

FURSTENBERG'S THEOREM ON PRODUCTS OF I.I.D. 2×2 MATRICES

JAIRO BOCHI

ABSTRACT. This is a revised version of some notes written > 10 years ago. I thank Anthony Quas for pointing to a gap in the previous proof of Lemma 5 and for providing a correct proof.

These notes follow [BL].

We deal with Lyapunov exponents of products of random i.i.d. 2×2 matrices of determinant ± 1 . Let $\mathrm{SL}_{\pm}(2, \mathbb{R})$ denote the group of such matrices.

Let μ be a probability measure in $\mathrm{SL}_{\pm}(2, \mathbb{R})$ which satisfies the integrability condition¹

$$\int_{\mathrm{SL}_{\pm}(2, \mathbb{R})} \log \|M\| d\mu(M) < \infty.$$

If Y_1, Y_2, \dots are random independent matrices with distribution μ , then the limit

$$\gamma = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|Y_n \cdots Y_1\|$$

(the upper Lyapunov exponent) exists a.s. and is constant, by the subadditive ergodic theorem. We have $\gamma \geq 0$.

The Furstenberg theorem says that $\gamma > 0$ for “most” choices of μ . Let us see some examples where $\gamma = 0$:

- (1) If μ is supported in the orthogonal group $\mathrm{O}(2)$ then $\gamma = 0$.
- (2) If μ is supported in the abelian subgroup

$$\left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix}; t \in \mathbb{R} \setminus \{0\} \right\}$$

then $\gamma = \int \log \|M\| d\mu(M)$, which may be zero.

- (3) Assume that only two matrices occur:

$$\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix} \quad \text{and} \quad R_{\pi/2} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Then it is a simple exercise to show that $\gamma = 0$.

Furstenberg's theorem says that the list above essentially covers all possibilities where the exponent vanishes:

Date: May 9, 2016.

¹Note that $\|M\| = \|M^{-1}\| \geq 1$ if $M \in \mathrm{SL}_{\pm}(2, \mathbb{R})$.

Theorem. *Let μ be as above, and let G_μ be the smallest closed subgroup which contains the support of μ . Assume that:*

- (i) G_μ is non compact.
- (ii) *There is no finite set $\emptyset \neq L \subset \mathbb{P}^1$ such that $M(L) = L$ for all $M \in G_\mu$.*

Then $\gamma > 0$.

Remark. Condition (i) is equivalent to:

- (i') There is no $C \in \text{GL}(2, \mathbb{R})$ such that CMC^{-1} is an orthogonal matrix, for all $M \in G_\mu$.

Remark. Under the assumption (i), condition (ii) is equivalent to:

- (ii') There is no set $L \subset \mathbb{P}^1$ with $\#L = 1$ or 2 and such that $M(L) = L$ for all $M \in G_\mu$.

(This follows from the fact that if $M \in \text{SL}_\pm(2, \mathbb{R})$ fixes three different directions then $M = I$.)

NON-ATOMIC MEASURES IN \mathbb{P}^1

Let $\mathcal{M}(\mathbb{P}^1)$ be the space of probability Borel measures in \mathbb{P}^1 . A measure $\nu \in \mathcal{M}(\mathbb{P}^1)$ is called *non-atomic* if $\nu(\{x\}) = 0$ for all $x \in \mathbb{P}^1$.

We collect some simple facts for later use.

If $A \in \text{GL}(2, \mathbb{R})$ then we also denote by A the induced map $A: \mathbb{P}^1 \rightarrow \mathbb{P}^1$. If A is not invertible but $A \neq 0$ then there is only one direction $x \in \mathbb{P}^1$ for which Ax is not defined. In this case, it makes sense to consider the push-forward $A\nu \in \mathcal{M}(\mathbb{P}^1)$, if $\nu \in \mathcal{M}(\mathbb{P}^1)$ is non-atomic.

Lemma 1. *If $\nu \in \mathcal{M}(\mathbb{P}^1)$ is non-atomic and A_n is a sequence of non-zero matrices converging to $A \neq 0$, then $A_n\nu \rightarrow A\nu$ (weakly).*

The proof is easy.

Lemma 2. *If $\nu \in \mathcal{M}(\mathbb{P}^1)$ is non-atomic then*

$$H_\nu = \{M \in \text{SL}_\pm(2, \mathbb{R}); M\nu = \nu\}$$

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

Proof. Assume that there exists a sequence M_n in H_ν with $\|M_n\| \rightarrow \infty$. Up to taking a subsequence, we may assume that the sequence (of norm 1 matrices) $\|M_n\|^{-1}M_n$ converges to a matrix C . Since $C \neq 0$, Lemma 1 gives $C\nu = \nu$. On the other hand,

$$\det C = \lim \frac{1}{\|M_n\|^2} = 0.$$

Thus C has rank one and $\nu = C\nu$ must be a Dirac measure, contradiction. \square

μ -STATIONARY MEASURES IN \mathbb{P}^1

If $\nu \in \mathcal{M}(\mathbb{P}^1)$, let the *convolution* $\mu * \nu \in \mathcal{M}(\mathbb{P}^1)$ is the push-forward of the measure $\mu \times \nu$ by the natural map $\text{ev}: \text{SL}_{\pm}(2, \mathbb{R}) \times \mathbb{P}^1 \rightarrow \mathbb{P}^1$. If $\mu * \nu = \nu$ then ν is called μ -stationary. By a Krylov–Bogolioubov argument, μ -stationary measures always exist.

Lemma 3. *If μ satisfies the assumptions of Furstenberg's theorem then every μ -stationary $\nu \in \mathcal{M}(\mathbb{P}^1)$ is non-atomic.*

Proof. Assume that

$$\beta = \max_{x \in \mathbb{P}^1} \nu(\{x\}) > 0.$$

Let $L = \{x \in \mathbb{P}^1; \nu(\{x\}) = \beta\}$. If $x_0 \in L$ then

$$\begin{aligned} \beta = \nu(\{x_0\}) &= (\mu * \nu)(\{x_0\}) = \iint \chi_{\{x_0\}}(Mx) d\mu(M) d\nu(x) \\ &= \int \nu(\{M^{-1}(x_0)\}) d\mu(M). \end{aligned}$$

But $\nu(\{M^{-1}(x_0)\}) \leq \beta$ for all M , so $\nu(\{M^{-1}(x_0)\}) \leq \beta$ for μ -a.e. M . We have proved that $M^{-1}(L) \subset L$ for μ -a.e. M . This contradicts assumption (ii). \square

From now on we assume that μ satisfies the assumptions of Furstenberg's theorem, and that ν is a (non-atomic) μ -stationary measure in \mathbb{P}^1 .

ν AND γ

The shift $\sigma: \text{SL}_{\pm}(2, \mathbb{R})^{\mathbb{N}} \leftarrow$ in the space of sequences $\omega = (Y_1, Y_2, \dots)$ has the ergodic invariant measure $\mu^{\mathbb{N}}$.

Consider the skew-product map $T: \text{SL}_{\pm}(2, \mathbb{R})^{\mathbb{N}} \times \mathbb{P}^1 \leftarrow$, $T(\omega, x) = (\sigma(\omega), Y_1(\omega)x)$. Consider $f: \text{SL}_{\pm}(2, \mathbb{R})^{\mathbb{N}} \times \mathbb{P}^1 \rightarrow \mathbb{R}$ given by

$$f(\omega, x) = \log \frac{\|Y_1(\omega)x\|}{\|x\|}.$$

(The notation is obvious). Then

$$\frac{1}{n} \sum_{j=0}^n f(T^j(\omega, x)) = \frac{1}{n} \log \frac{\|Y_n(\omega) \cdots Y_1(\omega)x\|}{\|x\|}.$$

by Oseledets' theorem, for a.e. ω and for all $x \in \mathbb{P}^1 \setminus \{E^-(x)\}$,² the quantity on the right hand side tends to γ as $n \rightarrow \infty$. In particular, this convergence holds for $\mu^{\mathbb{N}} \times \nu$ -a.e. (ω, x) . We conclude that

$$(1) \quad \gamma = \iint f d\mu^{\mathbb{N}} d\nu = \iint \log \frac{\|Mx\|}{\|x\|} d\mu(M) d\nu(x).$$

² $E^-(x)$ is the direction associated to the exponent $-\gamma$, if $\gamma > 0$.

CONVERGENCE OF PUSH-FORWARD MEASURES

Let $S_n(\omega) = Y_1(\omega) \cdots Y_n(\omega)$.

Lemma 4. *For $\mu^{\mathbb{N}}$ -a.e. ω , there exists $\nu_\omega \in \mathcal{M}(\mathbb{P}^1)$ such that*

$$S_n(\omega)\nu \rightarrow \nu_\omega.$$

Proof. Fix $f \in C(\mathbb{P}^1)$. Associate to f the function $F: \mathrm{SL}_\pm(2, \mathbb{R}) \rightarrow \mathbb{R}$ given by

$$F(M) = \int f(Mx) d\nu(x).$$

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. Also

$$\begin{aligned} \mathbb{E}(F(S_{n+1}) \mid \mathcal{F}_n) &= \int F(S_n M) d\mu(M) \\ &= \iint f(S_n Mx) d\mu(M) d\nu(x) \\ &= \int f(S_n y) d\nu(y) = F(S_n) \quad (\text{since } \mu * \nu = \mu). \end{aligned}$$

This shows that the sequence of functions $\omega \mapsto F(S_n(\omega))$ is a bounded martingale. Therefore the limit

$$\Gamma f(\omega) = \lim_{n \rightarrow \infty} F(S_n(\omega))$$

exists for a.e. ω .

Now let f_k ; $k \in \mathbb{N}$ be a countable dense subset of $C(\mathbb{P}^1)$. Take ω in the full-measure set where $\Gamma f_k(\omega)$ exists for all k . Let ν_ω be a (weak) limit point of the sequence of measures $S_n(\omega)\nu$. Then

$$\int f_k d\nu_\omega = \lim_{n \rightarrow \infty} \int f_k d(S_n\nu) = \lim_{n \rightarrow \infty} \int f_k \circ S_n d\nu = \Gamma f_k(\omega).$$

Since the limit is the same for all subsequences, we have in fact that $S_n(\omega)\nu \rightarrow \nu_\omega$. \square

Let's explore the construction of the measures to obtain more information about them:

Lemma 5. *The measures ν_ω from Lemma 4 satisfy*

$$S_n(\omega)M\nu \rightarrow \nu_\omega \quad \text{as } n \rightarrow \infty \text{ for } \mu\text{-a.e. } M.$$

Proof. We show that for any fixed f_k from the sequence above that for μ -a.e. M , that $\int f_k(S_n(\omega)Mx) d\nu(x) \rightarrow \int f_k d\nu_\omega(x)$ for $\mu^{\mathbb{N}}$ almost every ω . Let $F_k: \mathrm{SL}_\pm(2, \mathbb{R}) \rightarrow \mathbb{R}$ be the function introduced above corresponding to f_k . Given this, by taking the intersection over countably many sets of full μ -measure, we obtain a set of M 's of full measure on which $\int f(S_n(\omega)Mx) d\nu(x) \rightarrow \int f d\nu_\omega(x)$ for $\mu^{\mathbb{N}}$ -a.e. ω for all $f \in C(\mathbb{P}^1)$, that is, weak convergence of $S_n(\omega)M\nu$ to ν_ω .

We consider the expression

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

where \mathbb{E} denotes integration in the ω variable with respect to $\mu^{\mathbb{N}}$.

We establish below that $I < \infty$. From this it follows that the quantity in the brackets is finite for $\mu^{\mathbb{N}}$ -a.e. ω and μ -a.e. M . It follows that for $\mu^{\mathbb{N}}$ -a.e. ω and μ -a.e. M , $\int f_k(S_n(\omega)Mx) d\nu(x) - \int f_k(S_n(\omega)x) d\nu(x) \rightarrow 0$. However for $\mu^{\mathbb{N}}$ -a.e. ω , we have $\int f_k(S_n(\omega)x) d\nu(x) \rightarrow \int f_k(x) d\nu_{\omega}(x)$. Combining these, we obtain the desired result.

To prove that $I < \infty$, we note that

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

Now define I_n by

$$\begin{aligned} I_n &= \int \mathbb{E} \left(\int f_k(S_n(\omega)Mx) d\nu(x) - \int f_k(S_n(\omega)x) d\nu(x) \right)^2 d\mu(M) \\ &= \int \mathbb{E} \left(F_k(S_n(\omega)M) - F_k(S_n(\omega)) \right)^2 d\mu(M) \\ &= \int \mathbb{E} \left(F_k(S_n(\omega)M)^2 - 2F_k(S_n(\omega)M)F_k(S_n(\omega)) + F_k(S_n(\omega))^2 \right) d\mu(M). \end{aligned}$$

Notice that the distribution of conditional distribution of $S_{n+1}(\omega)$ given $S_n(\omega)$ is the same as that of $S_n(\omega)M$. Hence

$$I_n = \mathbb{E} \left(F_k(S_{n+1}(\omega))^2 - 2F_k(S_{n+1}(\omega))F_k(S_n(\omega)) + (F_k(S_n(\omega)))^2 \right).$$

Notice

$$\begin{aligned} \mathbb{E} \left(F_k(S_{n+1}(\omega))F_k(S_n(\omega)) \right) &= \mathbb{E} \left(\mathbb{E} \left(F_k(S_{n+1}(\omega))F_k(S_n(\omega)) \middle| \mathcal{F}_n \right) \right) \\ &= \mathbb{E} \left(F_k(S_n(\omega))\mathbb{E}(F_k(S_{n+1}(\omega)) | \mathcal{F}_n) \right) \\ &= \mathbb{E}(F_k(S_n(\omega))^2), \end{aligned}$$

where we used the tower law for conditional expectations for the second equality. Hence $I_n = \mathbb{E}F_k(S_{n+1}(\omega))^2 - \mathbb{E}F_k(S_n(\omega))^2$. Now

$$I = \lim_{N \rightarrow \infty} \sum_{n=1}^N I_n = \lim_{N \rightarrow \infty} \mathbb{E}F_k(S_{N+1}(\omega))^2 - \mathbb{E}F_k(S_1(\omega))^2 \leq \|f_k\|^2. \quad \square$$

THE LIMIT MEASURES ARE DIRAC

Lemma 6. *For $\mu^{\mathbb{N}}$ -a.e. ω , there exists $Z(\omega) \in \mathbb{P}^1$ such that $\nu_{\omega} = \delta_{Z(\omega)}$.*

Proof. Fix a $\mu^{\mathbb{N}}$ -generic ω . By Lemma 5 we have, for μ -a.e. M ,

$$\lim S_n \nu = \lim S_n M \nu.$$

Let B be a limit point of the sequence of norm 1 matrices $\|S_n\|^{-1}S_n$. Since $\|B\| = 1$, we can apply Lemma 1:

$$B\nu = BM\nu.$$

If B were invertible, this would imply $\nu = M\nu$. That is, a.e. M belongs to the compact group H_ν (see Lemma 2) and therefore $G_\nu \subset H_\nu$, contradicting hypothesis (i). So B is non-invertible. Since $B\nu = \nu_\omega$, we conclude that ν_ω is Dirac. \square

CONVERGENCE TO DIRAC IMPLIES NORM GROWTH

Lemma 7. *Let $m \in \mathcal{M}(\mathbb{P}^1)$ be non-atomic and let (A_n) be a sequence in $\mathrm{SL}_\pm(2, \mathbb{R})$ such that $A_n m \rightarrow \delta_z$, where $z \in \mathbb{P}^1$. Then*

$$\|A_n\| \rightarrow \infty.$$

Moreover, for all $v \in \mathbb{R}^2$,

$$\frac{\|A_n^*(v)\|}{\|A_n\|} \rightarrow |\langle v, z \rangle|.$$

Proof. We may assume that the sequence $A_n/\|A_n\|$ converges to some B . Since $\|B\| = 1$, we can apply Lemma 1 to conclude that $Bm = \delta_z$. If B were invertible then we would have that $m = \delta_{B^{-1}z}$ would be atomic. Therefore $\det B = 0$ and

$$\frac{1}{\|A_n\|^2} = \left| \det \frac{A_n}{\|A_n\|} \right| \rightarrow |\det B| = 0.$$

So $\|A_n\| \rightarrow \infty$.

Notice that the range of B must be the z direction.

Let v_n, u_n be unit vectors such that $A_n v_n = \|A_n\| u_n$. Then

$$u_n = \frac{A_n(v_n)}{\|A_n\|}.$$

Since $A_n/\|A_n\| \rightarrow B$ and $\|B\| = 1$, we must have $u_n \rightarrow z$ (up to changing signs). Moreover, u_n is the direction which is most expanded by A_n^* . The assertion follows. (For a more elegant proof, see [BL, p. 25].) \square

CONVERGENCE TO ∞ CANNOT BE SLOWER THAN EXPONENTIAL

We shall use the following abstract lemma from ergodic theory:

Lemma 8. *Let $T: (X, m) \leftrightarrow$ be a measure preserving transformation of a probability space (X, m) . If $f \in L^1(m)$ is such that*

$$\sum_{j=0}^{n-1} f(T^j x) = +\infty \quad \text{for } m\text{-almost every } x,$$

then $\int f d\mu > 0$.

Proof. ³ For any function g , let \tilde{g} denote the limit of Birkhoff averages of g . Then $\tilde{f} \geq 0$. Assume, for a contradiction, that $\int f = 0$. Then $\tilde{f} = 0$ a.e.

Let $s_n = \sum_{j=0}^{n-1} f \circ T^j$. For $\varepsilon > 0$, let

$$A_\varepsilon = \{x \in X; s_n(x) \geq \varepsilon \forall n \geq 1\} \quad \text{and} \quad B_\varepsilon = \bigcup_{k \geq 0} T^{-k}(A_\varepsilon).$$

Fix $\varepsilon > 0$ and let $x \in B_\varepsilon$. Let $k = k(x) \geq 0$ be the least integer such that $T^k x \in A_\varepsilon$. We compare the Birkhoff sums of f and χ_{A_ε} :

$$\sum_{j=0}^{n-1} f(T^j x) \geq \sum_{j=0}^{k-1} f(T^j x) + \sum_{j=k}^{n-1} \varepsilon \chi_{A_\varepsilon}(T^j x) \quad \forall n \geq 1.$$

Dividing by n and making $n \rightarrow \infty$ we get

$$0 = \tilde{f}(x) \geq \varepsilon \widetilde{\chi_{A_\varepsilon}}(x)$$

Therefore

$$\mu(A_\varepsilon) = \int \widetilde{\chi_{A_\varepsilon}} = \int_{B_\varepsilon} \widetilde{\chi_{A_\varepsilon}} = 0.$$

Thus $\mu(B_\varepsilon) = 0$ for every $\varepsilon > 0$ as well.

On the other hand, if $s_n(x) \rightarrow \infty$ then $x \in \bigcup_{\varepsilon > 0} B_\varepsilon$. We have obtained a contradiction. \square

End of the proof of the theorem. Replace everywhere Y_i by Y_i^* . Note that μ^* also satisfies the hypothesis of the theorem if μ does.⁴

Let T and f be as in page 3. By Lemmas 6 and 7 we have

$$\sum_{j=0}^n f(T^j(\omega, x)) = \log \frac{\|S_n^*(\omega)x\|}{\|x\|} \rightarrow \infty$$

for a.e. ω and all $x \in \mathbb{P}^1 \setminus \{Z(\omega)^\perp\}$. In particular, convergence holds $\mu^{\mathbb{N}} \times \nu$ -a.e. By Lemma 8, this implies $\int f > 0$. Then, by (1), $\gamma > 0$. \square

REFERENCES

- [BL] P. Bougerol and J. Lacroix. *Products of random matrices with applications to Schrödinger operators*. Birkhäuser, 1985.
- [F] H. Furstenberg. Non-commuting random products. *Trans. AMS*, 108: 377–428, 1963.

³This proof is a bit simpler than that in [BL].

⁴Because $A(v) = w \Rightarrow A^*(w^\perp) = v^\perp$.