

Random perturbations of matrix cocycles

L.-S. YOUNG

*Department of Mathematics, Michigan State University, East Lansing, MI 48824,
USA and Universität Bielefeld*

(Received 23 September 1985)

Abstract. Lyapunov exponents for two-dimensional matrix cocycles are shown to be stable under certain stochastic perturbations.

This paper concerns the effect of stochastic perturbations on Lyapunov exponents in dimension two. Let X be a compact metric space and let $f: X \rightarrow X$ be a homeomorphism preserving an ergodic Borel probability measure m . We consider a continuous map $A: X \rightarrow GL(2, \mathbb{R})$ and let $\lambda_1 \geq \lambda_2$ denote the Lyapunov exponents of A over f . Our main result is the stability of these exponents when A is perturbed randomly and in a relatively unbiased way.

For definiteness, let us first consider the case where A is followed by a small random rotation at each step. More precisely, fix some $\varepsilon > 0$ and let $\dots \omega_{-1}, \omega_0, \omega_1, \dots$ be chosen independently with the uniform distribution on $[-\varepsilon, \varepsilon]$. Let R_α denote rotation by angle α . Then there are numbers $\lambda_1^\varepsilon \geq \lambda_2^\varepsilon$ such that for a.e. x and a.e. choice of $\dots \omega_{-1}, \omega_0, \omega_1, \dots$, the sequence of matrices

$$\{R_{\omega_n} \circ A(f^n x)\}_{n=-\infty}^{\infty} \quad (*)$$

has Lyapunov exponents $\lambda_1^\varepsilon \geq \lambda_2^\varepsilon$. (This is just Oseledec's theorem [O].) We show that $\lambda_i^\varepsilon \rightarrow \lambda_i$ ($i = 1, 2$) as $\varepsilon \rightarrow 0$. We also show that if $\lambda_1 > \lambda_2$ and ε is sufficiently small, then for a.e. x and $\dots \omega_{-1}, \omega_0, \omega_1, \dots$, the Oseledec splittings associated with (*) are very near the corresponding ones for $\{A(f^n x)\}_{n=-\infty}^{\infty}$ for most values of n . Precise statements of these and slightly more general results are given in § 1.

In general we assume that the induced action of the stochastic perturbations on projective spaces have densities with uniform bounds. This is explained in § 1. The fact that some conditions of this type are necessary to ensure stability will also be discussed.

It should be mentioned that *random walks* on $GL(n, \mathbb{R})$ have been studied by a number of people (see e.g. [F], [FK] and [GR]). In particular, the question of stability of exponents has been answered in the affirmative when the limit measure possesses certain irreducibility properties. The reducible case was considered in [KS].

Finally, my primary motivation for studying random perturbations of matrix cocycles is a desire to understand the effect of noise on smooth dynamical systems, i.e. when X is a manifold, f is a diffeomorphism of X and $A = Df$. It would be more satisfactory to obtain some results on how Lyapunov exponents are affected

by perturbations of the map f itself, as opposed to just Df as we have done here. Our methods in this paper alone seem inadequate for that purpose.

I wish to thank M. Denker, F. Ledrappier and C. Preston for helpful discussions. This research was partially supported by NSF.

1. *Definitions and statements of theorems*

Throughout this paper we let X denote a compact metric space, m a Borel probability measure on X and $f: X \rightarrow X$ a homeomorphism of X onto itself preserving m . Unless declared otherwise (as in theorem 3') we shall assume that m is ergodic. This causes no loss of generality since all of our arguments apply to one ergodic component at a time. Let $A_0: X \rightarrow GL(2, \mathbb{R})$ be continuous. For $n > 0$, we write

$$A_0^n(x) = A_0(f^{n-1}x) \circ \dots \circ A_0(x)$$

and

$$A_0^{-n}(x) = A_0(f^{-n}x)^{-1} \circ \dots \circ A_0(f^{-1}x)^{-1}.$$

The *Lyapunov exponents* of A_0 over $f: (X, m) \rightarrow (X, m)$ are defined to be

$$\lambda_1 = \lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_0^n(x)\|$$

and

$$\lambda_2 = -\lim_{n \rightarrow \infty} \frac{1}{n} \log \|A_0^{-n}(x)\|,$$

these limits being constant *a.e.* Moreover, if $\lambda_1 > \lambda_2$, then there is a measurable splitting of the \mathbb{R}^2 -bundle over X into $E^1 \oplus E^2$ with the property that for m -a.e. x and every $v \neq 0 \in E^i(x)$,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \|A_0^n(x)v\| = \lambda_i.$$

We first state our results for a special case which is easy to describe and which the reader may wish to keep in mind. Afterwards we give the more general version stating the hypotheses exactly as we use them. The proofs of these two versions are practically identical.

Fix $\varepsilon > 0$ and let $[-\varepsilon, \varepsilon]^{\mathbb{Z}}$ denote the bi-infinite product of the interval $[-\varepsilon, \varepsilon]$ with itself. Let $\omega_n = (\omega)_n$ be the n 'th coordinate of $\omega \in [-\varepsilon, \varepsilon]^{\mathbb{Z}}$ and let $\sigma: [-\varepsilon, \varepsilon]^{\mathbb{Z}} \rightarrow [-\varepsilon, \varepsilon]^{\mathbb{Z}}$ be the shift operator, i.e. $(\sigma\omega)_n = (\omega)_{n+1}$. If ν_ε is the normalized Lebesgue measure on $[-\varepsilon, \varepsilon]$ and $\nu_\varepsilon^{\mathbb{Z}}$ is its product with itself, then $(f \times \sigma): X \times [-\varepsilon, \varepsilon]^{\mathbb{Z}} \rightarrow X \times [-\varepsilon, \varepsilon]^{\mathbb{Z}}$ preserves the product measure $m \times \nu_\varepsilon^{\mathbb{Z}}$, which is ergodic since the second factor is Bernoulli. Consider the matrix map

$$(x, \omega) \rightarrow R_{\omega_0} \circ A_0(x),$$

with R_α denoting rotation by angle α . We call its Lyapunov exponents $\lambda_1^\varepsilon \geq \lambda_2^\varepsilon$. That is, for m -a.e. x and $\nu_\varepsilon^{\mathbb{Z}}$ -a.e. ω , if $v \neq 0 \in \mathbb{R}^2$ then

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|(R_{\omega_n} \circ A_0(f^n x)) \circ \dots \circ (R_{\omega_0} \circ A_0(x))v\| = \lambda_1^\varepsilon \quad \text{or} \quad \lambda_2^\varepsilon.$$

THEOREM 1'. $\lambda_i^\varepsilon \rightarrow \lambda_i (i = 1, 2)$ as $\varepsilon \rightarrow 0$.

Now suppose $\lambda_1 > \lambda_2$, so that for ε sufficiently small, we have $\lambda_1^\varepsilon > \lambda_2^\varepsilon$ as well. Let $\mathbb{R}^2_{(x,\omega)} = E^1(x, \omega) \oplus E^2(x, \omega)$ denote the Oseledec decomposition of the above cocycle over $(x, \omega) \in X \times [-\varepsilon, \varepsilon]^Z$. Note that if $\omega = \dots 000 \dots$, then $E^i(x, \omega)$ is also written $E^i(x)$. The cardinality of a set is denoted by $\#$, and angle by \angle .

THEOREM 2'. Let $\delta > 0$ be given. Then for all sufficiently small $\varepsilon > 0$, we have for $(m \times \nu_\varepsilon)$ -a.e. (x, ω) ,

$$\lim_{n \rightarrow \infty} \frac{1}{2n+1} \cdot \# \{ |k| \leq n : \angle(E^i(f^k x, \sigma^k \omega), E^i(f^k x)) < \delta, i = 1, 2 \} > 1 - \delta.$$

Next we change our perspective a little: keep $\varepsilon > 0$ fixed and let f and A_0 vary. One scenario that might make sense is the following. Let M be a compact oriented surface endowed with a Riemannian metric and let m be its normalized Riemannian measure. For $f \in \text{Diff}_m^1(M)$, the space of C^1 diffeomorphisms of M preserving m , we consider the Lyapunov exponents of Df composed with $\leq \varepsilon$ -rotations (i.e. $A_0 = Df$ in the preceding discussion). Here no ergodicity is assumed, and we write

$$\lambda_i^\varepsilon(f) = \int \lambda_i^\varepsilon(f; x, \omega) d(m \times \nu_\varepsilon^Z), \quad i = 1, 2.$$

THEOREM 3'. For fixed $\varepsilon > 0$, $f \mapsto \lambda_i^\varepsilon(f)$ is continuous ($i = 1, 2$).

This statement is to be contrasted with the case $\varepsilon = 0$, where Mañé's theorem [M] says that away from Anosov components $\lambda_1(f) = \lambda_2(f) = 0$ for a residual subset of f .

We now begin to state our theorems in greater generality. Let $f: (X, m) \rightarrow (X, m)$ be as before. Let Ω be a compact metric space with a special point $\hat{\omega} \in \Omega$, and let $A: X \times \Omega \rightarrow \text{GL}(2, \mathbb{R})$ be a continuous map. Let $A_0(x) = A(x, \hat{\omega})$. We think of Ω as a parameter space and A as a parametrized family of perturbations of $x \mapsto A_0(x)$. The Lyapunov exponents of A_0 are denoted by $\lambda_1 \geq \lambda_2$. For each $\varepsilon > 0$, we let ν_ε be a Borel probability measure on Ω and let $\lambda_1^\varepsilon \geq \lambda_2^\varepsilon$ be the exponents of $(x, \omega) \mapsto A(x, \omega)$ over the dynamical system $(f \times \sigma): (X \times \Omega^Z, m \times \nu_\varepsilon^Z) \rightarrow (X \times \Omega^Z, m \times \nu_\varepsilon^Z)$.

One could rephrase the definition of λ_i^ε as follows. Take an m -typical point $x \in X$. For $n = 0, 1, 2, \dots$, the map $\cdot \mapsto A(f^n x, \cdot)$ induces a sequence of probabilities on $\text{GL}(2, \mathbb{R})$. Call them $\{\tau_n = \tau_n(x, \varepsilon)\}$. If B_0, B_1, B_2, \dots are independent $\text{GL}(2, \mathbb{R})$ -valued random variables with distributions $\tau_0, \tau_1, \tau_2, \dots$ respectively, then with probability one

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \|B_n \cdots B_0\| = \lambda_1^\varepsilon.$$

We wish to say that $\lambda_i^\varepsilon \rightarrow \lambda_i$ if $\nu_\varepsilon \rightarrow \delta_{\hat{\omega}}$, the dirac measure at $\hat{\omega}$, but will need some extra hypotheses which we now describe. Think of our process as taking place on $X \times \mathbb{P}^1$, where \mathbb{P}^1 is the one-dimensional projective space. Let \mathbb{P}_x^1 denote the copy of \mathbb{P}^1 over $x \in X$, and let \bar{u} denote the direction represented by the unit vector $u \in \mathbb{R}^2$. Now the rules of our game are that $\bar{u} \in \mathbb{P}_x^1$ may be sent to $\overline{A(x, \omega)u} \in \mathbb{P}_{f x}^1$, ω varying over Ω . We assume that for m -a.e. x and every $\bar{u} \in \mathbb{P}_x^1$, the image of the measure ν_ε

on $\mathbb{P}_{f_x}^1$ under the map $\cdot \mapsto \overline{A(x, \cdot)u}$ is absolutely continuous with respect to Lebesgue (written $\ll \text{Leb.}$) with density $\rho_\varepsilon(x, \bar{u})$.

Uniform density condition (UDC). $\exists K > 0$ s.t. for every $\varepsilon > 0$, the following is satisfied for a.e. x and every $\bar{u} \in \mathbb{P}^1$:

- (1) $\text{supp } \rho_\varepsilon(x, \bar{u})$ is contained in the ε -neighbourhood of $\overline{A_0(x)u}$.
- (2) $\rho_\varepsilon(x, \bar{u}) \leq K/\varepsilon$.

THEOREM 1. *Let $A: X \times \Omega \rightarrow \text{GL}(2, \mathbb{R})$ and $\{\nu_\varepsilon\}$ satisfy (UDC). Then $\lambda_i^\varepsilon \rightarrow \lambda_i^\varepsilon$ ($i = 1, 2$) as $\varepsilon \rightarrow 0$.*

THEOREM 2. *Under the same hypotheses as in theorem 1, the conclusion of theorem 2' holds.*

We mention a couple of relatively natural examples to which these generalized hypotheses apply. Take $f: (X, m) \curvearrowright$ and $A_0: X \rightarrow \text{GL}(2, \mathbb{R})$ to be anything one desires. Let $\Omega = [-1, 1]^4$ and let ν_ε be normalized Lebesgue measure with support $[-\varepsilon, \varepsilon]^4 \subset [-1, 1]^4$. Define $A: X \times \Omega \rightarrow \text{GL}(2, \mathbb{R})$ by

$$A\left(x, \begin{pmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_3 & \varepsilon_4 \end{pmatrix}\right) = \begin{pmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_3 & \varepsilon_4 \end{pmatrix} + A_0(x)$$

$$\text{or } \left(I + \begin{pmatrix} \varepsilon_1 & \varepsilon_2 \\ \varepsilon_3 & \varepsilon_4 \end{pmatrix}\right) \cdot A_0(x).$$

It is easy to verify (UDC). Instead of assuming that ν_ε is normalized Lebesgue measure on $[-\varepsilon, \varepsilon]^4$, one could also choose an L^∞ -function $\varphi: \mathbb{R}^4 \rightarrow \mathbb{R}$ with $\varphi \geq 0$, $\text{supp } \varphi \subset [-1, 1]^4$ and $\int \varphi dx = 1$ (here $x \in \mathbb{R}^4$ and $dx = d(\text{Leb.})$) and let ν_ε be defined by

$$\nu_\varepsilon(E) = \int_E \frac{1}{\varepsilon^4} \varphi\left(\frac{1}{\varepsilon} x\right) dx.$$

We now discuss the necessity of some regularity assumptions on ν_ε . First, keeping $f: (X, m) \curvearrowright$ fixed, the map $A_0 \mapsto \lambda_1(A_0)$ is known to be upper semi-continuous but not always continuous. Mañé has techniques for ‘jumping’ λ_1 ([M]). (Our situation is simpler since he perturbs f and so has spatial considerations that are irrelevant here.) Thouvenot also has examples in which $\lambda_1(A_0) > 0$ but multiplication of A_0 by a suitably chosen constant matrix near the identity can bring λ_1 arbitrarily close to zero [T].

Once we have *deterministic* perturbations A_ε of A_0 with $A_\varepsilon \rightarrow A_0$ and $\lim_{\varepsilon \rightarrow 0} \lambda_1(A_\varepsilon) < \lambda_1(A_0)$, we can approximate A_ε by processes for which the functions $\rho_\varepsilon(\cdot, \cdot)$ are well defined and still have the larger exponent bounded away from $\lambda_1(A_0)$ as $\varepsilon \rightarrow 0$. Thus in general some conditions must be imposed on $\rho_\varepsilon(\cdot, \cdot)$ to ensure stability. We do not know what the weakest natural assumptions are.

The organization of the rest of this paper is as follows. Most of the work is contained in the proof of theorem 1, which is given in §§ 2 and 3. The proof of theorem 2 occupies § 4 and theorem 3' is proved in § 5.

2. Proof of theorem 1

For fixed $\varepsilon > 0$ we define a Markov chain with state space $X \times \mathbb{P}^1$ as follows. Let $x \in X$, $\bar{u} \in \mathbb{P}^1_x$, and let E be a Borel subset of $X \times \mathbb{P}^1$. The transition probabilities are given by

$$P^\varepsilon((x, \bar{u}), E) = \nu_\varepsilon\{\omega: \overline{A(x, \omega)u} \in E \cap \mathbb{P}^1_{fx}\}.$$

Definition 2.1. A Borel probability measure μ on $X \times \mathbb{P}^1$ is called an *invariant measure* for $P^\varepsilon(\cdot, \cdot)$ if μ satisfies:

- (1) $\mu E = \int P^\varepsilon((x, \bar{u}), E) d\mu(x, \bar{u})$ for every Borel subset $E \subset X \times \mathbb{P}^1$, and
- (2) $p\mu = m$ if $p: X \times \mathbb{P}^1 \rightarrow X$ is the projection map.

Invariant measures of this type always exist. Let μ_ε be one such measure for the process $P^\varepsilon(\cdot, \cdot)$ and let $\{(\mu_\varepsilon)_x, x \in X\}$ be a canonical system of conditional probability measures of μ_ε on the \mathbb{P}^1 -fibres over X . Since we assume that $\rho_\varepsilon(x, \bar{u})$ exists for a.e. x and every \bar{u} , it follows that $(\mu_\varepsilon)_x \ll \text{Leb. on } \mathbb{P}^1_x$ for a.e. x .

LEMMA 2.2. $\lambda_1^\varepsilon = \int \log |A(x, \omega)u| d(\mu_\varepsilon \times \nu_\varepsilon)$.

Here u is understood to be one of the unit vectors in the equivalent class \bar{u} .

Proof. Consider the map $F: X \times \mathbb{P}^1 \times \Omega^\mathbb{N} \ni$ given by $F(x, \bar{u}, \omega) = (fx, \overline{A(x, \omega_0)u}, \sigma\omega)$. (With \mathbb{N} denoting the set of natural numbers it is obvious what $\Omega^\mathbb{N}$ means.) We leave it to the reader to verify that F preserves $\mu_\varepsilon \times \nu_\varepsilon^\mathbb{N}$. Let $\varphi: X \times \mathbb{P}^1 \times \Omega^\mathbb{N} \rightarrow \mathbb{R}$ be defined by $\varphi(x, \bar{u}, \omega) = \log |A(x, \omega_0)u|$. Then

$$\sum_{i=0}^{n-1} F^i(x, \bar{u}, \omega) = \log |A(f^{n-1}x, \omega_{n-1}) \circ \dots \circ A(x, \omega_0)u|$$

and we have, for m-a.e. x and all except possibly one $\bar{u} \in \mathbb{P}^1_x$,

$$\frac{1}{n} \sum_{i=0}^{n-1} F^i(x, \bar{u}, \omega) \rightarrow \lambda_1^\varepsilon$$

as $n \rightarrow \infty$. Since $(\mu_\varepsilon)_x \ll \text{Leb.}$ for a.e. x , the above convergence occurs for $(\mu_\varepsilon \times \nu_\varepsilon^\mathbb{N})$ -a.e. (x, \bar{u}, ω) . The Birkhoff ergodic theorem applied to φ then gives

$$\lambda_1^\varepsilon = \int \varphi d(\mu_\varepsilon \times \nu_\varepsilon^\mathbb{N})$$

and hence the desired result. □

Now if μ is a weak limit point of μ_ε as $\varepsilon \rightarrow 0$, then μ is left invariant by the transformation $F_0: X \times \mathbb{P}^1 \ni$ defined by $F_0(x, \bar{u}) = (fx, \overline{A_0(x)u})$.

COROLLARY 2.3. *If $\mu_\varepsilon \rightarrow \mu$ as $\varepsilon \rightarrow 0$, then*

$$\lambda_1^\varepsilon \rightarrow \int \log |A_0(x)u| d\mu.$$

If $\lambda_1 = \lambda_2$, then theorem 1 is automatically true. One can conclude that from general principles, namely the upper semi-continuity of exponents, or alternatively one could see that from corollary 2.3 with μ being any limit point of μ_ε , $\varepsilon \rightarrow 0$. So from

now on we may assume that $\lambda_1 > \lambda_2$. Let $E^1 \oplus E^2$ denote the splitting of $X \times \mathbb{R}^2$ associated with A_0 .

The set of Borel probability measures on $X \times \mathbb{P}^1$ that are F_0 -invariant and that project onto m has exactly two ergodic components, which we denote by μ^1 and μ^2 . They have the property that if $\{\mu_x^i, x \in X\}$ are their respective conditional measures on \mathbb{P}^1 -fibres, $i = 1, 2$, then for m -a.e. x , $\mu_x^i = \delta_{E^i(x)}$ (a benign abuse of notation). We then see from corollary 2.3 that showing $\lambda_1^\epsilon \rightarrow \lambda_1$ is equivalent to showing $\mu_\epsilon \rightarrow \mu^1$. Once this is proved we have $\lambda_2^\epsilon \rightarrow \lambda_2$ immediately since the additivity of the $\log |\det(\cdot)|$ function guarantees that $\lambda_1^\epsilon + \lambda_2^\epsilon \rightarrow \lambda_1 + \lambda_2$.

The idea of working with invariant measures on $X \times \mathbb{P}^1$ to obtain information on Lyapunov exponents probably goes back to [F]. We remark also that much of the preceding discussion, including the statement of lemma 2.2, works quite generally for perturbations of $GL(n, \mathbb{R})$ cocycles. We restrict ourselves to the $n = 2$ case in this paper because of the estimates in showing $\mu_\epsilon \rightarrow \mu^1$.

Let us indicate briefly how our argument for $\mu_\epsilon \rightarrow \mu^1$ goes in the 2×2 case. The estimates are carried out in the next section. First we choose and fix a measurable set $\mathcal{N} \subset X \times \mathbb{P}^1$ with the property that for a.e. x , $\mathcal{N} \cap \mathbb{P}_x^1$ contains an open neighbourhood of $\delta_{E^2(x)}$. Our objective is to show that $\mu_\epsilon \mathcal{N} \rightarrow 0$ as $\epsilon \rightarrow 0$.

To do that, consider the operator \mathcal{L} on $\mathcal{M}(X \times \mathbb{P}^1)$, the space of Borel probability measures on $X \times \mathbb{P}^1$, defined by

$$\mathcal{L}_\epsilon(\cdot)E = \int P^\epsilon((x, \bar{u}), E) d(\cdot).$$

As before, let us denote the conditional measure of $\mu \in \mathcal{M}(X \times \mathbb{P}^1)$ on \mathbb{P}_y^1 by μ_y . We will prove

PROPOSITION 2.4 (MAIN PROPOSITION). *There is a function $\alpha(\epsilon)$ tending to 0 as $\epsilon \rightarrow 0$ such that for any $\mu \in \mathcal{M}(X \times \mathbb{P}^1)$ with $p\mu = m$, we have, for m -a.e. x*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} (\mathcal{L}_\epsilon^k \mu)_{f^k x} \mathcal{N} < \alpha(\epsilon).$$

By taking $\mu_\epsilon = \mu$ in the above inequality, for instance, one sees immediately that $\mu_\epsilon \mathcal{N} \leq \alpha(\epsilon)$, which finishes the proof of theorem 1.

3. Proof of theorem 1 (continued)

Dividing by $|\det A_0|$, we may and will assume throughout this section that $\lambda_1 = \lambda > 0$ and $\lambda_2 = -\lambda$.

(A) *Change of coordinates.* We fix $\epsilon_1 > 0$ with $e^{2\lambda - 5\epsilon_1} > 1$.

LEMMA 3.1. *There exist measurable functions $a, b, c, \tau: X \rightarrow \mathbb{R}$ such that*

- (1) $0 < |a| \leq 1$, $b^2 + c^2 \geq 1$; and
- (2) if $\psi_x: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is given by

$$\psi_x = R_{\tau(x)} \circ \begin{pmatrix} a(x) & b(x) \\ 0 & c(x) \end{pmatrix}$$

then

$$\psi_{f_x}^{-1} \circ A_0(x) \circ \psi_x = \begin{pmatrix} B_x^1 & 0 \\ 0 & B_x^2 \end{pmatrix}$$

where

$$|B_x^1| \geq e^{\lambda - \epsilon_1}, \quad |B_x^2| \leq e^{-\lambda + \epsilon_1}.$$

This is nothing more than a slight modification of the standard Lyapunov metric used in the theory of non-uniformly hyperbolic systems [P]. We proceed to define the functions anyway.

Proof of lemma 3.1. Let $\{e_1, e_2\}$ denote the usual basis of \mathbb{R}^2 . We shall define for a.e. x a linear transformation $\psi_x: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $\psi_x e_1 \in E^1(x)$ and $\psi_x e_2 \in E^2(x)$. The role of $R_{\tau(x)}$ in the statement of the lemma is simply to rotate e_1 to the direction of $E^1(x)$.

For $i = 1, 2$ pick a measurable function $x \mapsto v_i(x)$ such that $v_i(x) \in E^i(x)$ and $|v_i(x)| = 1$. Recall that for $n > 0$, $A_0^n(x) = A_0(f^{n-1}x) \circ \dots \circ A_0(x)$ and $A_0^{-n}(x) = A_0(f^{-n}x)^{-1} \circ \dots \circ A_0(f^{-1}x)^{-1}$. For convenience we write $A_0^0 = \text{Id}$. Let

$$\psi_x e_1 = \left(\sum_{n=0}^{\infty} \frac{|A_0^{-n}(x)v_1(x)|}{e^{n(-\lambda + \epsilon_1)}} \right)^{-1} v_1(x).$$

Since the $n = 0$ term is equal to 1, we have $|a(x)| = |\psi_x e_1| < 1$. A simple computation gives

$$|B_x^1| = |(\psi_{f_x}^{-1} \circ A_0(x) \circ \psi_x) e_1| \geq e^{\lambda - \epsilon_1}.$$

To define $\psi_x e_2$, we first let

$$\tilde{\psi}_x e_2 = \left(\sum_{n=0}^{\infty} \frac{|A_0^n(x)v_2(x)|}{e^{n(-\lambda + \frac{1}{2}\epsilon_1)}} \right)^{-1} v_2(x).$$

As before, we have $|(\tilde{\psi}_{f_x}^{-1} \circ A_0(x) \circ \tilde{\psi}_x) e_2| \leq e^{-\lambda + \frac{1}{2}\epsilon_1}$. To obtain $|\psi_x e_2| \geq 1$ however, we consider the function $l: X \rightarrow \mathbb{R}^+$ defined by

$$l(x) = \sup_{m,n \in \mathbb{Z}} \frac{|A_0^{m+n}(x)v_2(x)|}{|A_0^m(x)v_2(x)|} e^{n\lambda} e^{-((|m|+|n|)\epsilon_1)/4}.$$

It is straightforward to verify that l so defined is a finite valued measurable function with the following properties:

- (1) $\frac{l(x)}{l(fx)} \leq e^{\frac{1}{2}\epsilon_1}$ for a.e. x
- (2) $|\tilde{\psi}_x e_2| \geq l(x)^{-1} (1 - e^{(-\epsilon_1)/4})$ a.e. x .

We then set

$$\psi_x e_2 = l(x) (1 - e^{(-\epsilon_1)/4})^{-1} \tilde{\psi}_x e_2$$

and check that $|\psi_x e_2| \geq 1$ and

$$|(\psi_{f_x}^{-1} \circ A_0(x) \circ \psi_x) e_2| \leq \frac{l(x)}{l(fx)} \cdot e^{-\lambda + \frac{1}{2}\epsilon_1} \leq e^{-\lambda + \epsilon_1}.$$

(B) *New coordinates on projective spaces.* Recall that for each $x \in X$ we are concerned with matrices $A(x, \omega)$, $\omega \in \Omega$, which we view as perturbations of the matrix $A_0(x)$.

First we decided to work in projective spaces. That is, we look at the induced action of $A(x, \omega)$ from \mathbb{P}_x^1 to \mathbb{P}_{fx}^1 and view them as perturbations of the map $\bar{u} \rightarrow \bar{A}_0(x)\bar{u}$. Since we will be dealing with projectivized maps, let us fix the following notation: if $L: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a linear transformation, then $\bar{L}: \mathbb{P}^1 \rightarrow \mathbb{P}^1$ is defined to be $\bar{L}\bar{u} = \overline{Lu}$.

Next we make a change of coordinates on \mathbb{P}_x^1 for m -a.e. x . This is done via the maps $\{\psi_x\}$ defined in lemma 3.1 and will be referred to as ψ -coordinates in the future. Let

$$B(x, \omega) = \psi_{fx}^{-1} \circ A(x, \omega) \circ \psi_x \quad \text{and} \quad \bar{B}_0(x) = \psi_{fx}^{-1} \circ A_0(x) \circ \psi_x.$$

Then $\bar{B}(x, \omega) = \bar{\psi}_{fx}^{-1} \circ \bar{A}(x, \omega) \circ \bar{\psi}_x$ is the representation of $\bar{A}(x, \omega)$ in ψ -coordinates and from now on we will be working with $\{\bar{B}(x, \omega), \omega \in \Omega\}$ as perturbations of $\bar{B}_0(x): \mathbb{P}_x^1 \rightarrow \mathbb{P}_{fx}^1$.

Let us write $\mathbb{P}^1 \cong S^1/\{\theta, \theta + \pi\} \cong [0, \pi]/\{0, \pi\}$. Then for a.e. x , $\bar{B}_0(x)$ has exactly two fixed points, namely 0 and $\pi/2$. Moreover, $|(\bar{B}_0(x))'(0)| \leq e^{-2\lambda+2\varepsilon_1}$ and $|(\bar{B}_0(x))'(\pi/2)| \geq e^{2\lambda-2\varepsilon_1}$. The price we pay for having these nice uniform properties for \bar{B}_0 is the loss of uniformness in the perturbations.

We need some control on the derivative of $\bar{\psi}_x$.

LEMMA 3.2. $\exists \delta_0 > 0$ s.t. for m -a.e. x , if $\theta_i \in \mathbb{P}_x^1$ satisfies $|\theta_i - (\pi/2)| \leq \delta_0, i = 1, 2$, then

$$\frac{\bar{\psi}'_x(\theta_1)}{\bar{\psi}'_x(\theta_2)} \leq e^{\varepsilon_1}.$$

Proof. Since this estimate is uniform in x , we fix an arbitrary point and omit all mention of it. Recall that

$$\psi = R_\tau \circ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$$

where $0 < |a| < 1$ and $b^2 + c^2 \geq 1$. A small computation gives

$$\bar{\psi}(\theta) = \tau + \tan^{-1} \frac{c \sin \theta}{a \cos \theta + b \sin \theta}$$

and

$$\bar{\psi}'(\theta) = \frac{ac}{a^2 \cos^2 \theta + 2ab \sin \theta \cos \theta + (b^2 + c^2) \sin^2 \theta}$$

Let $T_1(\theta), T_2(\theta)$ and $T_3(\theta)$ be the three terms in the denominator of $\bar{\psi}'(\theta)$. We claim that $T_1(\theta)/T_3(\theta)$ and $T_2(\theta)/T_3(\theta) \rightarrow 0$ as $\theta \rightarrow \pi/2$. The first comparison is obvious, as is the second if we argue the cases $|b| \geq 1$ and $|b| < 1$ separately.

Now for θ_1, θ_2 near $\pi/2, T_3(\theta_1)/T_3(\theta_2) \approx 1$. Hence $\bar{\psi}'(\theta_1)/\bar{\psi}'(\theta_2) \approx 1$. □

Choose $\delta_1 < \delta_0$ such that for a.e. $x, \bar{B}_0(x)\theta \in ((\pi/2) - \delta_0, (\pi/2) + \delta_0)$ and $|(\bar{B}_0(x))'(\theta)| \geq e^{2\lambda-3\varepsilon_1}$ for every $\theta \in ((\pi/2) - \delta_1, (\pi/2) + \delta_1)$. We define \mathcal{N} to be a subset of $X \times \mathbb{P}^1$ with the property that $\bar{\psi}_x^{-1}(\mathcal{N} \cap \mathbb{P}_x^1) = ((\pi/2) - \delta_1, (\pi/2) + \delta_1)$ for a.e. x . The remainder of § 3 is devoted to showing $\mu_\varepsilon \mathcal{N} \rightarrow 0$ as $\varepsilon \rightarrow 0$.

(C) *Transformation of densities.* Let $\varepsilon > 0$ be fixed throughout this subsection. We assume that it is sufficiently small (see case 1). Let μ be any Borel probability measure on $X \times \mathbb{P}^1$ that projects onto m and whose conditional measures on \mathbb{P}^1 -fibres are $\ll \text{Leb}$. We consider a generic point $x \in X$ (to which all of our previous ‘a.e.’

statements apply). Let η_0 denote the density of μ_x on \mathbb{P}_x^1 in ψ -coordinates, and let η_1 denote the density of $(\mathcal{L}_\varepsilon \mu)_{f_x}$ on $\mathbb{P}_{f_x}^1$, also in ψ -coordinates. Let $\tilde{\eta}_i = \max\{\eta_i(\theta) : |\theta - (\pi/2)| \leq \delta_i\}$, $i = 0, 1$. Let δ_2 be chosen so that $\delta_1 < \delta_2 \leq \delta_0$ and $e^{2\lambda - 3\varepsilon_1} \delta_1 \geq \delta_2$. We distinguish between the following two cases:

Case 1. For $t = (\pi/2) \pm \delta_1 \in \mathbb{P}_x^1$, $\bar{\psi}_{f_x}^{-1}(\text{supp } \rho_\varepsilon(x, \bar{\psi}_x t))$ is disjoint from the δ_1 -neighbourhood of $\pi/2$ in $\mathbb{P}_{f_x}^1$.

Claim. $\tilde{\eta}_1 \leq \tilde{\eta}_0 \cdot e^{-2\lambda + 5\varepsilon_1}$.

We assume that ε is sufficiently small that for m -a.e. x and ν_ε -a.e. $\omega \in \Omega$, $|(\bar{A}_0(x) \circ \bar{A}(x, \omega)^{-1})'| \leq e^{\varepsilon_1}$. This is legitimate since by hypothesis the support of $\rho_\varepsilon(x, \bar{u})$ is contained in the ε -neighbourhood of $\bar{A}_0(x)\bar{u}$, and this forces $A(x, \omega)$ to be very near $A_0(x)$ for ν_ε -a.e. $\omega \in \Omega$. (We are using implicitly the fact that $\{A_0(x), x \in X\}$ is a compact subset of $\text{GL}(2, \mathbb{R})$.)

To prove the claim we note that η_0 and η_1 are related by

$$\eta_1(\theta) = \int \eta_0(\bar{B}(x, \omega)^{-1}\theta) \cdot |(\bar{B}(x, \omega)^{-1})'\theta| \, d\nu_\varepsilon(\omega).$$

Since

$$\bar{B}(x, \omega)^{-1} = (\bar{\psi}_x^{-1} \circ \bar{A}_0(x)^{-1} \circ \bar{\psi}_{f_x}) \circ (\bar{\psi}_{f_x}^{-1} \circ \bar{A}_0(x) \circ \bar{A}(x, \omega)^{-1} \circ \bar{\psi}_{f_x}),$$

we have for $\theta \in ((\pi/2) - \delta_1, (\pi/2) + \delta_1)$,

$$|(\bar{B}(x, \omega)^{-1})'\theta| \leq |(\bar{B}_0(x)^{-1})'\theta_1| \cdot |(\bar{A}_0(x) \circ \bar{A}(x, \omega)^{-1})'| \cdot \left| \frac{\bar{\psi}'_{f_x} \theta}{\bar{\psi}'_{f_x} \theta_1} \right|$$

where $\theta_1 = \bar{\psi}_{f_x}^{-1} \circ \bar{A}_0(x) \circ \bar{A}(x, \omega)^{-1} \circ \bar{\psi}_{f_x}(\theta)$. That is, if $t \in \mathbb{P}_x^1$ is such that $\bar{B}_0(x)t = \theta_1$, then $\bar{B}(x, \omega)t = \theta$. It then follows from the hypothesis of case 1 that $t \in ((\pi/2) - \delta_1, (\pi/2) + \delta_1)$, so that $|(\bar{B}_0(x)^{-1})'\theta_1| \leq e^{-2\lambda + 3\varepsilon_1}$. The second and third terms are each $\leq e^{\varepsilon_1}$. This proves the claim.

Case 2. Not case 1.

Claim. $\tilde{\eta}_2 \leq K/(\delta_2 - \delta_1) \cdot e^{\varepsilon_1}$.

We suppose for definiteness that $\bar{B}_0(x)((\pi/2) + \delta_1) > (\pi/2) + \delta_2$ and that $\bar{\psi}_{f_x}^{-1} \text{supp } \rho_\varepsilon(x, \bar{\psi}_x((\pi/2) + \delta_1))$ intersects the δ_1 -neighbourhood of $\pi/2$ in $\mathbb{P}_{f_x}^1$. It then follows from our hypotheses on $\rho_\varepsilon(\cdot, \cdot)$ and the mean value theorem that there exists $\theta_0 \in ((\pi/2) + \delta_1, (\pi/2) + \delta_2) \subset \mathbb{P}_{f_x}^1$ such that $|\bar{\psi}'_{f_x} \theta_0| \leq \varepsilon/(\delta_2 - \delta_1)$. By lemma 3.1, we have $|\bar{\psi}'_{f_x} \theta| \leq \varepsilon e^{\varepsilon_1}/(\delta_2 - \delta_1)$ for all $\theta \in ((\pi/2) - \delta_0, (\pi/2) + \delta_0)$. Now for $\theta \in \mathbb{P}_{f_x}^1$,

$$\eta_1(\theta) \leq \max_{t \in \mathbb{P}_x^1} \{\rho_\varepsilon(x, \bar{\psi}_x t)(\bar{\psi}_{f_x} \theta) \cdot |\bar{\psi}'_{f_x} \theta|\}.$$

Thus for $\theta \in ((\pi/2) - \delta_1, (\pi/2) + \delta_1)$,

$$\eta_1(\theta) \leq \frac{K}{\varepsilon} \cdot \frac{\varepsilon}{\delta_2 - \delta_1} \cdot e^{\varepsilon_1}.$$

(D) *Counting.* Let μ be a measure on $X \times \mathbb{P}^1$ that projects onto m . For $x \in X$, let $\eta_k^\varepsilon(x)$ denote the density of $(\mathcal{L}_\varepsilon^k \mu)_{f^k x}$ on $\mathbb{P}_{f^k x}^1$ in ψ -coordinates, and let $\tilde{\eta}_k^\varepsilon(x) = \max\{\eta_k^\varepsilon(x)\theta : |\theta - (\pi/2)| \leq \delta_1\}$. Note that independent of the measure class of μ , $\tilde{\eta}_k^\varepsilon$ is well defined a.e. for $k \geq 1$. Also, $(\mathcal{L}_\varepsilon^k \mu)_{f^k x} \mathcal{N} \leq 2\delta_1 \cdot \tilde{\eta}_k^\varepsilon(x)$.

Now let $\Lambda_\varepsilon = \{x \in X : \text{case 1 applies}\}$. Let $C = e^{-2\lambda + 5\varepsilon_1} < 1$ and $D = Ke^\varepsilon / (\delta_2 - \delta_1)$. It is important that these constants are independent of ε . In the last subsection we showed that for ε sufficiently small, $\tilde{\eta}_{k+1}^\varepsilon(x) \leq C\tilde{\eta}_k^\varepsilon(x)$ if $f^k x \in \Lambda_\varepsilon$, while $\tilde{\eta}_{k+1}^\varepsilon(x) \leq D$ if $f^k x \notin \Lambda_\varepsilon$. If $m\Lambda_\varepsilon = 1$, then clearly $(1/n) \sum_{k=0}^{n-1} (\mathcal{L}_\varepsilon^k \mu)_{f^k x} \mathcal{N} \rightarrow 0$ as $n \rightarrow \infty$. Suppose then $m\Lambda_\varepsilon < 1$. Let $N_\varepsilon = (1 - m\Lambda_\varepsilon)^{-1}$. Since for a.e. x , $(1/n) \sum_{k=0}^{n-1} \chi_{\Lambda_\varepsilon}(f^k x) \rightarrow m\Lambda_\varepsilon$ as $n \rightarrow \infty$, $(1/n) \sum_{k=0}^{n-1} (\mathcal{L}_\varepsilon^k \mu)_{f^k x} \mathcal{N}$ is no bigger than

$$+ \frac{1}{N_\varepsilon} \sum_{k=0}^{\lfloor N_\varepsilon \rfloor} DC^k \leq \frac{D}{N_\varepsilon(1-C)}$$

for large n . With $m\Lambda_\varepsilon \rightarrow 1$ as $\varepsilon \rightarrow 0$, this completes the proof of proposition 2.4. □

4. Proof of theorem 2

We assume throughout this section that $\lambda_1 > \lambda_2$. By theorem 1, we have $\lambda_1^\varepsilon > \lambda_2^\varepsilon$ for all sufficiently small $\varepsilon > 0$. Let one such small ε be fixed. Consider the map $F : X \times \mathbb{P}^1 \times \Omega^Z \rightarrow X \times \mathbb{P}^1 \times \Omega^Z$ defined by

$$F(x, \bar{u}, \omega) = (fx, \bar{A}(x, \omega_0)\bar{u}, \sigma\omega).$$

For $i = 1, 2$, let μ_ε^i be the Borel probability measure on $X \times \mathbb{P}^1 \times \Omega^Z$ with the following properties:

- (1) if $p_1 : X \times \mathbb{P}^1 \times \Omega^Z \rightarrow X \times \Omega^Z$ is the projection map, then $p_1 \mu_\varepsilon^i = m \times \nu_\varepsilon^Z$, $i = 1, 2$;
- (2) if $\{(\mu_\varepsilon^i)_{(x,\omega)}, (x, \omega) \in X \times \Omega^Z\}$ is a canonical system of conditional measures of μ_ε^i on \mathbb{P}^1 -fibres, then $(\mu_\varepsilon^i)_{(x,\omega)} = \delta_{E^i(x,\omega)}$ for a.e. (x, ω) .

LEMMA 4.1. *If $p_2 : X \times \mathbb{P}^1 \times \Omega^Z \rightarrow X \times \mathbb{P}^1$ is the projection map, then $p_2 \mu_\varepsilon^1 = \mu_\varepsilon$.*

Proof. Since μ_ε has absolutely continuous conditional measures on \mathbb{P}^1 -fibres, it is easy to see (using coordinates induced from the so-called Lyapunov metrics, for instance) that $F^n(\mu_\varepsilon \times \nu_\varepsilon^Z) \rightarrow \mu_\varepsilon^1$ as $n \rightarrow \infty$. One then verifies that $p_2(F^n(\mu_\varepsilon \times \nu_\varepsilon^Z)) = \mu_\varepsilon$ for all $n \geq 0$, which essentially follows from the definition of μ_ε . □

Proof of theorem 2. For small $\delta > 0$, let

$$\mathcal{N}_\delta = \{(x, \bar{u}) : x \in X, |\bar{u} - E^1(x)| < \delta\}.$$

Since $p_2 \mu_\varepsilon^1 = \mu_\varepsilon \rightarrow \mu^1$ (see § 2 for the definition of μ^1), it follows that $\mu_\varepsilon^1(\mathcal{N}_\delta \times \Omega^Z) \rightarrow 1$ as $\varepsilon \rightarrow 0$. That $E^1(x, \omega) \rightarrow E^1(x)$ in the manner described in the statement of theorem 2 is a consequence of the ergodic theorem applied to the transformation $F : (X \times \mathbb{P}^1 \times \Omega^Z, \mu_\varepsilon^1) \rightarrow (X \times \mathbb{P}^1 \times \Omega^Z, \mu_\varepsilon^1)$ and the function $\chi_{\mathcal{N}_\delta}$.

The same argument is valid for showing $E^2(x, \omega) \rightarrow E^2(x)$, provided we can argue that $p_2 \mu_\varepsilon^2 \rightarrow \mu^2$. As in § 2, let $\varphi : X \times \mathbb{P}^1 \times \Omega^Z \rightarrow \mathbb{R}$ be defined by

$$\varphi(x, \bar{u}, \omega) = \log |A(x, \omega_0)u|.$$

Then for $i = 1, 2$ and $\varepsilon > 0$, $\lambda_i^\varepsilon = \int \varphi d\mu_\varepsilon^i$. We have already shown that $\lambda_2^\varepsilon \rightarrow \lambda_2$. Hence $p_2 \mu_\varepsilon^2 \rightarrow \mu^2$. □

5. Proof of theorem 3'

Let $\varepsilon > 0$ be fixed. Referring the reader to § 1 for the definition of this random rotation process, we first observe that if $f \in \text{Diff}_m^1(M)$ has the property that $\lambda_1^\varepsilon(f; x, \omega) > \lambda_2^\varepsilon(f; x, \omega)$, for a.e. (x, ω) , then the invariant measure of this process

is unique. To see this let $\mu_{\epsilon, f}^1$ be as defined at the beginning of § 4. Then $\mu_{\epsilon, f} = p_2 \mu_{\epsilon, f}^1$ by lemma 4.1.

Next we observe that if $\varphi_f(x, \bar{u}) = \log |(Df)_{x, \bar{u}}|$ and $\mu_{\epsilon, f}$ is any invariant measure, then $\lambda_1^\epsilon(f) = \int \varphi_f d\mu_{\epsilon, f}$. This follows from lemma 2.2 and is true regardless of whether the exponents are distinct or not.

Now suppose $f_n \rightarrow f_0$ as $n \rightarrow \infty$. For each n , let μ_{ϵ, f_n} be an invariant measure for the rotation process associated with f_n . Then any accumulation point of μ_{ϵ, f_n} as $n \rightarrow \infty$ is an invariant measure for the process defined using f_0 . Assume first that the exponents of f_0 are distinct a.e. Then the uniqueness of μ_{ϵ, f_0} forces $\mu_{\epsilon, f_n} \rightarrow \mu_{\epsilon, f_0}$. Since $\varphi_{f_n} \rightarrow \varphi_{f_0}$ uniformly, we can conclude that

$$\lambda_1^\epsilon(f_n) = \int \varphi_{f_n} d\mu_{\epsilon, f_n} \rightarrow \int \varphi_{f_0} d\mu_{\epsilon, f_0} = \lambda_1^\epsilon(f_0).$$

If $Z = \{x : \lambda_1^\epsilon(f_0; x, \omega) = \lambda_2^\epsilon(f_0; x, \omega)\}$ has positive measure, then the above argument works even if μ_{ϵ, f_0} restricted to the unit tangent bundle over Z is not unique. This completes the proof. □

The not so surprising message contained in theorem 3' is that infinitesimal noise tends to render dynamical invariants more continuous than they really are. In fact we have the following picture: Let $\mathcal{D} = \text{Diff}_m^1(M)$ and let $\lambda_1 : \mathcal{D} \times [0, \infty) \rightarrow \mathbb{R}$ be given by $\lambda_1(f, \epsilon) = \lambda_1^\epsilon(f)$. Then

- (1) $\lambda_1 | \mathcal{D} \times (0, \infty)$ is continuous (same proof as above);
- (2) for each $f \in \mathcal{D}$, $\lambda_1(f, \epsilon) \rightarrow \lambda_1(f, 0)$ as $\epsilon \rightarrow 0$ (theorem 1);
- (3) $\lambda_1 | \mathcal{D} \times \{0\}$ is continuous on Anosov components and generically zero away from them (Mañé's theorem [M]).

REFERENCES

[F] H. Furstenberg. Noncommuting matrix products. *Trans. Amer. Math. Soc.* **108** (1963), 377-428.
 [FK] H. Furstenberg & Y. Kifer. Random matrix products and measures on projective spaces. *Israel J. of Math.* **46** Nos. 1-2 (1983), 12-32
 [GR] Y. Guivarc'h & Raugi. Products of random matrices: convergence theorems. Preprint
 [KS] Y. Kifer & E. Slud. Perturbations of random matrix products in a reducible case. *Ergod. Th. & Dynam. Syst.* **2** (1982), 367-382.
 [M] R. Mañé. The Lyapunov exponents of generic area preserving diffeomorphisms. In preparation
 [O] V. I. Oseledec. A multiplicative ergodic theorem; Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19** (1968) 197-221
 [P] Ya. B. Pesin. Characteristic Lyapunov exponents and smooth ergodic theory. *Russ. Math. Surveys* **32:4** (1977), 55-114
 [T] J.-P. Thouvenot. Personal communication