LYAPUNOV EXPONENTS OF PROBABILITY DISTRIBUTIONS WITH NON-COMPACT SUPPORT

ADRIANA SÁNCHEZ AND MARCELO VIANA

ABSTRACT. A recent result of Bocker–Viana asserts that the Lyapunov exponents of compactly supported probability distributions in $GL(2, \mathbb{R})$ depend continuously on the distribution. We investigate the general, possibly non-compact case. We prove that the Lyapunov exponents are semi-continuous with respect to the Wasserstein topology, but not with respect to the weak* topology. Moreover, they are not continuous with respect to the Wasserstein topology.

1. INTRODUCTION

Let $M = SL(2, \mathbb{R})^{\mathbb{Z}}$ and $f : M \to M$ be the shift map over M defined by

$$(\alpha_n)_n \mapsto (\alpha_{n+1})_n.$$

Consider the function

$$A : M \to SL(2, \mathbb{R}), \quad (\alpha_n)_n \mapsto \alpha_0,$$

and define its n-th iterate by

$$A^n((\alpha_k)_k) = \alpha_{n-1} \cdots \alpha_0.$$

Given any probability measure p in $SL(2, \mathbb{R})$, the associated Bernoulli measure $\mu = p^{\mathbb{Z}}$ on M is invariant under f. Let $L^1(\mu)$ denote the space of μ-integrable functions on M. Assuming that $\log^+ \|A^{\pm}\| \in L^1(\mu)$, it follows from Furstenberg-Kesten [5, Theorem 2] that

$$\lambda_+(x) = \lim_n \frac{1}{n} \log \|A^n(x)\| \quad \text{and} \quad \lambda_-(x) = \lim_n \frac{1}{n} \log \|A^{-n}(x)\|^{-1},$$

exist for μ-almost every $x \in M$. Furthermore, the functions $x \mapsto \lambda_{\pm}(x)$ are constant on the orbits of f. Thus, by the ergodicity of the Bernoulli measure, they are constant on a full μ-measure set. We call these constants the *Lyapunov exponents* of μ, and we represent them as $\lambda_{\pm}(p)$.

The way Lyapunov exponents depend on the corresponding probability distribution p has been studied by several authors. Furstenberg, Kifer [6] proved that this dependence is continuous at every quasi-irreducible point, that is, such that there is at most one subspace invariant under every matrix in the support of p.

Also under an irreducibility condition, LePage [7] proved Hölder continuity of the Lyapunov exponents.

More recently, Bocker and Viana [2] proved that continuity actually holds at every point, in the realm of compactly supported probability distributions. Backes, Butler and Brown [1] have much extended this result, to linear cocycles with invariant holonomies, over hyperbolic systems with local product structure on compact spaces. Moreover, still in the compactly supported case, Tall and Viana [10] proved that Hölder continuity holds whenever the exponents $\lambda_-(p)$ and $\lambda_+(p)$ are distinct. In general, when the exponents coincide, one has a weaker but still quantitative, modulus of continuity, called log-Hölder continuity.

For much more on this topic, see Viana [9, Chapter 10], Duarte, Klein [4], Viana [8] and the references therein.

The aim of this note is to study the (semi)continuity of the Lyapunov exponents in the general case, when the support of the probability measure p is not necessarily compact. Our main result reads as follows (Theorem 3.1 and Theorem 3.2 provide more detailed statements):

Theorem A. *The function $p \mapsto \lambda_+(p)$ is upper semi-continuous for the Wasserstein topology but not for the weak* topology. The same remains valid for $p \mapsto \lambda_-(p)$ with lower semi-continuity.*

Regarding continuity of Lyapunov exponents we prove (see Theorem 4.1 for a more detailed statement):

Theorem B. *The functions $p \mapsto \lambda_\pm(p)$ are not continuous for the Wasserstein topology.*

2. Wasserstein Topology

Let (M, d) be a Polish space, that is a complete separable metric space. The *Wasserstein space* is the space of probability measures on M with finite moments of order 1, that is,

$$P_1(M) = \left\{ \mu \in P(M) : \int_M d(x_0, x) d\mu(x) < +\infty \right\},$$

where $x_0 \in M$ is arbitrary and $P(M)$ denotes the space of Borel probability measures on M. This space does not depend on the choice of the point x_0.

We use $\mu_k \xrightarrow{\ *\ } \mu$ to mean convergence in the weak* topology. Moreover, $\mu_k \xrightarrow{\ W\ } \mu$ will mean convergence in the Wasserstein topology, characterized by the following definition (see Definition 6.8 in Villani [11]):

Definition 2.1. *Let (M, d) be a Polish space and $(\mu_k)_{k \in \mathbb{N}}$ be a sequence in $P_1(M)$. Given any $\mu \in P_1(M)$, we say that $\mu_k \xrightarrow{\ W\ } \mu$ if one of the following equivalent conditions is satisfied for some (and then any) $x_0 \in M$:*

(1) *$\mu_k \xrightarrow{\ *\ } \mu$ and $\int d(x_0, x) d\mu_k(x) \to \int d(x_0, x) d\mu(x)$;*
(2) *$\mu_k \xrightarrow{\ *\ } \mu$ and*

$$\limsup_{k \to \infty} \int d(x_0, x) d\mu_k(x) \leq \int d(x_0, x) d\mu(x);$$

(3) $\mu_k \xrightarrow{*} \mu$ and
$$\lim_{R \to \infty} \limsup_{k \to \infty} \int_{d(x_0,x) \geq R} d(x_0, x) d\mu_k(x) = 0;$$

(4) *For every continuous function φ with $|\varphi(x)| \leq C(1 + d(x_0, x))$, $C \in \mathbb{R}$, one has*
$$\int \varphi(x) d\mu_k(x) \to \int \varphi(x) d\mu(x).$$

Let (M, μ) and (N, ν) be two probability spaces. A *coupling* of μ and ν is a measure π on $M \times N$ such that π projects to μ and ν on the first and second coordinate, respectively. When $\mu = \nu$ we call π a *self-coupling*.

If (M, d) is a Polish space, the *Wasserstein distance* between two probability measures μ and ν on M is defined by

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_M d(x, y) d\pi(x, y), \tag{1}$$

where the infimum is taken over the set $\Pi(\mu, \nu)$ of all the couplings of μ and ν. This distance metrizes the Wasserstein topology (see [11, Theorem 6.18]):

Theorem 2.1. *Let (M, d) be a Polish space. Then the Wasserstein distance W_1 metrizes the Wasserstein topology in the space $P_1(M)$:*

$$\mu_k \xrightarrow{W} \mu \text{ if and only if } W_1(\mu_k, \mu) \to 0.$$

Moreover, with this metric $P_1(M)$ is also a complete separable metric space and, every $\mu \in P_1(M)$ can be approximated by probability measures with finite support.

This also implies that W_1 is continuous on $P_1(M)$.

3. SEMICONTINUITY

It is a well-known fact that when the measures have compact support, the Lyapunov exponents are semicontinuous with the weak* topology (see for example [9, Chapter 9]). However, in the non compact setting this is no longer true. If they were semicontinuous then every measure with vanishing Lyapunov exponent would be a point of continuity. The next theorem shows that this is not the case.

Theorem 3.1. *There exist a measure q and a sequence of measures $(q_n)_n$ on $SL(2, \mathbb{R})$ converging to q in the weak* topology, such that $\lambda_+(q_n) \geq 1$ for n large enough but $\lambda_+(q) = 0$.*

Proof. Define a function $\alpha : \mathbb{N} \to SL(2, \mathbb{R})$ by

$$\alpha(2k - 1) = \begin{pmatrix} \sigma_k & 0 \\ 0 & \sigma_k^{-1} \end{pmatrix} \quad \text{and} \quad \alpha(2k) = \begin{pmatrix} \sigma_k^{-1} & 0 \\ 0 & \sigma_k \end{pmatrix}$$

where $(\sigma_k)_k$ is an increasing sequence such that $\sigma_1 > 1$ and $\sigma_k \to +\infty$.

Let $\mu = q^{\mathbb{Z}}$ be a measure in M where q is the measure on $SL(2, \mathbb{R})$ given by

$$q = \sum_{k \in \mathbb{N}} p_k \delta_{\alpha(k)},$$

with $\sum p_k = 1$, $0 < p_k < 1$ for all $k \in \mathbb{N}$.

The key idea to construct this example is to find p_k and σ_k such that $\log \|A\| \in L^1(\mu)$ and satisfying the hypothesis above. Hence, consider $0 < r < 1/2 < s < 1$,

and $l = s/r > 1$. Let us take $\sigma_k = e^{l^k}$ for all k, which is an increasing sequence provided that $l > 1$.

For $k \geq 2$ take $p_{2k-1} = p_{2k} = r^k$. Since $0 < r < 1/2$ it is easy to see that

$$\sum_{k \geq 3} p_k = 2 \sum_{k \geq 2} p_{2k} = 2 \sum_{k \geq 2} r^k = 2 \frac{r^2}{1 - r} < 1$$

We have to choose p_1 and p_2 such that $\sum p_k = 1$. Then, it is enough to take

$$p_1 = p_2 = \frac{1}{2} \left(1 - 2 \frac{r^2}{1 - r} \right).$$

We continue by showing that $\log \|A\| \in L^1(\mu)$. This is an easy computation,

$$\int_M \log \|A\| d\mu = 2 p_2 \log \sigma_1 + 2 \sum_{k \geq 2} p_{2k} \log \sigma_k = 2 p_2 l + 2 \sum_{k \geq 2} s^k.$$

Since $0 < s < 1$ this geometric series is convergent. More over, since $p_{2k-1} = p_{2k}$ for all k then $\lambda_+(q) = 0$.

What is left is to construct the sequence q_n. Fix $n_0 > 1$ large enough so $\frac{1}{2} \left(1 - 2 \frac{r^2}{1-r} \right) > l^{-n}$ for all $n \geq n_0$, and consider $q_n = \sum_k q_k^n \delta_{\alpha(k)}$ where for $n \geq n_0$

$$\begin{aligned} q_{2n}^n &= l^{-n} + r^n, \\ q_2^n &= \frac{1}{2} \left(1 - \frac{r^2}{1 - r} \right) - l^{-n} \\ q_k^n &= p_k \text{ other case.} \end{aligned}$$

Thus, since $q_k^n \to p_k$ when $n \to \infty$ for all k, it is easy to see that q_n converges in the weak* topology to q.

The proof is completed by showing that $\lambda_+(q_n) \geq 1$ for n large enough. It follows easily since,

$$\lambda_+(q_n) = |q_{2n-1}^n - q_{2n}^n| \log \sigma_n + |q_1^n - q_2^n| \log \sigma_1 = l^{-n} l^n + l^{-n+1},$$

which is equal to $1 + l^{-n+1} \geq 1$ for all $n \geq n_0$.   □

We now consider the Wasserstein topology in $P_1(SL(2, \mathbb{R}))$ which is stronger than the weak* topology, as stated in Definition 2.1. The advantage of using this topology is that all probability measures in $P_1(SL(2, \mathbb{R}))$ have finite moment of order 1, and so their Lyapunov exponents always exist. This observation is a direct consequence of the fact that $\log : [1, \infty) \to \mathbb{R}$ is a 1-Lipschitz function and $\|\alpha\| \geq 1$ for every matrix $\alpha \in SL(2, \mathbb{R})$, because

$$\int \log \|A(x)\| d\mu = \int \log \|\alpha\| dp \leq \int d(\alpha, \mathrm{id}) dp < \infty.$$

The convergence of the moments of order 1, allow us to control the weight of integrals outside compact sets, and prove semicontinuity of the Lyapunov exponents in $P_1(SL(2, \mathbb{R}))$. We are thus led to the following result:

Theorem 3.2. *The function defined on $P_1(SL(2, \mathbb{R}))$ by $p \to \lambda_+(p)$ is upper semicontinuous. The same remains valid for the function $p \to \lambda_-(p)$ with lower semicontinuity.*

Before beginning the proof of Theorem 3.2 we need to recalled some important results regarding the relationship between Lyapunov exponents and stationary measures.

A probability measure η on $\mathbb{P}^1$ is called a p-*stationary* if

$$\eta(E) = \int \eta(\alpha^{-1}E)dp(\alpha),$$

for every measurable set $E \in \mathbb{P}^1$ and $\alpha^{-1}E = \{[\alpha^{-1}v] : [v] \in E\}$.

Roughly speaking, the following result shows that the set of stationary measures for a measure p is close for the weak* topology.

Proposition 3.3. *Let $(p_k)_k$ be probability measures in $SL(2,\mathbb{R})$ converging to p in the weak* topology. For each k, let η_k be p_k-stationary measures and η_k converges to η in the weak* topology. Then η is a stationary measure for p.*

Furthermore, in our context it is well-known that

$$\lambda_+(p) = \max\left\{\int \Phi dp \times \eta : \eta \ \ p-\text{stationary}\right\},$$

where $\Phi : SL(2,\mathbb{R}) \times \mathbb{P}^1 \to \mathbb{R}$ is given by

$$\Phi(\alpha, [v]) = \log\frac{\|\alpha v\|}{\|v\|}.$$

For more details see for example [9, Proposition 6.7].

We now proceed to the proof of Theorem 3.2.

Proof of Theorem 3.2. We will prove that $\lambda_+(p)$ is upper semi-continuous. The case of $\lambda_-(p)$ is analogous.

Let $(p_k)_k$ be a sequence in the Wasserstein space $P_1(M)$ converging to p, i.e. $W(p_k, p) \to 0$. For each $k \in \mathbb{N}$ let η_k a stationary measure that realizes the maximum in the identity above. That is:

$$\lambda_+(p_k) = \int \Phi dp_k d\eta_k.$$

Since $\mathbb{P}^1$ is compact, passing to a subsequence if necessary, we can suppose η_k converges in the weak* topology to a measure η which, as established in Proposition 3.3, is a p-stationary measure.

Let $\epsilon > 0$, we want to prove that there exist a constant $k_0 \in \mathbb{N}$ such that for each $k > k_0$

$$\left|\int \Phi dp_k d\eta_k - \int \Phi dp d\eta\right| < \epsilon.$$

In order to do this we need to consider some properties of the Wasserstein topology. First of all, since the first moment of p is finite there exist K_1 a compact set of $SL(2,\mathbb{R})$ such that

$$\int_{K_1^c} d(\alpha, \text{id})dp < \frac{\epsilon}{36}. \tag{2}$$

Moreover, according to Definition 2.1, since $p_k \xrightarrow{W} p$ there exist $R' > 0$ satisfying

$$\limsup_k \int_{d(\alpha,\text{id})>R'} d(\alpha, \text{id})dp_k < \frac{\epsilon}{36},$$

then, there exist $k' > 0$ such that for every $k > k'$

$$\int_{d(\alpha, \mathrm{id}) > R'} d(\alpha, \mathrm{id}) dp_k < \frac{\epsilon}{36}. \tag{3}$$

Take $R > 0$ big enough so $B(\mathrm{id}, R') \cup K_1 \subset B(\mathrm{id}, R)$ and define the compact set $K = \bar{B}(\mathrm{id}, R)$. Since the function $\log : [1, \infty) \to \mathbb{R}$ is 1-Lipschitz and $\|\alpha\| \geq 1$ for all $\alpha \in SL(2, \mathbb{R})$, then

$$|\Phi(\alpha, [v])| = \left| \log \frac{\|\alpha v\|}{\|v\|} \right| \leq \log \|\alpha\| \leq |\|\alpha\| - \|\mathrm{id}\|| \leq d(\alpha, \mathrm{id}). \tag{4}$$

Our proof starts with the observation that

$$\left| \int \Phi dp_k d\eta_k - \int \Phi dp d\eta \right|$$

$$\leq \left| \int_{K \times \mathbb{P}^1} \Phi dp_k d\eta_k - \int_{K \times \mathbb{P}^1} \Phi dp d\eta \right| + \left| \int_{K^c \times \mathbb{P}^1} \Phi dp_k d\eta_k \right| + \left| \int_{K^c \times \mathbb{P}^1} \Phi dp d\eta \right|.$$

From (3) it follows that

$$\left| \int_{K^c \times \mathbb{P}^1} \Phi dp_k d\eta_k \right| \leq \int_{K^c} d(\alpha, \mathrm{id}) dp_k < \frac{\epsilon}{3}. \tag{5}$$

Furthermore, (2) implies that

$$\left| \int_{K^c \times \mathbb{P}^1} \Phi dp d\eta \right| \leq \int_{K^c} d(\alpha, \mathrm{id}) dp < \frac{\epsilon}{3}. \tag{6}$$

We now proceed to analyze the integral:

$$\left| \int_{K \times \mathbb{P}^1} \Phi dp_k d\eta_k - \int_{K \times \mathbb{P}^1} \Phi dp d\eta \right|$$

$$\leq \left| \int_{K \times \mathbb{P}^1} \Phi dp_k d\eta_k - \int_{K \times \mathbb{P}^1} \Phi dp_k d\eta \right| + \left| \int_{K \times \mathbb{P}^1} \Phi dp_k d\eta - \int_{K \times \mathbb{P}^1} \Phi dp d\eta \right|.$$

Consider $\Phi_K = \Phi|_{K \times \mathbb{P}^1}$ the restriction of Φ to the compact space $K \times \mathbb{P}^1$. Thus, Φ_K is uniformly continuous with the product metric. Hence, there exist $\delta = \delta(\epsilon)$ such that for every $[v] \in \mathbb{P}^1$ and every $\alpha, \beta \in K$ satisfying $d(\alpha, \beta) < \delta$ we have

$$|\Phi_K(\alpha, [v]) - \Phi_K(\beta, [v])| < \frac{\epsilon}{18}.$$

Moreover, by the compactness of the set K we can find $\alpha_1, ..., \alpha_N \in K$ such that $K \subset \cup_{i=1}^N B(\alpha_i, \delta)$. Therefore, the convergence of $(\eta_k)_k$ to η in the weak* topology implies that for each $i = 1, ..., N$ there exist $k_i > 0$ such that

$$\left| \int_{\mathbb{P}^1} \Phi_K(\alpha_i, [v]) d\eta_k - \int_{\mathbb{P}^1} \Phi_K(\alpha_i, [v]) d\eta \right| < \frac{\epsilon}{18}.$$

Take $k'' = \max\{k_1, ..., k_N\}$. From the above it follows that given $\alpha \in K$ there exist i such that $d(\alpha, \alpha_i) < \delta$ and for every $k > k''$ if $\Phi_i = \Phi_K(\alpha_i, [v])$ then

$$\left| \int_{\mathbb{P}^1} \Phi_K d\eta_k - \int_{\mathbb{P}^1} \Phi_K d\eta \right| \leq$$

$$\leq \int_{\mathbb{P}^1} |\Phi_K - \Phi_i| \, d\eta_k + \left| \int_{\mathbb{P}^1} \Phi_i d\eta_k - \int_{\mathbb{P}^1} \Phi_i d\eta \right| + \int_{\mathbb{P}^1} |\Phi_i - \Phi_K| \, d\eta < \frac{\epsilon}{6}.$$

Since this convergence is uniform in α, this implies that

$$\left| \iint_{K^c \times \mathbb{P}^1} \Phi_K(\alpha, [v]) d\eta_k dp_k - \int_{K^c \times \mathbb{P}^1} \Phi_K(\alpha, [v]) d\eta dp_k \right| < \frac{\epsilon}{6}, \qquad (7)$$

for all $k > k''$.

Now define $L = SL(2, \mathbb{R}) \backslash B(id, R + 1)$ and, consider the Urysohn function f_0 : $SL(2, \mathbb{R}) \to [0, 1]$ given by

$$f_0(\alpha) = \frac{d(\alpha, L)}{d(\alpha, L) + d(\alpha, K)}.$$

It is easily seen that f_0 is continuous, equal to zero in L and equal to 1 in K. Therefore, the functions

$$\psi(\alpha) = \int_{\mathbb{P}^1} \Phi(\alpha, [v]) d\eta \cdot f_0(\alpha)$$

are continuous. Then since $|\psi(\alpha)| \le d(\alpha, id)$ and, by the definition of the Wasserstein topology, there exist k''' such that for every $k > k'''$

$$\left| \int \psi dp_k - \int \psi dp \right| < \frac{\epsilon}{18}.$$

Moreover, since $\psi - \varphi = 0$ in K and

$$|\psi - \varphi| \le \log \|\alpha\| |f_0(\alpha) - \chi_K(\alpha)| \le 2d(\alpha, id).$$

Hence, by (2) and (3)

$$\int |\psi - \varphi| dp \le 2 \int_{K^c} d(\alpha, id) dp < \frac{\epsilon}{18},$$

$$\int |\psi - \varphi| dp_k \le 2 \int_{K^c} d(\alpha, id) dp_k < \frac{\epsilon}{18}$$

for each $k > k'$. Thus, if $k > \max\{k', k'''\}$ we get

$$\left| \iint_{K \times \mathbb{P}^1} \Phi dp_k d\eta - \int_{K \times \mathbb{P}^1} \Phi dp d\eta \right| < \frac{\epsilon}{6}. \qquad (8)$$

Finally, taking $k_0 = \max\{k', k'', k'''\}$, we conclude that for every $k > k_0$

$$\left| \int \Phi dp_k d\eta_k - \int \Phi dp d\eta \right| < \epsilon.$$

We just proved that

$$\lambda_+(p_k) = \int \Phi dp_k d\eta_k \to \int \Phi dp d\eta \le \lambda_+(p),$$

which concludes our proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 3.4. *Theorem 3.1 does not contradicts Theorem 3.2, since p is not in $P_1(SL(2, \mathbb{R}))$. To see this, take $x_0 = $ id and note that*

$$\int d(x, x_0) dp = \sum_{k=0}^{\infty} p_k \|\alpha_k - \mathrm{id}\| = 2 \sum_{k=0}^{\infty} r^k (e^{l^k} - 1),$$

which diverges.

4. Proof of Theorem B

At this section we are going to describe a construction of points of discontinuity of the Lyapunov exponents as functions of the probability measure, relative to the Wasserstein topology.

Theorem 4.1. *There exist a measure q and a sequence of measures $(q_n)_n$ on $SL(2, \mathbb{R})$ converging to q in the Wasserstein topology, such that $\lambda_+(q_n) = 0$ for all $n \in \mathbb{N}$ but $\lambda_+(q) > 0$.*

Proof. Consider the function $\alpha : \mathbb{N} \to SL(2, \mathbb{R})$ defined by the hyperbolic matrices

$$\alpha(k) = \begin{pmatrix} k & 0 \\ 0 & k^{-1} \end{pmatrix}$$

Take $m \in \mathbb{N}$ the smallest natural number bigger than 1 such that $\sum_{n \geq m} e^{-\sqrt{n}} < 1$, which exist since $\sum_k e^{-\sqrt{k}}$ is convergent, and define

$$p_k = e^{-\sqrt{k}}, \text{ if } k \geq m,$$

$$p_1 = 1 - \sum_{n \geq m} e^{-\sqrt{n}},$$

$$p_k = 0, \text{ otherwise.}$$

It is obvious from the definition that $\sum_k p_k = 1$. Hence, we define the probability measure $q = \sum p_k \delta_{\alpha(k)}$. We need to see that $q \in P_1(SL(2, \mathbb{R}))$, in order to do so notice that if $x_0 = \mathrm{id}$

$$\int d(x, x_0) dq = \sum_k p_k \| \alpha(k) - \mathrm{id} \| = \sum_k e^{-\sqrt{k}} (k - 1)$$

which is convergent by the Cauchy condensation test. Moreover, since

$$\sum_k e^{-\sqrt{k}} \log k < \sum_k k e^{-\sqrt{k}}$$

then if $\mu = q^{\mathbb{Z}}$ we have $\log \|A\| \in L^1(\mu)$ and,

$$\lambda_+(q) = \sum_k e^{-\sqrt{k}} \log k > 0.$$

Now, consider $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and for each n consider

$$\beta_n(k) = \begin{cases} \alpha(k) & \text{if } k \neq n, \\ B & \text{if } k = n. \end{cases}$$

With this we define the probability measures $q_n = \sum_k p_k \delta_{\beta_n(k)}$. In a similar way as above, we can see that for all n these measures belong to $P_1(SL(2, \mathbb{R}))$ and, $\log \|A\| \in L^1(q_n)$. We proceed to show that q_n converges to q in the Wasserstein topology. This follows since

$$W(q_n, q) \leq p_n d(\alpha(n), B) \sim n e^{-\sqrt{n}}$$

which goes to 0 if n goes to ∞.

It remains to prove that $\lambda_+(q_n) = 0$ for every n. In order to do this we proceed by contradiction. Suppose there exist N such that $\lambda_+(q_N) > 0 > \lambda_-(q_N)$. We will consider the distance in the projective space $\mathbb{P}^1$ given by

$$\delta([v], [w]) := \frac{\|v \wedge w\|}{\|v\|\|w\|} = \sin(\angle(v, w)).$$

Consider the family $\mathcal{F} = \{V, H\}$, where $V = [e_2]$ and $H = [e_1]$ are the vertical and horizontal axis respectively. By the definition of the measure q_N, it is clear that this family is invariant by every matrix in the support of this measure. Moreover, if $x = (x_k)_k \in M$ then we can see that for every m

$$\delta(A^m(x)H, A^m(x)V) \geq \delta(H, V) = 1. \tag{9}$$

On the other hand, we have for every unit vectors v and w

$$\|A^m(x)v \wedge A^m(x)w\| \leq \|\wedge^2 A^m(x)\|.$$

It is widely known that, for q_N-almost every $x \in M$

$$\lambda_+(q_N) + \lambda_-(q_N) = \lim_m \frac{1}{m} \log \|\wedge^2 A^m(x)\|.$$

Note that q_N is irreducible, this means that there is no proper subspace of $\mathbb{R}^2$ invariant under all the matrices in the support of q_N. Therefore, we have

$$\lambda_+(q_N) = \lim_m \frac{1}{m} \log \|A^m(x)w\|,$$

for every unit vector w. For a more detailed discussion of the two results mention above we refer the reader to [3, Chapter III].

Thus, we have for every unit vectors v and w

$$\lim_m \frac{1}{m} \log \frac{\|\wedge^2 A^m(x)\|}{\|A^m(x)v\|\|A^m(x)w\|} = \lambda_-(q_N) - \lambda_+(q_N) < 0,$$

and hence

$$\lim_m \delta(A^m(x)H, A^m(x)V) \leq \lim_m \frac{\|\wedge^2 A^m(x)\|}{\|A^m(x)e_1\|\|A^m(x)e_2\|}$$
$$= \lim_m \exp\left(m \cdot \frac{1}{m} \log \frac{\|\wedge^2 A^m(x)\|}{\|A^m(x)e_1\|\|A^m(x)e_2\|}\right) = 0,$$

which is a contradiction with (9) and, we finish our proof.

$\square$

Notice that this example shows that the Wasserstein topology is not enough to guarantee continuity of the Lyapunov exponents. The main problem is that the support of the measures q_n move further apart from the support of q. Thus, this suggest that we need to add some hypothesis guaranteeing the "convergence" of the supports. An assumption of this type was made by Bocker, Viana in [2] in order to prove the continuity for measures with compact support. However in the non compact setting we will also need to control the iteratives of the cocycles. Those hypotheses are very restrictive so we are left with the following conjecture:

Conjecture 4.1. *The Lyapunov exponent functions $\lambda_\pm : P_1(SL(2, \mathbb{R})) \to \mathbb{R}$ are not continuous with respect to the Wasserstein–Hausdorff topology.*

In the next two sections we are going to describe a construction of points of discontinuity of the Lyapunov exponents as functions of the measure, relative to the Wasserstein topology. However, in each of them the support of the measures are arbitrarily close. These constructions were inspired by the discontinuity example presented by Bocker Viana in [2, Section 7.1], and do not prove our conjecture since they are not examples in $SL(2, \mathbb{R})$.

4.1. Discontinuity example in $SL(2, \mathbb{R})^5$.

Let us recall that $M = (SL(2, \mathbb{R}))^{\mathbb{Z}}$, $f : M \to M$ is the shift map over M defined by

$$(\alpha_n)_n \mapsto (\alpha_{n+1})_n.$$

And the linear cocycle A is the product of random matrices which is defined by

$$A : M \to SL(2, \mathbb{R}), \quad (\alpha_n)_n \mapsto \alpha_0.$$

Given an invariant measure p in $SL(2, \mathbb{R})$ we can define $\mu = p^{\mathbb{Z}}$ which is an invariant measure in M.

Now consider $X = SL(2, \mathbb{R})^5$ with the product metric

$$d_\infty((\alpha_1, ..., \alpha_5), (\beta_1, ..., \beta_5)) = \max\{d(\alpha_1, \beta_1), ..., d(\alpha_5, \beta_5)\}.$$

Let $N = X^{\mathbb{Z}}$ be the space of sequences over X and $g : N \to N$ the shift map over N. We can identify N with M using the function $\iota : M \to N$ by $\iota((\alpha_n)_n) = (\beta_n)_n$ where

$$\beta_n = (\alpha_{5n}, \alpha_{5n+1}, \alpha_{5n+2}, \alpha_{5n+3}, \alpha_{5n+4})_n.$$

It is easy to see that ι defines a bijection between N and M. Moreover, we have the following identity

$$g(\iota((\alpha_n)_n)) = f^5((\alpha_n)_n).$$

Also we can consider the linear cocycle induced by A in N, that is the function $B : N \to SL(2, \mathbb{R})$ given by

$$B((\iota((\alpha_n)_n)) = A^5((\alpha_n)_n).$$

So in this context we have the following result.

Theorem 4.2. *There exist a measure q and a sequence of measures $(q_n)_n$ on X converging to q in the Wasserstein topology, such that*

$$\lambda_+(B, q_n) \nrightarrow \lambda_+(B, q).$$

The main idea of the proof is to construct a measure on N whose Lyapunov exponents are positive and approximate it, in the Wasserstein topology, by measures with zero Lyapunov exponents. In order to do that, define the function $\alpha : \mathbb{N} \to SL(2, \mathbb{R})$ as

$$\alpha(2k - 1) = \begin{pmatrix} k & 0 \\ 0 & k^{-1} \end{pmatrix} \quad \text{and} \quad \alpha(2k) = \begin{pmatrix} k^{-1} & 0 \\ 0 & k \end{pmatrix}$$

As in the example before take $m \in \mathbb{N}$ the smallest natural (odd) number bigger than 3 such that $\sum_{k \geq m} e^{-\sqrt{k}} < 1$, which exist since $\sum_k e^{-\sqrt{k}}$ is convergent, and

define

$$p_{2k} = p_{2k-1} = \frac{1}{2}e^{-\sqrt{k}}, \text{ if } 2k-1 \geq m,$$
$$p_3 = 1 - \sum_{n \geq m} e^{-\sqrt{n}},$$
$$p_k = 0, \text{ otherwise.}$$

Let $\mu = \tilde{q}^{\mathbb{Z}}$ be a measure in M where $\tilde{q}$ is the measure on $SL(2, \mathbb{R})$ given by

$$\tilde{q} = \sum_{k \in \mathbb{N}} p_k \delta_{\alpha(k)}.$$

Let us consider the space $\Omega = \mathbb{N}^5$ and define the measure on X by

$$q = \sum_{w \in \Omega} p_w \delta_{\alpha(w)},$$

where $\alpha(w) = (\alpha(w_1), \cdots, \alpha(w_5))$ and, $p_w = p_{w_1} \cdots p_{w_5}$ if $w = (w_1, ..., w_5)$.

Now consider the measure $\nu = q^{\mathbb{Z}}$ on N. First, we need to ensure that the measure q belong to $P_1(X)$. This is a direct consequence of the fact that $\sum e^{-\sqrt{n}}(n-1)$ is convergent equal to some positive constant c. Indeed, if $\alpha_0 = (\text{id}, ..., \text{id})$ and the notation $p_1 \cdots \hat{p}_i \cdots p_5$ denotes the product of p_1 through p_5 except p_i then

$$\int d_\infty(\alpha, \alpha_0)dq = \sum_w p_w d_\infty(\alpha(w), \alpha_0)$$
$$< \sum_{i=1}^5 \sum_{w_j, j \neq i} p_{w_1} \cdots \hat{p}_{w_i} \cdots p_{w_5} \left(\sum_{w_i} p_{w_i} d(\alpha(w_i), \text{id}) \right)$$
$$< c \sum_{i=1}^5 \sum_{w_j, j \neq i} p_{w_1} \cdots \hat{p}_{w_i} \cdots p_{w_5} = 5c$$

which proves our claim. Remember that this also guarantees the existence of $\lambda_\pm(B, q)$ as mention in Section 3.

It is easy to see that $\nu = \iota_* \mu$. Using this we have

$$\lambda_+(B, q) = \lim_n \frac{1}{n} \int_N \log \|B^n(x)\| d\nu = \lim_n \frac{1}{n} \int_M \log \|A^{5n}(x)\| d\mu$$
$$= 5\lambda_+(A, \tilde{q}) = 5p_3 \log 2 > 0.$$

The task is now to construct the sequence $(q_n)_n$. In order to do this, for each $n \in \mathbb{N}$ consider $w_n = (2n, 2n+2, 2n+1, 2n-1, 2n-1)$ and define

$$\beta(w_n) = (\alpha(2n)R_\epsilon, \alpha(2n+2), \alpha(2n+1)R_\delta, \alpha(2n-1), \alpha(2n-1)R_\epsilon),$$

where $\epsilon = n^{-1}(n+1)^{-1}$, $\delta = \arctan(\epsilon)$ and,

$$R_\epsilon = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix}, \quad R_\delta = \begin{pmatrix} \cos(\delta) & -\sin(\delta) \\ \sin(\delta) & \cos(\delta) \end{pmatrix}.$$

We proceed to define the sequence by

$$q_n = \sum_{w \neq w_n} p_w \delta_{\alpha(w)} + p_{w_n} \delta_{\beta(w_n)}.$$

We claim that $W(q_n, q) \to 0$ if n goes to infinite. Our proof starts with the observation that

$$\pi_n = \sum_{w \neq w_n} p_w \delta_{(\alpha(w), \alpha(w))} + p_{w_n} \delta_{(\alpha(w_n), \beta(w_n))}$$

is a coupling of q and q_n. Then,

$$W(q_n, q) \leq \int d_\infty(u, v) \pi_k(u, v) = p_{w_n} d(\alpha(w_n), \beta(w_n))$$

and the latter is bounded from above by

$$\max\{\|\alpha(2n)(1 - R_\epsilon)\|, \|\alpha(2n-1)(1 - R_\delta)\|, \|\alpha(2n+1)(1 - R_\epsilon)\|\} \leq \epsilon(n+1) = \frac{1}{n}$$

which proves our claim.

What is left is to show that $\lambda_+(B, q_n) = 0$ for all n. For this we need the following lema.

Lemma 4.3. *Let $H_x = \mathbb{R}(1, 0)$ and $V_x = \mathbb{R}(0, 1)$. If $Z_n = [0 : \beta(w_n)]$ then, for all $x \in Z_n$ we have $B(x)H_x = V_{g(x)}$ and $B(x)V_x = H_{g(x)}$*

Proof. Notice that for any $x \in Z_n$

$$B(x) = \begin{pmatrix} 0 & -\epsilon^{-2}\sin(\delta) \\ \epsilon^2 \sin(\delta) + \epsilon \cos(\delta) & 0 \end{pmatrix}.$$

Which completes the proof. $\square$

The rest of the proof follows the same arguments as the Bocker-Viana example. For more details we refer the reader to ([2, Section 7.1]).

4.2. **Discontinuity example in $GL(2, \mathbb{R})^2$.** Let $M = (GL(2, \mathbb{R}))^{\mathbb{Z}}$ let $f : M \to M$ be the shift map over M and $A : M \to GL(2, \mathbb{R})$ the product of random matrices. Now consider $X = GL(2, \mathbb{R})^2$ with the maximum norm, and let $N = X^{\mathbb{Z}}$ be the space of sequences over X and $g : N \to N$ the shift map over N. As before, we can identify N with M using the function $\iota : M \to N$ defined by $\iota((\alpha_n)_n) = (\beta_n)_n$ where $\beta_n = (\alpha_{2n}, \alpha_{2n+1})$ which is a bijection between N and M.

With the above definition we can see that

$$g(\iota((\alpha_n)_n)) = f^2((\alpha_n)_n)$$

and defined $B : N \to GL(2, \mathbb{R})$ the linear cocycle induced by A in N through $B(\iota((\alpha_n)_n)) = A^2((\alpha_n)_n)$. In a similar way as in the previous example, there exist a measure p and a sequence of measures $(p_k)_k$ on X converging to p in the Wasserstein topology, such that $\lambda_+(A, p_k) \not\to \lambda_+(A, p)$.

Indeed, let $\alpha : \mathbb{N} \to GL(2, \mathbb{R})$ be defined by

$$\alpha(2k - 1) = \begin{pmatrix} k & 0 \\ 0 & k^{-2} \end{pmatrix} \quad \text{and} \quad \alpha(2k) = \begin{pmatrix} k^{-2} & 0 \\ 0 & k \end{pmatrix}.$$

Take the weights p_k as in the previous section and let $\tilde{q} = \sum_{k \in \mathbb{N}} p_k \delta_{\alpha(k)}$, . Consider the space $\Omega = \mathbb{N}^2$ and define the measure on X by $q = \sum_{w \in \Omega} p_w \delta_{\alpha(w)}$, where $\alpha(w) = (\alpha(w_1), \alpha(w_2))$ and, $p_w = p_{w_1} p_{w_2}$ if $w = (w_1, w_2)$. Let $\nu = q^{\mathbb{Z}}$ a measure on N.

Analysis similar to that in Section 4.1 shows that $q \in P_1(X)$, and using that $\nu = i_*\mu$ we have

$$\lambda_+(B, q) = \lim_n \frac{1}{n} \int_N \log \|B^n(x)\| d\nu = \lim_n \frac{1}{n} \int_M \log \|A^{2n}(x)\| d\mu$$
$$= 2\lambda_+(A, \tilde{q}) = 2p_3 \log 2 > 0.$$

For each $n \in \mathbb{N}$ consider $w_n = (2n, 2n - 1)$ and define $\beta(w_n) = (\beta(2n), \beta(2n - 1))$, by

$$\beta(2n) = \begin{pmatrix} 1 & -\delta \\ 0 & 1 \end{pmatrix} \alpha(2n) \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix} = \begin{pmatrix} 0 & -n\delta \\ \epsilon n & n \end{pmatrix},$$

$$\beta(2n - 1) = \begin{pmatrix} 1 & 0 \\ \epsilon & 1 \end{pmatrix} \alpha(2n - 1) = \begin{pmatrix} n & 0 \\ \epsilon n & n^{-1} \end{pmatrix},$$

where $\delta = n^{-(1+\gamma)}$ with $0 < \gamma < 1$, $\epsilon = n^{-3}\delta^{-1} = n^{\gamma-2}$.

We proceed to define the sequence by

$$q_n = \sum_{w \neq w_n} p_w \delta_{\alpha(w)} + p_{w_n} \delta_{\beta(w_n)}.$$

To prove that $W(q_n, q) \to 0$ if n goes to infinite we consider the diagonal coupling of q_n and q

$$\pi_n = \sum_{w \neq w_n} p_w \delta_{(\alpha(w), \alpha(w))} + p_{w_n} \delta_{(\alpha(w_n), \beta(w_n))}$$

Hence, we have

$$W(q_n, q) \leq \int d_\infty(u, v)\pi_n(u, v) = p_{w_n} d_\infty(\alpha(w_n), \beta(w_n))$$
$$< \max\{\|\beta(2n) - \alpha(2n)\|, \|\beta(2n - 1) - \alpha(2n - 1)\|\}$$
$$\leq \max\{\epsilon\sigma_n, n^{-2} + n\delta\} = \max\{n^{\gamma-1}, n^{-2} + n^{-\gamma}\} \leq 2n^{-l}$$

where $l = \min\{\gamma, 1 - \gamma\} > 0$, which proves our claim.

The rest of the proof, that is proving that $\lambda_+(B, q_n) = 0$ for all n, runs as before by noticing that for any $x \in Z_n = [0 : \beta(w_n)]$

$$B(x) = \begin{pmatrix} 0 & -n^2\delta \\ \epsilon n^{-1} & 0 \end{pmatrix}.$$

Indeed, this guarantees that $B(x)H_x = V_{g(x)}$ and $B(x)V_x = H_{g(x)}$ where $H_x = \mathbb{R}(1, 0)$ and $V_x = \mathbb{R}(0, 1)$. Finally, applying the first return map argument in [2, Section 7.1], we conclude our proof.

Acknowledgements

References

[1] Lucas Backes, Aaron W Brown, and Clark Butler. Continuity of lyapunov exponents for cocycles with invariant holonomies. *J. Mod. Dyn.*, 12:223–260, 2018.

[2] Carlos Bocker-Neto and Marcelo Viana. Continuity of lyapunov exponents for random two-dimensional matrices. *Ergodic Theory and Dynamical Systems*, 37(5):1413–1442, 2017.

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

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

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

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

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

[8] M. Viana. (Dis)continuity of Lyapunov exponents. *Ergod. Theory Dynam. Systems*.

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

[10] Marcelo Viana and El Hadji Y. Tall. Moduli of continuity of lyapunov exponents for random $gl(2)$ cocycles. *final preparation*, 2018.

[11] Cédric Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.

ICMC-USP – Av. Trab. São-carlense 400, São Carlos, 13566-590 São Paulo, Brazil.
Email address: asanchez@icmc.usp.br

IMPA – Estrada D. Castorina 110, Jardim Botânico, 22460-320 Rio de Janeiro, Brazil.
Email address: viana@impa.br