Thesis for the Masters in Mathematical Finance

Products of i.i.d. random matrices and a theorem by Furstenberg

Author:

Martim da Costa

Supervisor:

João Lopes Dias

December 22, 2023

Contents

Acknowledgements

The idea of writing a thesis on this topic was originally proposed by Professor João Lopes Dias. His guidance was essential for the completion of this project and I express my gratitude for all the time he dedicated to answering my questions.

I thank Professor Jairo Bochi for authoring the paper on which this work essentially builds upon and for quickly answering all my questions despite having no formal responsibility to do so.

Finally, I thank my mother, Berta Alves, for her unconditional support.

Abstract

We study the Lyapunov exponents associated to the product of i.i.d. random matrices linear cocyle in $\mathrm{SL}_{\pm}(2,\mathbb{R})$. The existence of these quantities and conditions to guarantee strict positivity are established. These results are used to prove the exponential growth of a random Fibonacci sequence.

1 Introduction

The primary goal of this paper is to provide a reasonably self-contained, accessible to graduate students, introduction to the work of H. Furstenberg in [5] and [4], which laid the foundation for the study of Lyapunov exponents of linear cocyles. Lyapunov exponents quantify the exponential norm-growth of a dynamical system. The Furstenberg-Kesten theorem first established their existence for products of random matrices. If the associated Lyapunov exponent is positive, then we conclude that the system displays exponential growth. Under suitable conditions, this is precisely what Furstenberg's theorem gives a positive answer to.

The proofs that establish this result make use of mathematics which are not typically part of, or emphasized, in a standard introductory course on measure theory. The secondary goal of this paper is to include the detail and background which is often omitted from the literature, thus providing a gentler presentation of these topics, accessible to a wider audience. This accounts for the rather lengthy preliminary section which follows this introduction.

Following the section on preliminaries, we formally state the result and establish useful equivalent conditions. We use these to study the random Fibonacci sequence, and prove that its associated Lyapunov exponent is positive. The rest of the paper is dedicated to proving the stated result.

Our presentation follows [2] and [3] closely, as such, it is a particular version of the general result established by H. Furstenberg, which instead of $\mathrm{SL}_{\pm}(2, \mathbb{R})$ considers the general case of $\mathrm{SL}_{\pm}(n, \mathbb{R})$.

2 Preliminaries

Let $(\Omega, \mathcal{F}, \rho)$ be a probability space and $f : \Omega \to \Omega$ a measurable map which preserves ρ, i.e. $f_*\rho = \rho \circ f^{-1} = \rho$ on $\mathcal{F}$ (we also say that ρ is *f-invariant*). Moreover, we denote the set of all 2×2 matrices with determinant ± 1 by $\mathrm{SL}_{\pm}(2, \mathbb{R})$. If $A : \Omega \to \mathrm{SL}_{\pm}(2, \mathbb{R})$ is measurable, we construct a skew-product map given by

$$\begin{aligned} T : \Omega \times \mathbb{R}^2 &\to \Omega \times \mathbb{R}^2 \\ (\omega, v) &\mapsto (f(\omega), A(\omega)v). \end{aligned}$$

This is called the *linear cocyle* of A over f, and usually denoted by $T = (f, A)$. The orbit under T of the point $(\omega, v) \in \Omega \times \mathbb{R}^2$ is

$$T^n(\omega, v) = (f^n(\omega), A^{(n)}(\omega)v)$$

where

$$A^{(n)}(\omega) = A(f^{n-1}(\omega))A(f^{n-2}(\omega)) \cdots A(f(\omega))A(\omega).$$

In this paper we are interested in studying one particular linear cocyle. Let $(G, \mathcal{X}, \mu)$ be a probability space with $G \subseteq \mathrm{SL}_{\pm}(2, \mathbb{R})$. Define σ to be the shift map

$$\begin{aligned} \sigma : G^{\mathbb{N}} &\to G^{\mathbb{N}} \\ (\omega_1, \omega_2, \omega_3, \ldots) &\mapsto (\omega_2, \omega_3, \ldots) \end{aligned}$$

and $A : G^{\mathbb{N}} \to G$ defined by $(\omega_1, \omega_2, \ldots) \mapsto \omega_1$. Both maps are measurable with respect to the infinite-dimensional product space $(G^{\mathbb{N}}, \mathcal{X}^{\mathbb{N}}, \mu^{\mathbb{N}})$, where $\mathcal{X}^{\mathbb{N}}$ denotes the σ-algebra generated by the cylinder sets and $\mu^{\mathbb{N}}$ denotes the product probability measure. The cocyle (σ, A) is called the *product of i.i.d. random matrices*. Its dynamics are given by

$$T^n((\omega_1, \omega_2, \ldots), v) = ((\omega_{n+1}, \omega_{n+2}, \ldots), \omega_n \cdots \omega_1 v).$$

Throughout this text, M_n will denote the random variable defined by $M_n(\omega) = A(\sigma^{n-1}(\omega)) = \omega_n$.

2.1 Ergodic theory

The map $\tau_v : \mathbb{R}^n \to \mathbb{R}^n$ defined by $x \mapsto \tau_v(x) = x + v$ for a fixed $v \in \mathbb{R}^n$ is called a *translation*. The Lebesgue measure λ is known to be the unique measure on $\mathbb{R}^n$ which is τ_v-invariant for all v.

Definition 1. A group $(G, \cdot)$ together with a topology $\mathcal{T}$ is a *topological group* if the maps

$$\begin{aligned} G \times G &\to G & \qquad G &\to G \\ (x, y) &\mapsto x \cdot y, & \qquad x &\mapsto x^{-1} \end{aligned}$$

are continuous.

Definition 2. Let G be a topological group and κ be a measure on $\mathcal{B}(G)$[1]. The measure κ is said to be:

[1]This is the Borel σ-algebra on G.

- *left-translation-invariant* if $\kappa(gA) = \kappa(A)$ for all $A \in \mathcal{B}(G)$ and all $g \in G$;
- *right-translation-invariant* if $\kappa(Ag) = \kappa(A)$ for all $A \in \mathcal{B}(G)$ and all $g \in G$.

Observe that $\mathbb{R}^n$ taken with its usual addition is a topological group. The usual translation of a set $A \subset \mathbb{R}^n$ by a vector v can thus be represented by vA, which is equal to Av by commutativity, so the Lebesgue measure is one example of a measure which is both left- and right-translation-invariant. It is then natural to wonder about the existence of analogous measures in other topological groups.

Theorem 3. *Let G be a compact topological group. There exists a unique probability measure on $\mathcal{B}(G)$ that is both left- and right-translation-invariant. We call it the Haar measure on G.*

Proof. See [7]. □

The following theorem is a fundamental result in ergodic theory. It establishes a connection between the long-term behaviour of a dynamical system and its expected value.

Theorem 4 (Birkhoff's ergodic theorem)**.** *Let $(X, \mathcal{X}, \kappa)$ be a probability space and $f : X \to X$ be a measurable map such that κ is f-invariant. If $\phi : X \to \mathbb{R}$ is κ-integrable, then the limit*

$$\phi_f(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \phi(f^k(x))$$

exists for κ-a.e. x and

$$\int_X \phi_f \, d\kappa = \int_X \phi \, d\kappa.$$

The sum $\phi(x) + \phi(f(x)) + \cdots + \phi(f^{n-1}(x))$ is called the Birkhoff sum of ϕ.

Proof. See [6]. $\square$

Definition 5. Let $(X, \mathcal{X}, \kappa)$ be a probability space and $T : X \to X$ a measurable transformation. The map T is said to be κ-*ergodic* if for all $A \in \mathcal{X}$

$$T^{-1}(A) = A \implies \kappa(A) = 1 \text{ or } \kappa(A) = 0.$$

Theorem 6 (Kingman's subadditive ergodic theorem)**.** *Let T be a measure preserving transformation on a probability space $(X, \mathcal{X}, \kappa)$ and $(g_n)_{n\in\mathbb{N}}$ be a sequence of integrable functions which is subadditive, i.e.*

$$g_{n+m}(\omega) \leq g_n(\omega) + g_m(T^n\omega).$$

Then,

$$\lim_{n\to\infty} \frac{g_n(\omega)}{n} = g(\omega) \in \mathbb{R} \cup \{-\infty\}$$

for κ-a.e. ω, where g is a T-invariant function. If T is ergodic, then g is constant.

Proof. See [8]. $\square$

The proof of the following lemma follows [9] as well as [2].

Proposition 7. *Let (X, κ) be a probability space and $T : X \to X$ a measurable transformation. Suppose κ is T-invariant and let $f \in L^1(\kappa)$ be a* *function which satisfies*

$$\lim_{n\to\infty}\sum_{j=0}^{n-1} f(T^j(x)) = +\infty \tag{1}$$

for κ-almost every x. Then

$$\int_X f\,d\kappa > 0.$$

Proof. Let $(s_n)_{n\in\mathbb{N}}$ be the sequence defined by

$$s_n = \sum_{j=0}^{n-1} f\circ T^j.$$

For $\varepsilon > 0$ we define the two following sets:

$$A_\varepsilon = \{x\in X : \forall n\in\mathbb{N} : s_n(x)\geq\varepsilon\} \text{ and } B_\varepsilon = \bigcup_{k\geq 0} T^{-k}(A_\varepsilon).$$

We begin by proving that

$$\kappa\left(\bigcup_{\varepsilon>0} B_\varepsilon\right) = 1. \tag{2}$$

Suppose that x is such that (1) is satisfied and that for every $\varepsilon > 0$ we have $x\notin B_\varepsilon$. In particular $x\notin B_{1/l^2}$ for any $l\geq 1$, i.e. $T^k(x)\notin A_{1/l^2}$ for all $k\geq 0$, or, equivalently, for all $k\geq 0$ there exists $n_l\in\mathbb{N}$ such that $s_{n_l}(T^k x) < {}^1\!/l^2$. Therefore

$$\lim_{l\to\infty} s_{n_1+\cdots n_l}(x) < \sum_{l=1}^{\infty}\frac{1}{l^2} = \frac{\pi^2}{6}$$

contradicting (1).

Now fix $\varepsilon > 0$ and let $x \in B_\varepsilon$. Then, there exists at least one $k \geq 0$ such that $T^k x \in A_\varepsilon$. Let k_x denote the smallest such k. This entails that

$$\begin{aligned} s_n(T^{k_x}x) &= \sum_{j=0}^{n-1} f(T^j(T^{k_x}x)) \\ &= \sum_{j=k_x}^{n-1} f(T^j x) \\ &\geq \varepsilon \end{aligned}$$

for every $n \geq k_x + 1$. Therefore

$$\sum_{j=0}^{n-1} f(T^j x) \geq \sum_{j=0}^{k_x-1} f(T^j x) + \sum_{j=k_x}^{n-1} \varepsilon \mathbb{1}_{A_\varepsilon}(T^j x). \tag{3}$$

Let φ and ψ denote the limit of the Birkhoff averages of f and $\mathbb{1}_{A_\varepsilon}$ respectively. Divide (3) by n and then let $n \to \infty$. We obtain

$$\varphi(x) \geq \varepsilon\psi(x). \tag{4}$$

Now note that

$$\begin{aligned}
\int \psi(x)\, d\kappa(x) &= \int \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \mathbb{1}_{A_\varepsilon}(T^j x)\, d\kappa(x) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \kappa(T^{-j}(A_\varepsilon)) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \kappa(A_\varepsilon) \\
&= \kappa(A_\varepsilon)
\end{aligned}$$

and, since $\varphi(x) = 0$ when $x \notin B_\varepsilon$,

$$\begin{aligned}
\int \psi(x)\, d\kappa(x) &= \int_{B_\varepsilon} \psi(x)\, d\kappa(x) \\
&= \kappa(B_\varepsilon).
\end{aligned}$$

By Birkhoff's ergodic Theorem,

$$\int f\, d\kappa = \int \varphi\, d\kappa \geq 0$$

so we only need to exclude the case of an equality, which is equivalent to saying $\varphi = 0$ almost everywhere. Assume this is the case. By (4) $0 = \varphi(x) \geq \varepsilon\kappa(B_\varepsilon)$ and therefore $\kappa(B_\varepsilon) = 0$ for all ε, which contradicts (2). The desired result follows from this contradiction. $\square$

2.2 Lyapunov exponents

Lyapunov exponents are quantities associated to a linear cocyle. For a linear cocyle (f, A), its (upper) Lyapunov exponent γ is defined as

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \left\| A^{(n)} \right\|.$$

If the cocycle in question is the product of 2×2 random matrices, which we intend to study, we obtain

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|M_n \cdots M_1\|. \tag{5}$$

Suppose such a quantity exists and consider another, arbitrary norm $\|\cdot\|_*$ on $\mathbb{R}^{2\times 2}$. From the finite dimension of $\mathbb{R}^{2\times 2}$ it follows that any two norms are equivalent. Therefore there exists a pair of real numbers $0 < C_1 < C_2$ such that the following inequality is satisfied

$$C_1\|M_n \cdots M_1\| \leq \|M_n \cdots M_1\|_* \leq C_2\|M_n \cdots M_1\|.$$

The logarithm function preserves the inequalities. We can then divide by n and take the limit to obtain

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \log \|M_n \cdots M_1\|_*.$$

This establishes that γ does not depend on the chosen norm, supposing its existence. We now turn to the question of whether or not γ is well-defined.

For an arbitrary function g define g^+ to be the map $x \mapsto \sup(f(x), 0)$. If

$\log^+ \|M_1\|$ is integrable and $n, p \geq 1$ then

$$\log \left\|A^{(n+p)}(\omega)\right\| \leq \log \|M_{n+p}(\omega) \cdots M_{n+1}(\omega)\| + \log \|M_n(\omega) \cdots M_1(\omega)\|$$
$$= \log \left\|A^{(p)}(\sigma^n(\omega))\right\| + \log \left\|A^{(n)}(\omega)\right\|$$

by the submultiplicative property of matrix norms. Consequently the sequence $(\log ||A^{(n)}||)_{n\in\mathbb{N}}$ is subadditive and integrable. By Theorem 6 we have

$$\frac{1}{n} \log \left\|A^{(n)}(\omega)\right\| \to \gamma(\omega) \in \mathbb{R} \cup \{-\infty\}$$

for μ-a.e. $\omega \in \Omega^{\mathbb{N}}$. Since σ is $\mu^{\mathbb{N}}$-ergodic, γ is almost surely constant. This establishes the conditions for the existence of the Lyapunov exponent γ.

Unless stated otherwise, we fix $\|M\|$ to be the spectral norm of a matrix M, i.e. the square root of the maximum eigenvalue of $M^\intercal M$ and $\|v\|$ the usual euclidean norm for a vector $v \in \mathbb{R}^2$.

2.2.1 Examples

Example 1. Consider a probability space $(\Omega, \mathcal{F}, \mu)$ such that $\text{supp}(\mu) = O(2)$[2]. We have

$$\begin{aligned}\gamma &= \int_{O(2)^{\mathbb{N}}} \lim_{n\to\infty} \frac{1}{n} \log \|M_n(\omega) \cdots M_1(\omega)\| \, d\mu^{\mathbb{N}}(\omega) \\ &\leq \lim_{n\to\infty} \frac{1}{n} \log(1) \\ &= 0\end{aligned}$$

[2]For a brief discussion of $O(2)$ see section 2.4.

using the fact that the product $M_n(\omega)\cdots M_1(\omega)$ is an element of $O(2)$ by closure of the group operation, so $\|M_n(\omega)\cdots M_1(\omega)\| = 1$.

Example 2. Suppose

$$\operatorname{supp}(\mu) = \left\{ \begin{bmatrix} t & 0 \\ 0 & {}^1\!/_t \end{bmatrix} : t \geq 1 \right\}.$$

Observe that

$$M_n \cdots M_1 = \begin{bmatrix} t_n & 0 \\ 0 & {}^1\!/_{t_n} \end{bmatrix} \cdots \begin{bmatrix} t_1 & 0 \\ 0 & {}^1\!/_{t_1} \end{bmatrix} = \begin{bmatrix} t_n \cdots t_1 & 0 \\ 0 & (t_n \cdots t_1)^{-1} \end{bmatrix}.$$

So $\|M_n \cdots M_1\| = \max\{t_n \cdots t_1, (t_n \cdots t_1)^{-1}\} = t_n \cdots t_1$. Hence,

$$\log(\|M_n \cdots M_1\|) = \log(t_n \cdots t_1).$$

By the usual law of large numbers

$$\gamma = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \log(t_i) = \mathbb{E}\left[\log \|M_1\|\right] > 0.$$

2.3 The one-dimensional projective space

We define the real projective space of dimension one by first defining an equivalence relation $\sim$ on $\mathbb{R}^2 \setminus \{0\}$, stipulating that $x \sim y$ iff there exists $\alpha \in \mathbb{R}$ such that $x = \alpha y$. The real projective space is defined as the quotient

$$\mathbb{RP}^1 = {}^{\mathbb{R}^2\setminus\{0\}}\!/_{\sim},$$

i.e. the set of all equivalence classes. The equivalence class, or *direction*, of $x \in \mathbb{R}^2 \setminus \{0\}$ will be denoted by $\bar{x}$ and may be thought of as a straight line passing through the origin or as the set of all linear combinations of x denoted by $\operatorname{span}\{x\}$. Such lines are entirely characterized by the angle they form with the horizontal axis. Therefore there is an intuitive identification between $\mathbb{RP}^1$ and the interval $[0, \pi)$ and the two sets may be regarded as interchangeable when convenient.

A matrix A in $\mathrm{GL}(2, \mathbb{R})$ induces a transformation on $\mathbb{RP}^1$ in a straightforward manner: we start with an element $\bar{x} \in \mathbb{RP}^1$ and consider $x \in \bar{x}$. We then perform the standard matrix multiplication Ax and take the equivalence class $\overline{Ax}$. This procedure results in a well-defined function, given that it does not depend on the choice of x. In other words, for a given matrix $A \in \mathrm{GL}(2, \mathbb{R})$, the induced map $\bar{A} : \mathbb{RP}^1 \to \mathbb{RP}^1$, is defined by $\bar{x} \mapsto \overline{Ax}$. We selectively adhere to this notational distinction between a matrix and the map it induces in the projective space, and likewise for a vector and its corresponding direction. More often than not, we simply use the same symbol for both, unless we deem it confusing.

The following lemma will be used later on.

Lemma 8. *If $A \in \mathbb{R}^{2\times 2}$ has rank one and ν is a probability measure on $\mathbb{RP}^1$, then $\bar{A}_*\nu$ is a Dirac measure.*

Proof. Since $\operatorname{rank}(A) = 1$, then there exists $x \in \mathbb{R}^2$ such that range $A = \operatorname{span}(x) = \bar{x}$. Let $B \subseteq \mathbb{RP}^1$ be a measurable set. Then

$$\bar{A}_*\nu(B) = \nu(\bar{A}^{-1}(B)) = \begin{cases} \nu(\mathbb{RP}^1) & \text{if } \bar{x} \in B \\ \nu(\emptyset) & \text{otherwise.} \end{cases}$$

Therefore $\bar{A}_* \nu = \delta_{\bar{x}}$. $\square$

Finally, we note that $\mathbb{RP}^1$ is a compact and separable topological space.

2.4 Orthogonal matrices

A matrix $A \in \mathbb{R}^{n \times n}$ is said to be *orthogonal* if $AA^\intercal = I$. We denote the set of all orthogonal $n \times n$ matrices by $O(n)$, which is easily checked to be a subgroup of $\mathrm{GL}(n, \mathbb{R})$.

Suppose $A \in O(n)$. Since $AA^\intercal = I$ implies that

$$\det(AA^\intercal) = \det(A)\det(A^\intercal) = (\det A)^2 = 1$$

then $|\det A| = 1$, or, equivalently, $A \in \mathrm{SL}_\pm(n, \mathbb{R})$.

We now particularize our discussion to 2×2 matrices. Consider

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

an arbitrary element of $O(2)$. Since $A^\intercal$ is an invertible matrix, its corresponding linear transformation $R_{A^\intercal} : \mathbb{R}^2 \to \mathbb{R}^2$ is an isomorphism, hence

$$\operatorname{span}\{(a, b), (c, d)\} = \operatorname{range} R_{A^\intercal} = \mathbb{R}^2,$$

from which it follows that the rows of the matrix A form a basis of $\mathbb{R}^2$.

Additionally, it is an orthonormal basis, because

$$AA^\intercal = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} \|(a,b)\|^2 & (a,b)\cdot(c,d) \\ (a,b)\cdot(c,d) & \|(c,d)\|^2 \end{bmatrix} = I.$$

The fact that $\|(a,b)\| = \|(c,d)\| = 1$ implies the existence of $\theta \in [0, 2\pi)$ such that $(a,b) = (\cos\theta, \sin\theta)$ and, since $(a,b) \perp (c,d)$, the vector (c,d) is either $(-\sin\theta, \cos\theta)$ or $(\sin\theta, -\cos\theta)$. We have proven that

$$O(2) = \left\{ \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} : \theta \in [0, 2\pi) \right\} \cup \left\{ \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} : \theta \in [0, 2\pi) \right\}.$$

Therefore, an abitrary matrix $A \in O(2)$ has norm equal to $\|A\| = 1$.

3 The statement of the theorem

We now deal exclusively with the product of random matrices, the linear cocycle (σ, A) described earlier and its associated Lyapunov exponent . We assume that the measure μ is such that the associated Lyapunov exponent γ exists, i.e. we assume that $\log^+ \|M\|$ is integrable.

Theorem 9 (Furstenberg)**.** *Let G_μ be the smallest closed subgroup which contains the support of μ. Assume that:*

i) G_μ is not compact.

ii) For every finite, non-empty $L \subseteq \mathbb{RP}^1$, there exists $M \in G_\mu$ such that $\bar{M}(L) \neq L$.

Then $\gamma > 0$.

The next two propositions establish equivalent conditions for the theorem, which are, in practice, simpler to check.

Proposition 10. *G_μ is compact if, and only if, there exists $C \in \mathrm{GL}(2,\mathbb{R})$ such that $CMC^{-1} \in O(2)$ for every $M \in G_\mu$.*

Proof. We begin by proving that

$$\boxed{\exists C \in \mathrm{GL}(2,\mathbb{R}), \forall M \in G_\mu : CMC^{-1} \in O(2) \implies G_\mu \text{ is compact.}}$$

Assume that the premise of the implication above holds and let A be an element of G_μ. Then $CAC^{-1} \in O(2)$ and there exists $R \in O(2)$ such that $A = C^{-1}RC$. Applying the norm to both sides of this equality yields

$$\|A\| = \left\|C^{-1}RC\right\| \leq \left\|C^{-1}\right\|\|R\|\|C\| = \left\|C^{-1}\right\|\|C\|.$$

Since the matrix C is fixed and our choice for A was arbitrary, this argument holds for all elements of G_μ. Consequently, G_μ is bounded and therefore it is compact.

We still have to prove the converse implication. To this end, suppose G_μ is compact. By Theorem 3, a probability measure h, known as Haar measure, exists on G_μ which is both left- and right-translation-invariant. Define the quadratic form $Q_0 : \mathbb{R}^2 \to \mathbb{R}$ by $Q_0(v) = v^\intercal I v$ and $Q : \mathbb{R}^2 \to \mathbb{R}$ by

$$Q(v) = \int_{G_\mu} Q_0(gv)\, dh(g) = \int_{G_\mu} \|gv\|^2\, dh(g) \geq 0$$

which is a positive quadratic form. There exists a positive semidefinite matrix

B such that $Q(v) = v^\intercal Bv$ for every v. So $B = C^\intercal C$ for some matrix C and

$$Q_0(Cv) = v^\intercal C^\intercal Cv = Q(v). \tag{6}$$

Note that C is invertible. Now let $T_{g_0} : G_\mu \to G_\mu$ be a translation map defined by $g \mapsto gg_0$, then

$$\begin{aligned} Q(v) &= \int_{G_\mu} Q_0(gv)\, dh(g) \\ &= \int_{T_{g_0}^{-1}(G_\mu)} Q_0(gv)\, dh(g) \\ &= \int_{G_\mu} Q_0(gg_0v)\, dh(g) \\ &= Q(g_0v). \end{aligned}$$

This, together with (6), yields

$$Q_0(Cg_0C^{-1}w) = Q_0(w)$$

which means $Cg_0C^{-1} \in O(2)$ as desired. $\square$

Before we prove the other equivalence we referred to, we need a technical lemma.

Lemma 11. *If $M \in \mathrm{SL}_\pm(2, \mathbb{R})$ fixes three directions then $M = \pm I$.*

Proof. Let $M \in \mathrm{SL}_\pm(2, \mathbb{R})$. Suppose there exist distinct $\bar{x}_1, \bar{x}_2, \bar{x}_3 \in \mathbb{RP}^1$ such that $\bar{M}(\bar{x}_i) = \bar{x}_i$ for $i = 1, 2, 3$. Equivalently, there exists $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ such that $Mx_i = \lambda_i x_i$. The matrix M has at most two linearly independent eigenvectors, thus, without loss of generality, suppose

$$x_3 = \alpha x_1 + \beta x_2$$

for some $\alpha, \beta \in \mathbb{R} \setminus \{0\}$. Then

$$Mx_3 = \alpha M x_1 + \beta M x_2 = \alpha \lambda_1 x_1 + \beta \lambda_2 x_2$$

and

$$\lambda_3 x_3 = \alpha \lambda_1 x_1 + \beta \lambda_2 x_2.$$

By linear independence $\lambda_1 = \lambda_2 = \lambda_3$. Additionally, $|\det M| = \lambda_1^2 = 1$ implies that $\lambda_1 \in \{-1, 1\}$ and therefore $M = \pm I$ as desired. □

We will also need the following proposition, which we state without proof.

Proposition 12. *Let $\varphi : G \to G'$ be a group homomorphism. The quotient group ${}^{G}/_{Ker\,\varphi}$ is isomorphic to Im φ.*

Proof. See [1]. □

Proposition 13. *Assume G_μ is not compact. Condition ii) in Theorem 9 is true iff for every set $L \subseteq \mathbb{RP}^1$ with $\#L \in \{1, 2\}$ there exists $M \in G_\mu$ such that $\bar{M}(L) \neq L$.*

Proof. The $\implies$ direction is trivial. We prove the converse. Suppose $L \subseteq \mathbb{RP}^1$ is finite, i.e. $L = \{\bar{x}_1, \ldots, \bar{x}_n\}$ and $\#L = n$. By hypothesis

$$\bar{M}(L) = \{\bar{M}(\bar{x}_1), \bar{M}(\bar{x}_2), \ldots, \bar{M}(\bar{x}_n)\} = L.$$

Since $\#M(L) = \#L$, each matrix $M \in G_\mu$ induces a permutation φ_M of L. This allows us to define a group homomorphism $\varphi : G_\mu \to \mathrm{Perm}(L)$ where $\mathrm{Perm}(L)$ denotes the group of all permutations of L. The group $\mathrm{Perm}(L)$ is finite and G_μ must be infinite since we are assuming it is non-compact. By Proposition 12, ${}^{G_\mu}\!/_{\mathrm{Ker}\ \varphi}$ is isomorphic to $\mathrm{Im}\,\varphi \subseteq \mathrm{Perm}(L)$ and therefore ${}^{G_\mu}\!/_{\mathrm{Ker}\ \varphi}$ is also finite. Since

$$G_\mu = \bigcup_{H \in {}^{G_\mu}\!/_{\mathrm{Ker}\ \varphi}} H \ker \varphi$$

is a finite union, each class in ${}^{G_\mu}\!/_{\mathrm{Ker}\ \varphi}$ must be infinite. Consequently,

$$\mathrm{Ker}\ \varphi = \{M \in G_\mu : \varphi(M) = I\} = \{M \in G_\mu : M(x_i) = x_i \text{ for } i = 1, \dots, n\}$$

is an infinite set. If $n \geq 3$ then, by Lemma 11, $\mathrm{Ker}\ \varphi$ is finite. This is a contradiction and therefore $n \in \{1, 2\}$. $\square$

3.1 An application to the random Fibonacci sequence

Example 3. Consider the random Fibonacci sequence

$$F_n = \begin{cases} F_{n-1} + F_{n-2}, & \text{with probability } p \\ F_{n-1} - F_{n-2}, & \text{with probability } 1-p \end{cases}$$

for $n \geq 2$ and $F_0 = 0, F_1 = 1$. The classical Fibonacci sequence occurs when $p = 1$, and in this case F_n grows exponentially. We would like to see how F_n

evolves when $0 < p < 1$. Notice that

$$\begin{bmatrix} F_{n+2} \\ F_{n+1} \end{bmatrix} = \begin{bmatrix} 1 & \pm 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}.$$

Define

$$A_+ = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \text{ and } A_- = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}.$$

We consider the probability space $(\Omega, \mathcal{F}, \mu)$ where $\Omega = \{A_+, A_-\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ and $\mu = p\delta_{A_+} + (1-p)\delta_{A_-}$ with $0 < p < 1$.

In this case, the product of random matrices cocyle will be over the product space $(\Omega^{\mathbb{N}}, \mathcal{F}^{\mathbb{N}}, \mu^{\mathbb{N}})$.

Let G_μ be the smallest closed group which contains $\mathrm{supp}(\mu) = \Omega$. We denote the classical Fibonnaci sequence by $(C_n)_{n \in \mathbb{N}}$. Then

$$A_+^n = \begin{bmatrix} C_n & C_{n-1} \\ C_{n-1} & C_{n-2} \end{bmatrix}$$

for $n \geq 2$. Given that C_n grows exponentially, we have $\left\| A_+^n \right\| \to \infty$ as $n \to \infty$ and this proves that G_μ is not compact. We next check if the second condition of Theorem 9 is satisfied.

If L is made up of only one direction, then the fact that A_- has no real eigenvalues shows that the condition is satisfied.

Suppose $L = \{\bar{x}_1, \bar{x}_2\}$ with $\bar{x}_1 \neq \bar{x}_2$ and $\bar{M}(L) = L$ for every $M \in G_\mu$. The matrix A_- cannot fix both directions since this would again imply it has real eigenvalues. The remaining case is

$$\bar{A}_{-}\bar{x}_1 = \bar{x}_2 \text{ and } \bar{A}_{-}\bar{x}_2 = \bar{x}_1$$

which implies $\bar{A}_{-}^2\bar{x}_i = \bar{x}_i$ for $i \in \{1, 2\}$. Since A_{-}^2 has complex eigenvalues, no such set L exists. By Proposition 13 and Theorem 9, the associated Lyapunov exponent γ is positive.

4 Proof of the theorem

In this section we fix μ to be a measure satisfying the assumptions of Theorem 9.

4.1 Properties of measures

Let $(X, \mathcal{X}, \kappa)$ be a measure space. The measure κ is said to be *atomic* if there exists $x \in X$ such that $\kappa(\{x\}) \neq 0$. A well known example of an atomic measure is the Dirac measure.

If X is a topological space, we denote the space of all the probability measures on $(X, \mathcal{B}(X))$ by the symbol $\mathcal{M}(X)$ endowed with the weak* topology. A detailed study of this topology is beyond the scope of this text, but we will make use of the fact that $\mathcal{M}(X)$ is a compact space if X is compact. This is the case when $X = \mathbb{RP}^1$. Furthermore, weak* convergence is equivalent to the usual weak convergence of measures, that is, (ρ_n) converges to ρ iff

$$\lim_{n\to\infty} \int f(x)\, d\rho_n(x) = \int f(x)\, d\rho(x)$$

for every bounded and continuous function f on X. In this case we write $\rho_n \Rightarrow \rho$. Further details and proofs of these results can be found in [10].

Lemma 14. *If $\nu \in \mathcal{M}(\mathbb{RP}^1)$ is non-atomic and $(A_n \neq 0)_{n\in\mathbb{N}}$ is a sequence of matrices converging to $A \neq 0$, then $A_{n*}\nu \Rightarrow A_*\nu$.*

Proof. Let $f : \mathbb{RP}^1 \to \mathbb{R}$ be a continuous function, then

$$\begin{aligned}
\lim_{n\to\infty} \int_{\mathbb{RP}^1} f(x)\, dA_{n*}\nu(x) &= \lim_{n\to\infty} \int_{\mathbb{RP}^1} f \circ A_n(x)\, d\nu(x) \\
&= \int_{\mathbb{RP}^1} \lim_{n\to\infty} f \circ A_n(x)\, d\nu(x) \\
&= \int_{\mathbb{RP}^1} f \circ A(x)\, d\nu(x) \\
&= \int_{\mathbb{RP}^1} f(x)\, dA_*\nu(x),
\end{aligned}$$

where we have used the dominated convergence theorem and the continuity of f. $\square$

Whenever convenient we may omit the domain of integration. In this case, the reader should assume that the domain is the whole sample space corresponding to the respective measure.

Lemma 15. *If $\nu \in \mathcal{M}(\mathbb{RP}^1)$ is non-atomic, then the set of matrices which preserve ν, i.e.*

$$H_\nu = \{M \in \mathrm{SL}_\pm(2, \mathbb{R}) : M_*\nu = \nu\},$$

is a compact subgroup of $\mathrm{SL}_\pm(2, \mathbb{R})$.

Proof. It is a simple exercise in algebra to see that H_ν is a group. We prove compactness. H_ν is compact iff it is closed and bounded. Let $(M_n \in H_\nu)_{n\in\mathbb{N}}$

be a sequence converging to $M \in \mathbb{R}^{2\times 2}$. For each n we have $\det M_n = \pm 1$ which implies $M_n \neq 0$. As for the matrix M, note that

$$\det(M) = \det\Big(\lim_{n\to\infty} M_n\Big) = \lim_{n\to\infty} \det(M_n) = \pm 1,$$

and therefore $M \neq 0$. We can apply Lemma 14 to obtain $M_{n*}\nu \Rightarrow M_*\nu$ i.e. $\nu \Rightarrow M_*\nu$ and $\nu = M_*\nu$ as desired. We have proven that H_ν is a closed set.

Suppose, in order to arrive at a contradiction, that H_ν is not bounded, so there exists a sequence $(M_n \in H_\nu)_{n\in\mathbb{N}}$ which diverges. Consider the new sequence (X_n) given by $X_n = M_n\|M_n\|^{-1}$. Since (X_n) is a sequence in a compact subspace of $\mathbb{R}^{2\times 2}$ it has a convergent subsequence (X_{n_k}) with limit C. Again, by Lemma 14, $X_{n_k *}\nu \Rightarrow C_*\nu$ such that $\nu = C_*\nu$. Now note that,

$$\det C = \det\left(\lim_{k\to\infty} \frac{M_{n_k}}{\|M_{n_k}\|}\right) = \lim_{k\to\infty} \frac{\pm 1}{\|M_{n_k}\|^2} = 0.$$

By the fundamental theorem of linear maps, $\operatorname{rank}(C) = 1$. Lemma 8 would then imply that ν is an atomic measure, contradicting our assumption.

$\square$

4.2 Stationary measures

Definition 16. Let $\nu \in \mathcal{M}(\mathbb{RP}^1)$. We define $\mu * \nu$ to be the measure on $\mathbb{RP}^1$ which satisfies

$$\int f(x)\, d(\mu * \nu)(x) = \iint f(\bar{M}x)\, d\mu(M) d\nu(x)$$

for any bounded Borel function $f : \mathbb{RP}^1 \to \mathbb{R}$. The measure ν is said to be *μ-stationary* if $\mu * \nu = \nu$.

We define the evaluation map by

$$\begin{aligned} \mathrm{ev} : \mathrm{SL}_{\pm}(2, \mathbb{R}) \times \mathbb{RP}^1 &\to \mathbb{RP}^1 \\ (M, \bar{v}) &\mapsto \bar{M}\bar{v}. \end{aligned}$$

Let $\nu \in \mathcal{M}(\mathbb{RP}^1)$. Notice that if $B \subseteq \mathbb{RP}^1$ is a measurable set, then

$$\begin{aligned} \mu * \nu(B) &= \iint \mathbb{1}_B(\mathrm{ev}(A, \bar{x}))\, d\mu(A) d\mu(\bar{x}) \\ &= \int \mathbb{1}_{\mathrm{ev}^{-1}(B)}(A, \bar{x})\, d(\mu \times \nu)(A, \bar{x}) \\ &= \mu \times \nu(\{(A, \bar{x}) : \bar{A}\bar{x} \in B\}) \\ &= (\mu \times \nu)(\mathrm{ev}^{-1}(B)) \\ &= \mathrm{ev}_*(\mu \times \nu)(B). \end{aligned}$$

Lemma 17. *Every μ-stationary $\nu \in \mathcal{M}(\mathbb{RP}^1)$ is non-atomic.*

Proof. Suppose, so as to obtain a contradiction, that ν is atomic. Then, the quantity

$$\beta = \max_{x \in \mathbb{RP}^1} \nu(\{x\})$$

is positive. Let $L = \{x \in \mathbb{RP}^1 : \nu(\{x\}) = \beta\}$. If L has infinite cardinality, then we may consider a countable subset $L_1 = \{x_1, x_2, \ldots\}$, but this contradicts the assumption that ν is a probability measure since

$$\begin{aligned}\nu(L_1) &= \nu(\{x_1, x_2, \dots\}) \\ &= \sum_{i=1}^{\infty} \beta \\ &= \infty.\end{aligned}$$

Consequently, L must be finite. Now let $x_0 \in L$ and note that

$$\beta = \nu(\{x_0\}) = \iint \mathbb{1}_{\{M^{-1}x_0\}}(x)\, d\nu(x) d\mu(M) = \int \nu(\{M^{-1}x_0\})\, d\mu(M) \leq \beta$$

By definition, the inequality $\beta \geq \nu(\{M^{-1}x_0\})$ is true for every M, so $\nu(\{M^{-1}x_0\}) = \beta$ and thus $M^{-1}x_0 \in L$ for μ-a.e. M, i.e. $M(L) = L$ for μ-a.e. M. This means that the set

$$F_L = \{M \in \mathrm{SL}_{\pm}(2, \mathbb{R}) : M(L) = L\}$$

has full measure, i.e. $\mu(F_L) = 1$. Furthermore, F_L is closed, so $\mathrm{supp}(\mu) \subseteq F_L$, which implies that $G_\mu \subseteq F_L$. This contradicts assumption ii) of Theorem 9. $\square$

Remark 18. It can be proven that μ-stationary measures always exist (see Lemma 3.5 of [3]). By Lemma 17, any such measure on $\mathbb{RP}^1$ is non-atomic.

4.3 Convergence of μ-stationary measures

Let $S_n = M_1 \cdots M_n$.

Lemma 19. *Let $\nu \in \mathcal{M}(\mathbb{RP}^1)$ be μ-stationary. For $\mu^{\mathbb{N}}$-a.e. $\omega \in \Omega$, there exists $\nu_\omega \in \mathcal{M}(\mathbb{RP}^1)$ such that*

$$S_n(\omega)_*\nu \Rightarrow \nu_\omega.$$

Proof. Let $f \in C(\mathbb{RP}^1)$. Define

$$\begin{aligned} F_f : \mathrm{SL}_\pm(2,\mathbb{R}) &\to \mathbb{R} \\ M &\mapsto \int f(Mx)\, d\nu(x). \end{aligned}$$

Let $\mathcal{F}_n$ be the σ-algebra of $\mathrm{SL}_\pm(2,\mathbb{R})^{\mathbb{N}}$ formed by the cylinders of length n. Then $S_n(\cdot)$ is $\mathcal{F}_n$-measurable. Let $C \in \mathcal{F}_n$, then

$$\int_C \int_{\mathrm{SL}_\pm(2,\mathbb{R})} F_f(S_n(\omega)M)\, d\mu(M) d\mu^{\mathbb{N}}(\omega) = \int_C F_f(S_{n+1}(\omega))\, d\mu^{\mathbb{N}}(\omega).$$

By definition of conditional expectation, we obtain

$$\begin{aligned} \mathbb{E}\left[F_f(S_{n+1}) \mid \mathcal{F}_n\right] &= \int_{\mathrm{SL}_\pm(2,\mathbb{R})} F_f(S_n M)\, d\mu(M) \\ &= \iint f(S_n M x)\, d\nu(x) d\mu(M) \\ &= \int f(S_n y)\, d\nu(y) \\ &= F_f(S_n). \end{aligned}$$

Therefore the stochastic process $\{F_f(S_n)\}_{n\in\mathbb{N}}$ is a bounded martingale and as such it converges almost surely, i.e. the limit

$$\Gamma f(\omega) = \lim_{n\to\infty} F_f(S_n(\omega))$$

exists for a.e. $\omega \in \Omega^{\mathbb{N}}$. We now use this fact to prove $S_n(\omega)_*\nu \Rightarrow \nu_\omega$ almost surely for some $\nu_\omega \in \mathcal{M}(\mathbb{RP}^1)$.

By the compactness of $\mathbb{RP}^1$, the space $C(\mathbb{RP}^1)$ is separable. Let $\{f_k\}_{k\in\mathbb{N}}$ be a dense subset of $C(\mathbb{RP}^1)$. The limit $\Gamma f_k(\omega)$ exists in a set $\mathcal{L}_k$ of full measure for each $k \in \mathbb{N}$. Let

$$\mathcal{L} = \bigcap_{k\in\mathbb{N}} \mathcal{L}_k$$

then

$$\mu^{\mathbb{N}}(\mathcal{L}^{\complement}) = \mu^{\mathbb{N}}\left(\bigcup_{n\in\mathbb{N}} \mathcal{L}_k^{\complement}\right) \leq \sum_{n=1}^{\infty} \mu^{\mathbb{N}}(\mathcal{L}_k^{\complement}) = 0$$

and we conclude that $\mu^{\mathbb{N}}(\mathcal{L}) = 1$. Now consider $\omega \in \mathcal{L}$ and let ν_ω be a weak* limit point of the sequence of measures $S_n(\omega)_*\nu$. Then

$$\begin{aligned}
\int f_k \, d\nu_\omega &= \lim_{n\to\infty} \int f_k \, dS_n(\omega)_*\nu \\
&= \lim_{n\to\infty} \int f_k \circ S_n(\omega) \, d\nu \\
&= \lim_{n\to\infty} F_{f_k}(S_n(\omega)) \, d\nu \\
&= \Gamma f_k(\omega).
\end{aligned}$$

Since the limit is the same for all subsequences then $S_n(\omega)_*\nu \Rightarrow \nu_\omega$. $\square$

Lemma 20. *The measures ν and ν_ω from Lemma 19 satisfy*

$$S_n(\omega)_* M_* \nu \Rightarrow \nu_\omega \quad \text{as } n \to \infty$$

for μ-a.e. M.

Proof. Let $\ell = \{f_1, f_2, \ldots\}$ be a countable dense subset of $C(\mathbb{RP}^1)$ and fix $k \in \mathbb{N}$. We will prove that the following quantity is finite:

$$I = \int \mathbb{E}^{\mu^{\mathbb{N}}} \left[\sum_{n=1}^{\infty} \left(\int f_k(S_n(\omega)Mx)\, d\nu(x) - \int f_k(S_n(\omega)x)\, d\nu(x) \right)^2 \right] d\mu(M).$$

Note that

$$\begin{aligned} I &= \int \sum_{n=1}^{\infty} \mathbb{E}^{\mu^{\mathbb{N}}} \left[\left(\int f_k(S_n(\omega)Mx)\, d\nu(x) - \int f_k(S_n(\omega)x)\, d\nu(x) \right)^2 \right] d\mu(M) \\ &= \sum_{n=1}^{\infty} \int \mathbb{E}^{\mu^{\mathbb{N}}} \left[\left(\int f_k(S_n(\omega)Mx)\, d\nu(x) - \int f_k(S_n(\omega)x)\, d\nu(x) \right)^2 \right] d\mu(M). \end{aligned}$$

Define

$$\begin{aligned} I_n &= \int \mathbb{E}^{\mu^{\mathbb{N}}} \left[\left(\int f_k(S_n(\omega)Mx)\, d\nu(x) - \int f_k(S_n(\omega)x)\, d\nu(x) \right)^2 \right] d\mu(M) \\ &= \int \mathbb{E}^{\mu^{\mathbb{N}}} \left[(F_{f_k}(S_n(\omega)M) - F_{f_k}(S_n(\omega)))^2 \right] d\mu(M) \\ &= \int \mathbb{E}^{\mu^{\mathbb{N}}} \left[(F_{f_k}(S_n(\omega)M))^2 + (F_{f_k}(S_n(\omega)))^2 - 2F_{f_k}(S_n(\omega)M) F_{f_k}(S_n(\omega)) \right] d\mu(M) \\ &= \mathbb{E}^{\mu^{\mathbb{N}}} \left[\int (F_{f_k}(S_n(\omega)M))^2 + (F_{f_k}(S_n(\omega)))^2 - 2F_{f_k}(S_n(\omega)M) F_{f_k}(S_n(\omega))\, d\mu(M) \right] \\ &= \mathbb{E}^{\mu^{\mathbb{N}}} \left[(F_{f_k}(S_{n+1}(\omega)))^2 + (F_{f_k}(S_n(\omega)))^2 - 2F_{f_k}(S_{n+1}(\omega)) F_{f_k}(S_n(\omega)) \right], \end{aligned}$$

where the last equality comes from the fact that

$$\iint F_{f_k}(S_n(\omega)M)^2\, d\mu(M) d\mu^{\mathbb{N}}(\omega) = \int F_{f_k}(S_{n+1}(\omega))^2\, d\mu^{\mathbb{N}}(\omega).$$

Furthermore,

$$\begin{aligned}\mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_{n+1}(\omega)F_{f_k}(S_n(\omega))\right] &= \mathbb{E}^{\mu^{\mathbb{N}}}\left[\mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_{n+1}(\omega))F_{f_k}(S_n(\omega)) \mid \mathcal{F}_n\right]\right] \\ &= \mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_n(\omega))\mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_{n+1}(\omega)) \mid \mathcal{F}_n\right]\right] \\ &= \mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_n(\omega))^2\right],\end{aligned}$$

where we have used the law of total expectation in the first equality. We have shown that

$$I_n = \mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_{n+1}(\omega))^2\right] - \mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_n(\omega))^2\right].$$

Therefore, we obtain a telescopic sum

$$\begin{aligned} I &= \lim_{N\to\infty}\sum_{n=1}^{N} I_n \\ &= \lim_{N\to\infty}\mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_{N+1}(\omega))^2\right] - \mathbb{E}^{\mu^{\mathbb{N}}}\left[F_{f_k}(S_1(\omega))^2\right] \\ &= \lim_{N\to\infty}\mathbb{E}^{\mu^{\mathbb{N}}}\left[\left(\int f_k(S_{N+1}(\omega)x)\, d\nu(x)\right)^2\right] - \mathbb{E}^{\mu^{\mathbb{N}}}\left[\left(\int f_k(S_1(\omega)x)\, d\nu(x)\right)^2\right] \\ &\leq \|f_k\|_{C^0}^2.\end{aligned}$$

So $I < \infty$ and the series

$$\sum_{n=1}^{\infty} \left(\int f_k(S_n(\omega)Mx)\, d\nu(x) - \int f_k(S_n(\omega)x)\, d\nu(x) \right)^2$$

is convergent for $\mu^{\mathbb{N}}$-a.e. ω and μ-a.e. M. So

$$\begin{aligned} \lim_{n\to\infty} \int f_k(S_n(\omega)Mx)\, d\nu(x) &= \lim_{n\to\infty} \int f_k(S_n(\omega)x)\, d\nu(x) \\ &= \int f_k(x)\, d\nu_\omega(x). \end{aligned}$$

Since the set ℓ is dense, the result holds for any continuous function, and we have proven the desired result. $\square$

We now show that the measures ν_ω above are necessarily Dirac measures.

Lemma 21. *For $\mu^{\mathbb{N}}$-a.e. ω, there exists $Z(\omega) \in \mathbb{RP}^1$ such that $\nu_\omega = \delta_{Z(\omega)}$*

Proof. Fix an $\omega \in \Omega^{\mathbb{N}}$ in the full-measure set for which

$$S_n(\omega)_*\nu \Rightarrow \nu_\omega \text{ and } S_n(\omega)_* M_* \nu \Rightarrow \nu_\omega$$

as $n \to \infty$ for μ-a.e. M. The sequence $X_n(\omega) = S_n(\omega)\|S_n(\omega)\|^{-1}$ has a convergent subsequence because it is defined on a compact subspace of $\mathbb{R}^{2\times 2}$. Suppose its limit is $X(\omega)$. As reasoned before,

$$\lim_{n\to\infty} \frac{S_n(\omega)}{\|S_n(\omega)\|} = X(\omega) \implies \|X(\omega)\| = 1$$

because of the continuity of the norm. Consequently, each $X_n(\omega)$ and $X(\omega)$ itself are non-zero matrices. By Lemma 17, ν is non-atomic, and thus we are in a position to apply Lemma 14 to conclude that

$$X(\omega)_*\nu = X(\omega)_* M_*\nu = \nu_\omega$$

for μ-a.e. M.

Suppose $X(\omega)$ is invertible. This would mean that $\nu = M_*\nu$ and thus X is an element of H_ν as defined in Lemma 15 for μ-a.e. M, therefore $G_\mu \subseteq H_\nu$. We are already assuming that G_μ is closed and have now concluded that it is a subspace of a compact space, which means it must be compact, contradicting assumption (i) of Theorem 9. In conclusion, $X(\omega)$ must not be invertible, from which follows $\operatorname{rank}(X(\omega)) = 1$. By Lemma 8 $X(\omega)_*\nu = \nu_\omega$ is a Dirac measure. □

4.4 Norm growth

We now prove that convergence to a Dirac measure tells us something about the norm growth of our product of matrices.

Lemma 22. *Let $m \in \mathcal{M}(\mathbb{RP}^1)$ be non-atomic and let (A_n) be a sequence in $\mathrm{SL}_\pm(2,\mathbb{R})$ such that $A_{n*}m \Rightarrow \delta_{\bar{z}}$, where $\bar{z} \in \mathbb{RP}^1$. Then $\|A_n\| \to \infty$. Moreover, for all $v \in \mathbb{R}^2$,*

$$\frac{\|A_n^\intercal v\|}{\|A_n^\intercal\|} \to |\langle v, z\rangle|.$$

Proof. Suppose $A_n\|A_n\|^{-1}$ converges to A. Lemma 14 implies that $\overline{A_n}_* m \Rightarrow \bar{A}_* m$, hence $\bar{A}_* m = \delta_{\bar{z}}$. If $\det A \neq 0$ then $m = \overline{A^{-1}}_*\delta_{\bar{z}}$ is Dirac. Contradiction. Hence $\det A = 0$. Now note that

$$0 = |\det A| = \lim_{n\to\infty} \left| \frac{\det A_n}{\|A_n\|^2} \right| = \lim_{n\to\infty} \frac{1}{\|A_n\|^2}$$

so $\lim_{n\to\infty} \|A_n\| = \infty$ as desired. Furthermore, the fact that $A \neq 0$ tells us that $\operatorname{rank}(A) = 1$ and thus $\operatorname{range}(A)$ is a line. Suppose $\operatorname{range}(A) = \operatorname{span}\{y\} = \bar{y}$ for some $y \in \mathbb{R}^2$, then

$$\bar{A}m(\{\bar{y}\}) = m(\overline{A}^{-1}(\{\bar{y}\})) = m(\mathbb{RP}^1) = 1 = \delta_{\bar{z}}(\{\bar{y}\})$$

and $\bar{z} = \bar{y}$. Now suppose $\|z\| = 1$ and let $\{e_1, e_2\}$ be the canonical basis of $\mathbb{R}^2$. Then

$$Ae_1 = \pm\|Ae_1\|z \text{ and } Ae_2 = \pm\|Ae_2\|z.$$

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then $\|Ae_1\|^2 + \|Ae_2\|^2 = a^2 + c^2 + b^2 + d^2$. The eigenvalues of $A^\intercal A$ are $\lambda_1 = 0$ and $\lambda_2 = a^2 + b^2 + c^2 + d^2$ which, together with the fact that $\|A\| = 1$, implies that $\lambda_2 = 1 = \|Ae_1\|^2 + \|Ae_2\|^2$. Now let v be a vector in $\mathbb{R}^2$, then

$$\begin{aligned}
\|A^\intercal v\|^2 &= \langle A^\intercal v, e_1\rangle^2 + \langle A^\intercal v, e_2\rangle^2 \\
&= \langle v, Ae_1\rangle^2 + \langle v, Ae_2\rangle^2 \\
&= (\|Ae_1\|^2 + \|Ae_2\|^2)\langle v, z\rangle^2 \\
&= \langle v, z\rangle^2.
\end{aligned}$$

Thus,

$$\lim_{n\to\infty} \frac{\|A_n^\intercal v\|^2}{\|A_n^\intercal\|^2} = \langle v, z\rangle^2.$$

$\square$

4.5 Proof of Theorem 9

Define $P_n = M_1^\intercal \cdots M_n^\intercal$. Let $\nu \in \mathcal{M}(\mathbb{RP}^1)$ be μ-stationary. By Lemma 19 there exists a measure $\nu_\omega \in \mathcal{M}(\mathbb{RP}^1)$ such that $P_n(\omega)_*\nu \Rightarrow \nu_\omega$ for $\mu^{\mathbb{N}}$-a.e. ω. Then, by Lemma 21 there exists a direction $\bar{Z}(\omega) \in \mathbb{RP}^1$ such that $\nu_\omega = \delta_{\bar{Z}(\omega)}$ for $\mu^{\mathbb{N}}$-a.e. ω. Using Lemma 22 we obtain that

$$\lim_{n\to\infty} \|P_n^\intercal(\omega)\| = \lim_{n\to\infty} \|P_n(\omega)\| = \infty \tag{7}$$

and

$$\frac{\|P_n^\intercal(\omega)v\|}{\|P_n(\omega)\|} \to |\langle v, Z(\omega)\rangle|. \tag{8}$$

for $\mu^{\mathbb{N}}$-a.e. ω and every $v \in \mathbb{R}^2$. Define

$$\begin{aligned} T : \mathrm{SL}_\pm(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^1 &\to \mathrm{SL}_\pm(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^1 \\ (\omega, \bar{x}) &\mapsto (\sigma(\omega), M_1(\omega)x) \end{aligned}$$

and

$$\begin{aligned} f : \mathrm{SL}_\pm(2,\mathbb{R})^{\mathbb{N}} \times \mathbb{RP}^1 &\to \mathbb{R} \\ (\omega, \bar{x}) &\mapsto \log \frac{\|M_1(\omega)x\|}{\|x\|}. \end{aligned}$$

Then

$$\sum_{j=0}^{n-1} f(T^j(\omega, \bar{x})) = \log \frac{\|M_n(\omega)\cdots M_1(\omega)x\|}{\|x\|} \to \infty$$

for $\mu^{\mathbb{N}}$-a.e. ω and $\bar{x}$ non-orthogonal to $\bar{Z}(\omega)$ by (7) and (8). Since ν is non-atomic, the convergence holds $\mu^{\mathbb{N}} \times \nu$ almost everywhere. For any $w \neq 0$

$$\begin{aligned}
\gamma &= \lim_{n\to\infty} \frac{1}{n} \log \left(\sup_{x\neq 0} \frac{\|P_n^\intercal x\|}{\|x\|} \right) \\
&\geq \iint \lim_{n\to\infty} \frac{1}{n} \log \left(\frac{\|M_n(\omega)\cdots M_1(\omega)w\|}{\|w\|} \right) d\mu^{\mathbb{N}}(\omega) d\nu(w) \\
&= \int \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f \circ T^j(\omega, w)\, d(\mu^{\mathbb{N}} \times \nu)(\omega, w) \\
&= \int f(\omega, w)\, d(\mu^{\mathbb{N}} \times \nu)(\omega, w)
\end{aligned}$$

by Birkhoff's ergodic theorem. Finally, by Proposition 7

$$\gamma \geq \int f(\omega, w)\, d(\mu^{\mathbb{N}} \times \nu)(\omega, w) > 0.$$

References

[1] M. Artin. *Algebra.* Pearson Education, 2011.

[2] Jario Bochi. Furstenberg's theorem on products of i.i.d. 2×2 matrices, 2016.

[3] Philippe Bougerol et al. *Products of random matrices with applications to Schrödinger operators*, volume 8. Springer Science & Business Media, 2012.

[4] Harry Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(3):377–428, 1963.

[5] Harry Furstenberg and Harry Kesten. Products of random matrices. *The Annals of Mathematical Statistics*, 31(2):457–469, 1960.

[6] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems.* Number 54. Cambridge university press, 1995.

[7] Halsey Royden and Patrick Fitzpatrick. *Real Analysis.* Pearson, 2010.

[8] J Michael Steele. Kingman's subadditive ergodic theorem. In *Annales de l'IHP Probabilités et statistiques*, volume 25, pages 93–98, 1989.

[9] Marcelo Viana. *Lectures on Lyapunov exponents*, volume 145. Cambridge University Press, 2014.

[10] Marcelo Viana and Krerley Oliveira. *Foundations of ergodic theory.* Number 151. Cambridge University Press, 2016.