2 \notin 2J^\mu_l \right\}.
\]

Notice that taking \( R \) sufficiently large and applying Proposition 7.10 we have

\[
\mu(\Theta^\mu_l) \geq 1/2 - \beta \quad \text{and} \quad \mu(\Theta^\tau_l) \geq 1/2 - \beta.
\] (16)

Now, we finally define our subset of matchings as

\[
\mathcal{M}_l := \Theta^\mu_l \cap \sigma^{-l^3}([0; \tau_l]) \cap \sigma^{-(2l^3+l)}(\Theta^\tau_l).
\]

Lemma 8.1. For every \( l \in \mathbb{N}' \) and \( \omega \in \mathcal{M}_l \) there exists \( E^*_l \in I_l \) such that

1. \( A_{(E^*_l)}^{(3l^2+l)}(\omega) \hat{e}_1 = \hat{e}_2 \);
2. \( e^{(\lambda-\beta)l^3} \leq \| A_{(E^*_l)}^{(3l^2+l)}(\omega) e_1 \| \leq \| A_{(E^*_l)}^{(3l^2+l)}(\omega) \| \leq e^{(\lambda+\beta)l^3} \);
3. \( e^{(\lambda-\beta)l^3} \leq \| A_{(E^*_l)}^{-(3l^2-l)}(\sigma^{2l^3+l}\omega) e_2 \| \leq \| A_{(E^*_l)}^{-(3l^2-l)}(\sigma^{2l^3+l}\omega) \| \leq e^{(\lambda+\beta)l^3} \).
Moreover, \( \tilde{\mu}(\mathcal{M}_I) \geq (1/2 - \beta)^2 e^{-l(H(\mu)+\beta)}. \)

**Proof.** Fix \( \omega \in \mathcal{M}_I. \) By Proposition 7.14 we conclude that the equation

\[
A_E^{2l^3+l}(\omega) \hat{e}_1 = \hat{e}_2,
\]

has at least a solution \( E^*_I \subset I_l. \) In fact, as explained above the heteroclinic tangency \((P_{E_I}, T_{E_I}, S_{E_I})\) is \((\gamma, \gamma^{1/2}, \gamma^{1/7})\)-controlled so that the initial assumptions of Proposition 7.14 are automatically satisfied. Next consider the curves \( \hat{\nu}^+(E) := A_E^{3}(\omega) \hat{e}_1 \) and \( \hat{\nu}^-(E) := A_E^{3}(\sigma^{2l^3+l}\omega) \hat{e}_2. \) Assumption A1 holds because \( \omega \in \Theta^U \) and \( \sigma^{2l^3+l}\omega \in \Theta^U. \) Assumption A2 holds by Lemma 7.13.

The lower bounds in items 2 and 3 follow from item 1 of Proposition 7.11 and the fact that \( \omega \in \mathcal{G}_{l^3}(\hat{e}_1, \hat{e}_2, \beta, \tau, 0). \) From items 1 and 3 of the said proposition together with conclusion 2) of Proposition A.3 we get the upper bounds in items 2 and 3.

Taking unit vectors \( w_1 \in A_{E_I}^{3}(\omega) \hat{e}_1 \) and \( w_2 \in A_{E_I}^{l}(\sigma^{l^3}\omega) \hat{w}_1, \) by 2 above,

\[
\left\| A_{(E_I)}^{2l^3+l}(\omega) e_1 \right\| = \left\| A_{(E_I)}^{3}(\sigma^{l^3+l}\omega) w_2 \right\| \left\| A_{(E_I)}^{l}(\sigma^{l^3}\omega) w_1 \right\| \left\| A_{(E_I)}^{l}(\omega) e_1 \right\| 
\leq e^{-(\lambda-\beta)t^3} e^{Ct} e^{(\lambda+\beta)t^3} \leq e^{2\beta t^3+Ct} \leq e^{3\beta t^3},
\]

which proves item 4.

To finish, using the inequalities in (15) and (16) we have

\[
\tilde{\mu}(\mathcal{M}_I) = \tilde{\mu}(\Theta^U) \tilde{\mu}(\Theta^U) \tilde{\mu}([0; \tau]) \geq (1/2 - \beta)^2 e^{-l(H(\mu)+\beta)}.
\]

This completes the proof of the lemma. \( \Box \)

Now we can give a lower bound for the set of matchings. Recall the notation of Sect. 6.

**Corollary 8.2.** For all large \( l \in \mathbb{N}, \) if \( n_l := 4(2l^3 + l) \) then

\[
\tilde{\nu} \left( \Sigma(e^{-(\lambda-\beta)n_l}, n_l, I_l) \right) \geq \frac{1}{4} (1/2 - \beta)^2 e^{-l(\mu(\mu)+\beta)}.
\]

**Proof.** Let \( \pi_0 := \pi|_{\Sigma_0} : \Sigma_0 \to \Omega \) be the conjugation given by Lemma 5.2. We claim that \( \pi_0^{-1}(\mathcal{M}_I) \subset \Sigma(e^{-(\lambda-\beta)n_l}, n_l, I_l). \) Indeed, by Lemma 8.1 if \( \pi_0(\zeta) \in \mathcal{M}_I, \) there exist \( E^*_I \subset I_l \) such that

\[
A_{E^*_I}^{n_l}(\zeta) \hat{e}_1 = A_{E^*_I}^{2l^3+l}(\pi_0(\zeta)) \hat{e}_1 = \hat{e}_2
\]

Moreover,

\[
\tau_{n_l}(\zeta, E^*_I) \leq \left\| A_{E^*_I}^{n_l}(\zeta) e_1 \right\| = \left\| A_{E^*_I}^{2l^3+l}(\pi_0(\zeta)) e_1 \right\| \leq e^{3\beta l^3-(\lambda-\beta)l^3}.
\]

This proves that any \( \zeta \) is a \((e^{-(\lambda-\beta)n_l}, n_l, E^*_I)-\)matching for some \( E^*_I \subset I_l. \)
To finish, since $4\bar{\nu}$ is normalization of $\bar{\nu}$ to $\Sigma_0$,

$$
\bar{\nu} \left( \Sigma(e^{-l^3(\lambda-\beta)}, n_l, I_l) \right) \geq \bar{\nu}(\pi_0^{-1}(M_l)) = \frac{1}{4} \bar{\mu}(\mathcal{M}_l) \\
\geq \frac{1}{4} \left(1/2 - \beta\right)^2 e^{-l(H(\mu)+\beta)}.
$$

This completes the proof of the corollary. \(\square\)

9. Proof of the Results

We keep the notations of the previous section.

9.1. Proof of theorem A. We keep the notation of the previous section. Take $\alpha > \frac{H(\mu)}{L(\mu)}$ and choose $\delta > 0$ such that $\lambda := \min_{|E| \leq \delta} L(\mu_E)$ satisfies $\lambda \alpha > H(\mu) > 0$. Then take $0 < \beta < \lambda$ small enough so that $\lambda \alpha - H(\mu) > 2\beta + \alpha \beta$, which implies that

$$
-H(\mu) - \beta + \alpha (\lambda - \beta) > \beta. \tag{18}
$$

By Proposition 4.1, to prove Theorem A it is enough to prove that the integrated density of states $\mathcal{N}$ is not $\alpha$-Hölder continuous. By Corollaries 6.4 and 8.2, writing $\delta_l := e^{-l^3(\lambda-4\beta)}$,

$$
\Delta_{I_l + [-\delta_l, \delta_l]} \mathcal{N} \geq \frac{1}{n_l} \bar{\nu} (\Sigma(\delta_l, n_l, I_l)) \geq \frac{1}{4n_l} (1/2 - \beta)^2 e^{-l(H(\mu)+\beta)}.
$$

Thus by inequality (18),

$$
\frac{\Delta_{I_l + [-\delta_l, \delta_l]} \mathcal{N}}{|I_l + [-\delta_l, \delta_l]|^\alpha} \gtrsim e^{l(-H(\mu) - \beta + \alpha (\lambda - \beta))} \gtrsim e^{\beta l}.
$$

Taking $l \in \mathbb{N}$, $l \to \infty$ we conclude that $\mathcal{N}$ can not be $\alpha$-Hölder continuous. \(\square\)

9.2. Proof of corollary A. By [13, Theorem 4.1], if $\mu$ is not uniformly hyperbolic, then the semigroup generated by $\text{supp} \mu$ must contain a parabolic or elliptic matrix. In either case, by Proposition 7.8 and Proposition 7.12, we can approximate $\mu$ by measures with finite support admitting tangencies which are irreducible. The result follows by continuity of the quotient $E \mapsto \frac{H(\mu)}{L(\mu_E)}$. \(\square\)

9.3. Proof of corollary B. By Johnson’s theorem [29], if $E_0$ is an energy in the almost sure spectrum of the Schrödinger operator, then the associated Schrödinger cocycle $A_{E_0}$ is not uniformly hyperbolic. Therefore, we can again apply Propositions 7.8 and 7.12 to find energies close to $E_0$ such that the cocycle $A_{E_0}$ is irreducible and has heteroclinic tangencies. The result follows by continuity of the quotient $E \mapsto \frac{H(\mu)}{L(\mu_E)}$. \(\square\)
Appendix A: Some Linear Algebra Facts

In this appendix we state and prove a few results about the geometry of the projective action of a matrix $A \in SL_2(\mathbb{R})$ in the Euclidean space $\mathbb{R}^2$. Some of these results are well known. Others like Propositions A.3, A.6 and Lemma A.9 play a key role in logical architecture of our main results. For the reader’s convenience we also state here a version of the Avalanche Principle for $SL_2(\mathbb{R})$ matrices.

If $v = \{v_1, v_2\}$ is a basis of $\mathbb{R}^2$ then the dual basis of $v$ is the unique basis $v^b = \{v_1^b, v_2^b\}$ of $\mathbb{R}^2$ such that $\langle v_i^b, v_j \rangle = \delta_{ij}$, for $i, j = 1, 2$.

As usual let $J := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ denote the $90^\circ$ rotation matrix.

Lemma A.1. For a basis $v = \{v_1, v_2\}$ of $\mathbb{R}^2$, its dual basis is given by $v^b = \{ -\frac{Jv_2}{v_1 \wedge v_2}, \frac{Jv_1}{v_1 \wedge v_2} \}$. In particular, the trace of $A \in SL_2(\mathbb{R})$ is given by

$$\text{tr}(A) = \frac{1}{v_1 \wedge v_2} (Av_1 \wedge v_2 + v_1 \wedge Av_2).$$

Proof. For the first part it is enough to check the following relations, where we make extensive use the relation $\langle x, y \rangle = x \wedge J y$.

$$\begin{align*}
- \frac{Jv_2}{v_1 \wedge v_2}, v_1 & = - \frac{\langle Jv_2, v_1 \rangle}{v_1 \wedge v_2} = - \frac{Jv_2 \wedge Jv_1}{v_1 \wedge v_2} = - \frac{v_2 \wedge v_1}{v_1 \wedge v_2} = 1 \\
- \frac{Jv_2}{v_1 \wedge v_2}, v_2 & = - \frac{\langle Jv_2, v_2 \rangle}{v_1 \wedge v_2} = 0 \\
\frac{Jv_1}{v_1 \wedge v_2}, v_1 & = - \frac{\langle Jv_1, v_1 \rangle}{v_1 \wedge v_2} = 0 \\
\frac{Jv_1}{v_1 \wedge v_2}, v_2 & = \frac{\langle Jv_1, v_2 \rangle}{v_1 \wedge v_2} = \frac{Jv_1 \wedge Jv_2}{v_1 \wedge v_2} = \frac{v_1 \wedge v_2}{v_1 \wedge v_2} = 1
\end{align*}$$

Denoting the dual basis of $v = \{v_1, v_2\}$ by $v^b = \{v_1^b, v_2^b\}$, we have

$$\text{tr}(A) = v_1^b \cdot Av_1 + v_2^b \cdot Av_2.$$
If a
\[
\exists k
\]
\[
\frac{Jv_1}{v_1 \wedge v_2}, A v_1 + \frac{Jv_2}{v_1 \wedge v_2}, A v_2
\]
\[
= -\frac{Jv_2, A v_1}{v_1 \wedge v_2} + \frac{Jv_1, A v_2}{v_1 \wedge v_2} = -\frac{Jv_2 \wedge J A v_1}{v_1 \wedge v_2} + \frac{Jv_1 \wedge J A v_2}{v_1 \wedge v_2}
\]
\[
= -\frac{v_2 \wedge A v_1}{v_1 \wedge v_2} + \frac{v_1 \wedge A v_2}{v_1 \wedge v_2} = \frac{A v_1 \wedge v_2}{v_1 \wedge v_2} + \frac{v_1 \wedge A v_2}{v_1 \wedge v_2}.
\]

\[
\square
\]

From Lemma A.2 until Proposition A.6, we consider the projective distance,
\[
d(\hat{x}, \hat{y}) := \frac{|x \wedge y|}{\|x\| \|y\|} = |\sin \angle(x, y)|
\]
take a matrix \(A \in SL_2(\mathbb{R})\) and let \(\{v_1, v_2\}\) and \(\{v_1^*, v_2^*\}\) be singular orthonormal basis of \(A\) characterized by the relations \(A v_1 = \|A\| v_1^*, A v_2 = \|A\|^{-1} v_2^*, v_2 = J v_1\) and \(v_2^* = J v_1^*\).

\textbf{Lemma A.2.} For any \(\hat{x} \in \mathbb{P}^1\), if \(x \in \hat{x}\) is a unit vector,
\begin{enumerate}
\item \(\|A x\| \geq \|A\| \frac{d(\hat{x}, \hat{v}_2)}{1}\),
\item \(d(A \hat{x}, \hat{v}_1^*) \leq \frac{d(\hat{x}, \hat{v}_2)}{\|A\|^2}\).
\end{enumerate}

\textbf{Proof.} Writing \(x = \langle x, v_1 \rangle v_1 + \langle x, v_2 \rangle v_2\),
\[
\|A x\| = \left\| \langle x, v_1 \rangle \|A\| v_1^* + \langle x, v_2 \rangle \|A\|^{-1} v_2^* \right\|
\]
\[
\geq \|\langle x, v_1 \rangle \|A\| v_1^*\| \geq |x \wedge v_2| \|A\| = d(\hat{x}, \hat{v}_2) \|A\|.
\]
Hence
\[
d(A \hat{x}, \hat{v}_1^*) = \frac{|(A x) \wedge v_1^*|}{\|A x\|} \leq \frac{|\langle x, v_2 \rangle| \|A\|^{-1} |v_2^* \wedge v_1^*|}{d(\hat{x}, \hat{v}_2) \|A\|} \leq \frac{1}{d(\hat{x}, \hat{v}_2) \|A\|^2}.
\]
\[
\square
\]

We denote by \(\lambda(A)\) the absolute value of the unstable eigenvalue of \(A\).

\textbf{Proposition A.3.} If \(a := |v_1^* \wedge v_2|\) then
\begin{enumerate}
\item \(\|\text{tr}(A)\| \geq a \|A\|\).
\item If \(a \|A\| > 2\) then \(A\) is hyperbolic and \(\lambda(A) \geq \frac{1}{2} \left( a \|A\| + \sqrt{a^2 \|A\|^2 - 4} \right) \geq a \|A\|\) as \(\|A\| \to \infty\).
\item There exists a function \(k(A) := 1 + O\left(\frac{1}{a^2 \|A\|^2}\right)\) such that
\[
d(\hat{s}(A), \hat{v}_2) \leq \frac{k(A)}{a \|A\|^2} \quad \text{and} \quad d(\hat{u}(A), \hat{v}_1^*) \leq \frac{k(A)}{a \|A\|^2}.
\]
Moreover, for any \(\hat{x} \in \mathbb{P}^1\),
\[
d(\hat{u}(A), A \hat{x}) \leq \frac{k(A)}{\|A\|^2} \left( \frac{1}{a} + \frac{1}{|x \wedge v_2|} \right).
\]
Proof. Item (1): Consider the basis \( v = \{v^*_1, v_2\} \). By Lemma A.1,

\[
|\text{tr}(A)| = \frac{1}{|v_1^* \wedge v_2|} |Av_1^* \wedge v_2 + v_1^* \wedge Av_2|
\]

\[
= \frac{1}{|v_1^* \wedge v_2|} \left| (v_1^*, v_1) Av_1 + (v_1^*, v_2) \wedge Av_2 \right| v_2 + \|A\|^{-1} (v_1^* \wedge v_2^*)
\]

\[
= \frac{1}{|v_1^* \wedge v_2|} \left| (v_1^*, v_1) (v_1^* \wedge v_2) \|A\| + \left\{ (v_1^*, v_2)(v_2^* \wedge v_2) + (v_2^* \wedge v_1^*) \right\} \|A\|^{-1} \right|
\]

\[
= \frac{1}{|v_1^* \wedge v_2|} \left( (v_1^* \wedge v_2)^2 \|A\| + \left( 1 - \langle v_1^*, v_2 \rangle \right)^2 \|A\|^{-1} \right]
\]

\[
\geq |v_1^* \wedge v_2| \|A\| + |v_1^* \wedge v_2|^{-1} \left( 1 - \langle v_1^*, v_2 \rangle \right)^2 \|A\|^{-1} \geq a \|A\|.
\]

Item (2): If \( a \|A\| > 2 \) then \( |\text{tr}(A)| \geq a \|A\| > 2 \) and \( A \) is hyperbolic. Therefore

\[
a \|A\| \leq |\text{tr}(A)| = \lambda + \lambda^{-1} \quad \text{with} \quad \lambda = \lambda(A).
\]

Solving in \( \lambda \) we get

\[
\lambda \geq \frac{1}{2} \left( a \|A\| + \sqrt{a^2 \|A\|^2 - 4} \right).
\]

Item (3): Because \( v_1^* = \langle v_1^*, v_1 \rangle v_1 + \langle v_1^*, v_2 \rangle v_2 \),

\[
Av_1^* = \langle v_1^*, v_1 \rangle \|A\| v_1^* + \langle v_1^*, v_2 \rangle \|A\|^{-1} v_2
\]

and whence

\[
d(\hat{v}_1^*, A \hat{v}_1^*) = \frac{|\langle v_1^*, v_2 \rangle| \|A\|^{-1} |v_1^* \wedge v_2^*|}{\sqrt{a^2 \|A\|^2 + (1 - a^2) \|A\|^{-2}}} \leq \frac{1}{a \|A\|^2}.
\]

Since near \( \hat{u}(A) \) the projective map \( \hat{A} \) is a Lipschitz contraction with Lipschitz constant of order \( \lambda(A)^{-2} \),

\[
d(\hat{u}(A), \hat{v}_1^*) \leq \frac{1}{a \|A\|^2} \frac{1}{1 - O(\lambda(A)^{-2})} = \frac{1}{a \|A\|^2} \left( 1 + O \left( \frac{1}{a^2 \|A\|^2} \right) \right).
\]

The bound on \( d(\hat{s}(A), \hat{w}) \) follows from the previous inequality applied to \( A^{-1} \). By Lemma A.2(b),

\[
d(A \hat{x}, \hat{v}_1^*) \leq \frac{1}{d(\hat{x}, \hat{v}_2) \|A\|^2}.
\]

Hence by the triangle inequality

\[
d(\hat{u}(A), A \hat{x}) \leq d(\hat{u}(A), \hat{v}_1^*) + d(\hat{v}_1^*, A \hat{x}) \leq \frac{k(A)}{\|A\|^2} \left( \frac{1}{a} + \frac{1}{|x \wedge v_2|} \right).
\]

\[\square\]

Lemma A.4. If \( \|w\| = 1 \) and \( d(\hat{w}, \hat{v}_1) < \sqrt{1 - \|A\|^{-2}} \), then \( \|Aw\| \geq 1 \).
Proof. By Lemma A.2(a),
\[ \|A w\| \geq \|A\| d(\hat{w}, \hat{v}_2) = \|A\| \sqrt{1 - d(\hat{w}, \hat{v}_1)^2} \geq \|A\| \|A\|^{-1} = 1. \]
\[\square\]

Lemma A.5. If \( d(A \hat{v}, \hat{v}_1^\ast) \leq \|A\|^{-1} \), \( d(A \hat{v}, \hat{w}) \geq a > 0 \) and \( a \|A\| > 2 \) then \( d(A^{-1} \hat{w}, \hat{v}_2) < \|A\|^{-1} \).

Proof. Note that \( d(\hat{w}, \hat{v}_1^\ast) \geq d(A \hat{v}, \hat{w}) - d(A \hat{v}, \hat{v}_1^\ast) \). By assumption this implies that \( d(\hat{w}, \hat{v}_1^\ast) \geq a - \|A\|^{-1} \geq \|A\|^{-1} \), or equivalently \( d(\hat{w}, \hat{v}_1^\ast) \leq \sqrt{1 - \|A\|^{-2}} \). Applying Lemma A.4 to \( A^{-1} \) we conclude that \( \|A^{-1} w\| \geq 1 \). Therefore, writing
\[
\begin{align*}
v &= x_v v_1 + y_v v_2, \quad \text{with } x_v^2 + y_v^2 = 1, \\
w &= x_w v_1^* + y_w v_2^*, \quad \text{with } x_w^2 + y_w^2 = 1, \\
A v &= x_v \|A\| v_1^* + y_v \|A\|^{-1} v_2^* \\
A^{-1} w &= x_w \|A\|^{-1} v_1 + y_w \|A\| v_2,
\end{align*}
\]
we have that
\[
d(A^{-1} \hat{w}, \hat{v}_2) = \frac{|A^{-1} w \wedge v_2|}{\|A^{-1} w\|} = \frac{|x_w\|A\|^{-1}|}{\|A^{-1} w\|} \leq \|A\|^{-1}.
\]
\[\square\]

Proposition A.6. If \( d(A \hat{v}, \hat{w}) \geq a \), \( d(A^{-1} \hat{w}, \hat{v}) \geq a \) and \( a \|A\| > 2 \) then
\[
\max \left\{ d(A \hat{v}, \hat{v}_1^\ast), \ d(A^{-1} \hat{w}, \hat{v}_2) \right\} \leq \frac{2}{a \|A\|^2 + \|A\| \sqrt{a^2 \|A\|^2 - 4}} \times \frac{1}{a \|A\|^2}.
\]

Proof. Without loss of generality me may assume that
\[
a \leq d(A \hat{v}, \hat{w}) \leq d(A^{-1} \hat{w}, \hat{v}),
\]
for otherwise we would replace the roles of \( \hat{v} \) and \( \hat{w} \), respectively of \( A \) and \( A^{-1} \). To prove the inequalities above we derive a system of recursive inequalities, which by iteration lead to fixed point bound. For this scheme to work we need the following preliminary inequalities:
\[
d(A \hat{v}, \hat{v}_1^\ast) \leq \|A\|^{-1} \quad \text{and} \quad d(A^{-1} \hat{w}, \hat{v}_2) \leq \|A\|^{-1}. \quad (A1)
\]

Choose \( v \) to have norm 1 and normalize \( w \) so that \( A^{-1} w \) has norm 1. There exist coordinates \((x_v, y_v)\) and \((x_w, y_w)\) in the unit circle such that
\[
\begin{align*}
v &= x_v v_1^* + y_v v_2^*, \\
A^{-1} w &= x_w v_1 + y_w v_2, \\
A v &= x_v \|A\| v_1^* + y_v \|A\|^{-1} v_2^*, \\
w &= x_w \|A\| v_1^* + y_w \|A\|^{-1} v_2^*.
\end{align*}
\]
We have

\[ d(A\hat{v}, \hat{v}_1^*) = \frac{|Av \wedge v_1^*|}{\|Av\|} = \frac{|y_v| \|A\|^{-1}}{\sqrt{x_v^2 \|A\|^2 + y_v^2 \|A\|^{-2}}} \]

and similarly

\[ \frac{|x_v y_w - x_w y_v|}{\sqrt{x_v^2 \|A\|^2 + y_v^2 \|A\|^{-2}}} = d(A\hat{v}, \hat{w}) \leq d(A^{-1}\hat{w}, \hat{v}) = |x_v y_w - x_w y_v|. \]

This implies that \( x_v^2 \|A\|^2 + y_v^2 \|A\|^{-2} \geq 1 \) and whence

\[ d(A\hat{v}, \hat{v}_1^*) \leq |y_v| \|A\|^{-1} \leq \|A\|^{-1}. \]

This proves the first inequality in (A1). The second follows from Lemma A.5.

We establish next the mentioned recursive inequalities. Since

\[ a \leq d(A^{-1}\hat{w}, \hat{v}) \leq d(A^{-1}\hat{w}, \hat{v}_2) + d(\hat{v}, \hat{v}_2) \]

we have

\[ d(\hat{v}, \hat{v}_2) \geq a - d(A^{-1}\hat{w}, \hat{v}_2) \geq a - \|A\|^{-1} \geq \|A\|^{-1} > 0 \]

and by Lemma A.2 (b),

\[ \|A\| d(A\hat{v}, \hat{v}_1^*) \leq \frac{1}{d(\hat{v}, \hat{v}_2)} \|A\| \leq \frac{1}{a \|A\| - d(A^{-1}\hat{w}, \hat{v}_2) \|A\|}. \tag{A2} \]

Similarly,

\[ a \leq d(\hat{w}, A\hat{v}) \leq d(\hat{w}, \hat{v}_1^*) + d(\hat{v}_1^*, A\hat{v}) \]

implies that,

\[ d(\hat{w}, \hat{v}_1^*) \geq a - d(\hat{v}_1^*, A\hat{v}) \geq a - \|A\|^{-1} \geq \|A\|^{-1} > 0. \]

Hence, as before,

\[ \|A\| d(A^{-1}\hat{w}, \hat{v}_2) \leq \frac{1}{d(\hat{w}, \hat{v}_1^*) \|A\|} \leq \frac{1}{a \|A\| - d(\hat{v}_1^*, A\hat{v}) \|A\|}. \tag{A3} \]

To solve the recursive inequalities (A2) and (A3), consider the 1-parameter family of partial maps \( F_b : \mathbb{R}^2 \to \mathbb{R}^2 \), \( F_b(x, y) := \left( \frac{1}{b-x}, \frac{1}{b-y} \right) \). For \( b > 2 \), each component of \( F_b \) is a well-defined contraction of the interval \([0, 1]\). Hence \( F_b \) leaves the square \([0, 1]^2\) invariant and is a strict contraction with unique fixed point \((x_*, y_*) := \left( \frac{2}{b+\sqrt{b^2-4}}, \frac{2}{b+\sqrt{b^2-4}} \right)\). Moreover, the maps \( F_b \) preserve the usual partial order of \( \mathbb{R}^2 \), defined by

\[ (x, y) \geq (x', y') \text{ if } x \geq x' \text{ and } y \geq y'. \]

Setting \( b := a \|A\| \) and \((x_0, y_0) := (\|A\| d(A\hat{v}, \hat{v}_1^*), \|A\| d(A^{-1}\hat{w}, \hat{v}_2))\), (A2) and (A3) are equivalent to \((x_0, y_0) \leq F_b(x_0, y_0)\), while (A1) ensures that \((x_0, y_0) \in [0, 1]^2\). Hence we obtain, inductively, that \((x_0, y_0) \leq F^n_b(x_0, y_0)\) for all \(n \geq 1\), and taking the limit as \(n \to \infty\), \((x_0, y_0) \leq (x_*, y_*)\). This concludes the proof. \(\square\)
Remark 3. By the previous lemma, if $a^* \|A\| > 2$ then $A$ is hyperbolic and

\[
\lambda(A) \geq \frac{a^* \|A\| + \sqrt{a^* 2 \|A\|^2 - 4}}{2}.
\]

Given $A \in SL_2(\mathbb{R})$ with $\|A\| > 1$, denote by $\hat{v}_1(A)$, $\hat{v}_2(A)$, $\hat{v}_1^*(A)$ and $\hat{v}_2^*(A)$ the unique projective points such that taking unit vectors $v_i \in \hat{v}_i(A)$ and $v_j^* \in \hat{v}_j(A)$, with $i, j = 1, 2$, $\{v_1, v_2\}$ and $\{v_1^*, v_2^*\}$ are singular basis of $A$ characterized by the relations $A v_1 = \|A\| v_1^*$ and $A v_2 = \|A\|^{-1} v_2^*$.

Lemma A.7. Given $A, A' \in SL_2(\mathbb{R})$ with $\|A\|, \|A'\| > 1$,

\[
\frac{\|A' A\|}{\|A\| \|A'\|} \left(1 - \frac{\|\hat{v}_1^*(A)\|^{-4} + \|A'\|^{-4}}{\left(\|A' A\|\right)^2}\right) \leq d(\hat{v}_1^*(A), \hat{v}_2(A')) \leq \frac{\|A' A\|}{\|A\| \|A'\|}.
\]

Proof. See [19, Propositions 2.23 and 2.24].

\[\Box\]

Proposition A.8 (Avalanche Principle). There exist positive constants $c_i, i = 0, 1, 2$ such that given $0 < \kappa < c_0 \epsilon^2$ and $A_0, \ldots, A_n \in SL_2(\mathbb{R})$, if

1. $\min_j \|A_j\|^2 \geq \kappa^{-1}$;
2. $\min_j \frac{\|A_j A_{j-1}\|}{\|A_{j-1}\|} \geq \epsilon$.

Then, for $A^n := A_{n-1} \cdots A_0$,

\[
\max \left\{ d(\hat{v}_1^*(A^n), \hat{v}_1^*(A_{n-1})), d(\hat{v}_2(A^n), \hat{v}_2(A_0)) \right\} \leq c_1 \kappa \epsilon^{-1}.
\]

and

\[
e^{-c_2 \kappa \epsilon^{-1} n} \leq \frac{\|A_{n-1} \cdots A_1 A_0\| \|A_1\| \cdots \|A_{n-2}\|}{\|A_1 A_0\| \cdots \|A_{n-1} A_{n-2}\|} \leq e^{c_2 \kappa \epsilon^{-1} n}.
\]

Proof. See [19, Proposition 2.42] or [21, Theorem 2.1].

\[\Box\]

Lemma A.9. Given $\hat{v}, \hat{w} \in \mathbb{P}^1$ and $A_1, \ldots, A_n \in SL_2(\mathbb{R})$ assume that:

(a) $\lambda \gg \gamma \gg t$;
(b) $\min_j \|A_j\| \geq e^\lambda$;
(c) $\min_j \|A_j A_{j-1}\| \geq e^{\gamma}$;
(d) $\min \left\{ d(A_1 \hat{v}, \hat{w}), d(A_1^{-1} \hat{w}, \hat{v}) \right\} \geq e^{-t}$;
(e) $\min \left\{ d(A_n \hat{v}, \hat{w}), d(A_n^{-1} \hat{w}, \hat{v}) \right\} \geq e^{-t}$;
(f) $d(A_n \hat{v}, A_n^{-1} \hat{w}) \geq e^{-t}$.

Then for all $j = 1, \ldots, n - 1$,

1. $d(\hat{v}_1^*(A^n) \hat{v}_2(A^n)) \geq e^{-t}$;
2. $A^n$ is hyperbolic and $\lambda(A^n) \geq e^{(\lambda - 2\gamma)n}$;
3. $\|A^n v\| \geq e^{(\lambda - 2\gamma)n} \quad \text{and} \quad \|A^n v\|^{-1} \hat{w} \geq e^{(\lambda - 2\gamma)n}$;
Proof. Using the conditions (d), (e) and Proposition A.6 we have
\[ d(A_n \hat{v}, \hat{v}_1(A_n)) \lesssim e^{-2\lambda t} \quad \text{and} \quad d(A_n^{-1} \hat{v}, \hat{v}_2(A_n)) \lesssim e^{-2\lambda t}. \]
By the AP (Proposition A.8) with \( \kappa := e^{-2\lambda} \) and \( \epsilon := e^{-\gamma} \),
\[ d(\hat{v}_1(A^n), \hat{v}_1(A^n)) \leq e^{-2\lambda + \gamma} \quad \text{and} \quad d(\hat{v}_2(A^n), \hat{v}_2(A^n)) \leq e^{-2\lambda + \gamma}. \]
Applying triangular inequality with condition (f),
\[ d(\hat{v}_1(A^n), \hat{v}_2(A^n)) \gtrsim e^{-t} \quad 2e^{-2\lambda + \gamma} - 2e^{-2\lambda + t} \gtrsim e^{-t}, \] (A4)
which give us Item 1. The AP also implies that
\[ \|A^n\| \gtrsim \exp\left( -e^{-2(\lambda - \gamma)} \right) \frac{\|A_2 A_1 \cdots A_n A_{n-1}\|}{\|A_2\| \cdots \|A_{n-1}\|} \gtrsim e^{-n\gamma} \|A_1\| \cdots \|A_n\| \geq e^{(\lambda - 2\gamma)n}. \] (A5)
Item 2 follows from inequality (A4) and Proposition A.3 with \( \lambda(A^n) \gtrsim e^{-t} \|A^n\| \gtrsim e^{(\lambda - 2\gamma)n} \).

Using the bounds above for \( d(A_n^{-1} \hat{v}, \hat{v}_2(A_1)) \), \( d(\hat{v}_2(A^n), \hat{v}_2(A_1)) \) and condition (d) we have
\[ d(\hat{v}, \hat{v}_2(A^n)) \gtrsim e^{-t} - e^{-2\lambda + t} - e^{-2\lambda + \gamma} \gtrsim e^{-t}. \] (A6)

Hence, by Lemma A.2, \( \|A^n v\| \gtrsim e^{(\lambda - \gamma)n - t} \). Using similar arguments we conclude that, \( \|A^n^{-1} w\| \gtrsim e^{(\lambda - 2\gamma)n} \) which proves item 3. \( \square \)

Appendix B: Derivative of Projective Actions

We state and prove some general formulas for the derivatives of the action that will be used throughout this section. Given a non-zero vector \( w \in \mathbb{R}^2 \), let \( z(w) \) be the unique unit vector which makes \( \{w/\|w\|, z(w)\} \) an orthonormal basis.

Given \( A \in SL_2(\mathbb{R}) \), the derivative of its projective action \( \hat{A} : \mathbb{P}^1 \to \mathbb{P}^1 \) is
\[ D\hat{A}(\hat{w})v = \frac{(Aw \wedge Av)}{\|Aw\|^2} z(Aw). \] (B7)

If \( w \in \hat{w} \) is a unit vector and \( v \in T_w \mathbb{P}^1 \) is a unit and positive tangent vector then \( A w \wedge A v = w \wedge v = 1 \) and the norm of the derivative \( D\hat{A}(\hat{w}) \) is equal to \( \|Aw\|^{-2} \).

For a \( C^1 \) one parameter family \( \{A_t\}_{t \in I} \) of \( SL_2(\mathbb{R}) \) matrices, where \( I \subseteq \mathbb{R} \) is an interval containing 0,
\[ \left. \frac{d}{dt} A_t \hat{v} \right|_{t=0} = \frac{(A_0 v \wedge \hat{A}_0 v)}{\|A_0 v\|^2} z(A_0 v). \] (B8)

More generally, given a one parameter family of cocycles \( A_t : X \to SL_2(\mathbb{R}) \), defined in some interval \( I \subseteq \mathbb{R} \),
Proposition B.1. If \( v_j(t) := A^j_t(x) v / \| A^j_t(x) v \| \) for \( j = 0, 1, \ldots, n \), then

\[
\frac{d}{dt} \frac{A^n_t(x) v}{\| A^n_t(x) v \|} = \sum_{j=0}^{n-1} \frac{z(A^n_t v)}{\| A^n_t(x) v \|^2} \left( v_{n-j} \wedge (\dot{A}_t A_t^{-1})(T^{n-j-1}x) v_{n-j} \right)
\]

Proof. Since

\[
\frac{d}{dt} A^n_t(x) = \sum_{j=1}^{n} A_t^{n-j}(T^j x) \dot{A}_t(T^{j-1}x) A_t^{j-1}(x)
\]

we have

\[
\frac{d}{dt} \frac{A^n_t(x) v}{\| A^n_t(x) v \|} = \sum_{j=1}^{n} \frac{A_t^n(x) v \wedge A_t^{n-j}(T^j x) \dot{A}_t(T^{j-1}x) A_t^{j-1}(x) v}{\| A^n_t(x) v \|^2} z(A^n_t v)
\]

\[
= \sum_{j=1}^{n} \frac{A_t^j(x) v \wedge \dot{A}_t(T^{j-1}x) A_t^{j-1}(x) v}{\| A_t^n(x) v \|^2} \frac{1}{z(A^n_t v)}
\]

\[
= \sum_{j=1}^{n} \frac{1}{\| A_t^{n-j}(T^j x) v_j \|^2} \left( v_j \wedge (\dot{A}_t A_t^{-1})(T^{j-1}x) v_j \right) \frac{1}{z(A^n_t v)}
\]

\[
= \sum_{j=0}^{n-1} \frac{1}{\| A_t^{n-j}(T^j x) v_{n-j} \|^2} \left( v_{n-j} \wedge (\dot{A}_t A_t^{-1})(T^{n-j-1}x) v_{n-j} \right) \frac{1}{z(A^n_t v)}
\]

Lemma B.2. Given a compact interval \( I \subset \mathbb{R} \), there exist \( C, c > 0 \) such that for every \( x \in X \), \( \hat{v} \in \mathbb{P}^1 \) and \( n \in \mathbb{N} \)

\[
\left| \frac{d}{dt} A^n_t(x) \hat{v} \right| \leq C e^{cn}.
\]

Proof. Let \( M_0 = \sup_{t \in I} \| \dot{A}_t A_t^{-1} \|_\infty \) and \( M_1 = \sup_{t \in I} \| A_t \|_\infty \). Then, for every \( \hat{w} \in \mathbb{P}^1 \),

\[
\| A^j(x) \hat{w} \| \geq M_1^{-j}. \]

Therefore, by Proposition B.1,

\[
\left| \frac{d}{dt} A^n_t(x) \hat{w} \right| \leq M_0 \sum_{j=0}^{n-1} M_1^j \leq M_0 \frac{M_1^n}{M_1 - 1} = C e^{cn},
\]

where \( C = \frac{M_0}{M_1 - 1} \) and \( c = \log M_1 \). \( \square \)
For Schrödinger cocycles more can be said.

**Lemma B.3.** Given a Schrödinger cocycle \( A : X \to SL_2(\mathbb{R}) \) with continuous potential \( \phi : X \to \mathbb{R} \) and generated by the dynamical system \((X, T, \xi)\), there exists a constant \( c_* > 0 \) such that for all \( n \geq 2 \), all \( x \in X \), all \( E \in \mathbb{R} \) and all \( \hat{v} \in \mathbb{P}^1 \),

\[
\left| \frac{d}{dE} A^n_E(x) \hat{v} \right| = \frac{A^n_E(x) v \wedge \frac{d}{dE} A^n_E(x) v}{\| A^n_E(x) v \|^2} \geq c_*.
\]

and

\[
\left| \frac{d}{dE} A^{-n}_E(x) \hat{v} \right| = -\frac{A^{-n}_E(x) v \wedge \frac{d}{dE} A^{-n}_E(x) v}{\| A^{-n}_E(x) v \|^2} \geq c_*.
\]

**Proof.** For \( n = 1 \) a simple calculation gives

\[
\frac{A_E(x) v \wedge \dot{A}_E(x) v}{\| A_E(x) v \|^2} = \frac{v_1^2}{\| A_E(x) v \|^2}
\]

while using Proposition B.1 with \( n = 2 \)

\[
\frac{A^2_E(x) v \wedge \frac{d}{dE} A^2_E(x) v}{\| A^2_E(x) v \|^2} = \frac{v_1^2}{\| A^2_E(x) v \|^2} + \frac{((\phi(x) - E) v_1 - v_2)^2}{\| A_E(Tx) v \|^2}
\]

is positive and bounded away from 0 because the denominators are bounded and the numerators add up to a positive definite quadratic form \( v_1^2 + ((\phi(x) - E) v_1 - v_2)^2 \), for any \( x \in X \) and \( E \in \mathbb{R} \). The general case follows also from Proposition B.1 neglecting all terms but the last two.

Finally, since \( A^{-n}_E(x) = A^n_E(T^{-n}x)^{-1} \), if \( w = A^{-n}_E(x) v \) then

\[
A^{-n}_E(x) v \wedge \frac{d}{dE} A^{-n}_E(x) v = A^{-n}_E(x) v \wedge A^{-n}_E(x) \left( -\frac{d}{dE} A^n_E(T^{-n}x) \right) A^{-n}_E(x) v
\]

\[
= -v \wedge \left( \frac{d}{dE} A^n_E(T^{-n}x) \right) A^{-n}_E(x) v
\]

\[
= -A^n_E(T^{-n}x) w \wedge \frac{d}{dE} A^n_E(T^{-n}x) w < 0.
\]

Therefore the projective curve \( E \mapsto A^{-n}_E(x) \hat{v} \) winds in the opposite direction, and a similar argument gives that \( \left| \frac{d}{dE} A^{-n}_E(x) \hat{v} \right| \) is bounded away from 0. \( \square \)

The space \( \mathbb{P}^1 \) has a natural orientation. We say that a curve \( \hat{v} : I \to \mathbb{P}^1 \) winds positively, resp. negatively, when it is positively, resp. negatively, oriented.

**Proposition B.4.** If \( M_E \) is hyperbolic with \( \lambda(M_E) \geq \lambda_0 > 1 \) for all \( E \) in some compact interval \( I \), then there exists a positive constant \( K < \infty \) such that the projective curves \( E \mapsto \hat{v}(M_E) \) and \( E \mapsto \hat{s}(M_E) \) wind around \( \mathbb{P}^1 \) in opposite directions with non-zero speed bounded from above by \( K \). The unstable curve \( E \mapsto \hat{u}(M_E) \) winds positively, while the stable curve \( E \mapsto \hat{s}(M_E) \) winds negatively. Moreover

\[
\frac{d}{dE} \hat{u}(M_E) = \frac{\lambda(E) (u(E) \wedge \dot{M}_E u(E))}{\lambda(E)^2 - 1} \zeta(\hat{u}(M_E))
\]
and

\[
\frac{d}{dE} \hat{s}(M_E) = -\frac{\lambda(E)}{(\lambda(E))^2 - 1} (s(E) \wedge \hat{M}_E^{-1}s(E)) z(\hat{s}(M_E)).
\]

Proof. Breaking the interval \( I \) into finitely many sub-intervals, if necessary, we can take \( \hat{e}_1 \in \mathbb{P}^1 \) such that \( \hat{e}_1 \neq \hat{s}(M_E) \), for all \( E \in I \). Hence \( \hat{u}(M_E) = \lim_{k \to \infty} M_E^{-k} \hat{e}_1 \), with uniform convergence of the functions and their derivatives over the compact interval \( I \).

As \( k \to \infty \), the unit vectors \( v_k(E) := \frac{M_E^{k} e_1}{\|M_E^{k} e_1\|} \) converge geometrically and uniformly to an eigenvector \( u(E) \) of \( M_E \) in \( \hat{u}(M_E) \) and, denoting by \( \lambda(E) \) the absolute value of the corresponding eigenvalue, by Proposition B.1 we have

\[
\frac{d}{dE} \hat{u}(M_E) = \lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \frac{1}{\|M_E^{j} v_{k-j}\|^2} \left( v_{k-j} \wedge \hat{M}_E M_E^{-1} v_{k-j} \right) z(M_E^{k} \hat{e}_1)
\]

\[
= \sum_{j=0}^{\infty} \frac{1}{\|M_E^{j} u(E)\|^2} \left( u(E) \wedge \hat{M}_E M_E^{-1} u(E) \right) z(\hat{u}(M_E))
\]

\[
= \sum_{j=0}^{\infty} \frac{1}{\lambda(E)^{2j+1}} u(E) \wedge \hat{M}_E u(E) z(\hat{u}(M_E))
\]

\[
= \frac{\lambda(E) (u(E) \wedge \hat{M}_E u(E))}{\lambda(E)^2 - 1} z(\hat{u}(M_E)).
\]

This concludes the argument for the unstable curve \( \hat{u}(M_E) \). The stable curve \( \hat{s}(M_E) \) winds negatively because \( \hat{s}(M_E) = \lim_{k \to \infty} M_E^{-k} \hat{e}_2 \), for any vector \( e_2 \) such that \( \hat{e}_2 \neq \hat{u}(M_E) \), for all \( E \in I \). See Lemma B.3. The bound for the derivative of this stable curve is obtained in a similar way.

\[\square\]

References

1. Furstenberg, H., Kesten, H.: Products of random matrices. Ann. Math. Statist. 31, 457–469 (1960)
2. Furstenberg, H., Kifer, Y.: Random matrix products and measures on projective spaces. Israel J. Math. 46(1–2), 12–32 (1983). https://doi.org/10.1007/BF02760620
3. Furman, A.: Random walks on groups and random transformations. In: Handbook of Dynamical Systems, Vol. 1A, pp. 931–1014. North-Holland, Amsterdam, (2002). https://doi.org/10.1016/S1874-575X(02)80014-5
4. Bocker-Neto, C., Viana, M.: Continuity of Lyapunov exponents for random two-dimensional matrices. Ergod. Theory Dyn. Syst. 37, 1413 (2016)
5. Viana, M.: Lectures on Lyapunov Exponents. Cambridge Studies in Advanced Mathematics, p. 215. Cambridge University Press, (2014)
6. Le Page, E.: Régularité du plus grand exposant caractéristique des produits de matrices aléatoires indépendantes et applications. Ann. Inst. H. Poincaré Probab. Statist. 25(2), 109–142 (1989)
7. Baraviera, A., Duarte, P.: Approximating Lyapunov exponents and stationary measures. J. Dynam. Differ. Equ. 31(1), 25–48 (2019). https://doi.org/10.1007/s10884-018-9724-5
8. Tall, E.H.Y., Viana, M.: Moduli of continuity for the Lyapunov exponents of random \( \text{GL}(2) \)-cocycles. Trans. Am. Math. Soc. 373(2), 1343–1383 (2020). https://doi.org/10.1090/tran/7973
9. Duarte, P., Klein, S.: Large deviations for products of random two dimensional matrices. Commun. Math. Phys. 375(3), 2191–2257 (2020). https://doi.org/10.1007/s00220-019-03586-2
10. Simon, B., Taylor, M.: Harmonic analysis on \( \text{SL}(2, R) \) and smoothness of the density of states in the one-dimensional Anderson model. Commun. Math. Phys. 101(1), 1–19 (1985)
11. Duarte, P., Klein, S., Santos, M.: A random cocycle with non Hölder Lyapunov exponent. Discrete Contin. Dyn. Syst. 39(8), 4841–4861 (2019). https://doi.org/10.3934/dcds.2019197

12. Ruelle, D.: Analyticity properties of the characteristic exponents of random matrix products. Adv. Math. 32(1), 68–80 (1979). https://doi.org/10.1016/0001-8708(79)90029-X

13. Avila, A., Bochi, J., Yoccoz, J.-C.: Uniformly hyperbolic finite-valued SL(2, R)-cocycles. Comment. Math. Helv. 85(4), 813–884 (2010). https://doi.org/10.4171/CMH/212

14. Ledrappier, F., Young, L.-S.: The metric entropy of diffeomorphisms. Bull. Am. Math. Soc. 11(2), 343–346 (1984)

15. Ledrappier, F., Young, L.-S.: The metric entropy of diffeomorphisms: part ii: relations between entropy, exponents and dimension. Ann. Math. 122, 540–574 (1985)

16. Bárány, B.: On the Ledrappier-Young formula for self-affine measures. Math. Proc. Cambridge Philos. Soc. 159(3), 405–432 (2015). https://doi.org/10.1017/S0305004115000344

17. Ledrappier, F.: Une relation entre entropie, dimension et exposant pour certaines marches aléatoires. CR Acad. Sci. Paris Sér. I Math 296(8), 369–372 (1983)

18. Hochman, M., Solomyak, B.: On the dimension of Furstenberg measure for $SL_2(R)$ random matrix products. Invent. Math. 210(3), 815–875 (2017)

19. Duarte, P., Klein, S.: Lyapunov Exponents of Linear Cocycles; Continuity Via Large Deviations. Atlantis Studies in Dynamical Systems, vol. 3. Atlantis Press, (2016)

20. Kifer, Y.: Perturbations of random matrix products. Z. Wahrsch. Verw. Gebiete 61(1), 83–95 (1982). https://doi.org/10.1007/BF00537227

21. Duarte, P., Klein, S.: Continuity of the Lyapunov Exponents of Linear Cocycles. Publicações Matemáticas, 31° Colóquio Brasileiro de Matemática, p. 142. IMPA, (2017). available at https://impa.br/wp-content/uploads/2017/08/31CBM_02.pdf

22. Hoeffding, W.: Probability inequalities for sums of bounded random variables. J. Am. Statist. Assoc. 58, 13–30 (1963)

23. Damanik, D.: Schrödinger operators with dynamically defined potentials. Ergod. Theory Dynam. Syst. 37(6), 1681–1764 (2017). https://doi.org/10.1017/etds.2015.120

24. Goldstein, M., Schlag, W.: Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equations and averages of shifts of subharmonic functions. Ann. Math. 154(1), 155–203 (2001). https://doi.org/10.2307/3062114

25. Avila, A.: Density of positive lyapunov exponents for $SL(2, R)$-cocycles. J. Am. Math. Soc. 24(4), 999–1014 (2011)

26. Bougerol, P., Lacroix, J.: Products of Random Matrices with Applications to Schrödinger Operators. Progress in Probability and Statistics, vol. 8. Birkhäuser Boston, Inc., Boston, MA, (1985). https://doi.org/10.1007/978-1-4684-9172-2

27. Guivarc’h, Y.: Produits de matrices aléatoires et applications aux propriétés géométriques des sous-groupes du groupe linéaire. Ergod. Theory Dynam. Syst. 10(3), 483–512 (1990)

28. Benoist, Y., Quint, J.-F.: Random walks on reductive groups. In: Random Walks on Reductive Groups, pp. 153–167. Springer (2016)

29. Johnson, R.A.: Exponential dichotomy, rotation number, and linear differential operators with bounded coefficients. J. Differ. Equ. 61(1), 54–78 (1986)

Communicated by C. Liverani