ISRAEL JOURNAL OF MATHEMATICS **229** (2019), 461–485
DOI: 10.1007/s11856-018-1809-7

CONTINUITY OF LYAPUNOV EXPONENTS IN THE C^0 TOPOLOGY*

BY

Marcelo Viana

IMPA, Est. D. Castorina 110, 22460-320 Rio de Janeiro, Brazil
e-mail: viana@impa.br

AND

Jiagang Yang

Departamento de Geometria, Instituto de Matemática e Estatística
Universidade Federal Fluminense, Niterói, Brazil
e-mail: yangjg@impa.br

Dedicated to Welington de Melo, in memoriam.

ABSTRACT

We prove that the Bochi–Mañé theorem is false, in general, for linear cocycles over non-invertible maps: there are C^0-open subsets of linear cocycles that are not uniformly hyperbolic and yet have Lyapunov exponents bounded from zero.

1. Introduction

Bochi [4, 5] proved that every continuous SL(2)-cocycle over an aperiodic invertible system can be approximated in the C^0 topology by cocycles whose Lyapunov exponents vanish, unless it is uniformly hyperbolic. The (harder) version of the statement for derivative cocycles of area-preserving diffeomorphisms on surfaces had been claimed by Mañé [23] almost two decades before.

* Partially supported by Fondation Louis D.—Institut de France (project coordinated by M. Viana), CNPq and FAPERJ
Received November 13, 2017 and in revised form May 6, 2018

Bochi [4, 5] also completed the proof of this harder claim, based on an outline by Mañé. These results were then extended to arbitrary dimension by Bochi and Viana [7] and Bochi [6].

In this paper, we prove that the Bochi–Mañé theorem does not hold, in general, for cocycles over non-invertible systems: surprisingly, in the non-invertible setting there exist C^0-open sets of SL(2)-cocycles which are not uniformly hyperbolic and whose exponents are bounded away from zero. Indeed, we provide two different constructions of such open sets.

The first one (Theorem A) applies to Hölder continuous cocycles satisfying a bunching condition. The second one (Theorem B) has no bunching hypothesis but requires the cocycle to be $C^{1+\epsilon}$ and to be hyperbolic at some periodic point. In either case, we assume some form of irreducibility. A suitable extension of the Invariance Principle (Bonatti, Gómez-Mont and Viana [9], Avila and Viana [2]) that we prove here gives that these cocycles have non-zero Lyapunov exponents. We also prove that they are continuity points for the Lyapunov exponents, relative to the C^0 topology, and thus the claim follows.

The problem of continuity of Lyapunov exponents of linear cocycles was raised explicitly by Knill [19], and has been the object of considerable recent interest, especially with respective to finer topologies. See Viana [28, Chapter 10], Duarte and Klein [13] and references therein. It was conjectured by Viana [28] that Lyapunov exponents are always continuous among Hölder continuous fiber-bunched SL(2)-cocycles, and this has just been proved by Backes, Brown and Butler [3]. In fact, they prove a stronger conjecture also from Viana [28]: Lyapunov exponents vary continuously on any family of SL(2)-cocycles with continuous invariant holonomies. Improving on a construction of Bocker and Viana in [28, Chapter 9], Butler [12] also shows that the fiber-bunching condition is sharp for continuity in some cases.

These and many other related results require the cocycles to have some fair amount of regularity, starting from Hölder continuity. In view also of the Bochi–Mañé theorem, continuity in the C^0 topology (outside the uniformly hyperbolic realm) as we exhibit here comes as a bit of a surprise. It seems that the explanation should lie in the fact that existence of an invariant stable holonomy comes for free in the non-invertible case, but this requires further investigation.

Acknowledgement. The authors are grateful to the anonymous referee for a careful reading of the manuscript and several useful suggestions.

2. Statement of results

Let $f : M \to M$ be a continuous uniformly expanding map on a compact metric space. By this we mean that there are $\rho > 0$ and $\sigma > 1$ such that, for any $x \in M$,

(i) $d(f(x), f(y)) \geq \sigma d(x, y)$ for every $y \in B(x, \rho)$, and
(ii) $f(B(x, \rho))$ contains the closure of $B(f(x), \rho)$.

Take f to be topologically mixing and let μ be the equilibrium state of some Hölder continuous potential (see [29, Chapter 11]). Then μ is f-invariant and ergodic, and the support is the whole of$M$.

Let $\hat{M}$ be the space of sequences $\hat{x} = (x_{-n})_n$ such that $f(x_{-n}) = x_{-n+1}$ for every $n \geq 1$, and let $\pi : \hat{M} \to M$ denote the canonical map $\pi(\hat{x}) = x_0$. Consider the distance $\hat{d}$ defined on $\hat{M}$ by

$$\hat{d}(\hat{x}, \hat{y}) = \sum_{n=0}^{\infty} \tau^n d(x_{-n}, y_{-n}),$$

where $\tau > 0$ is a small constant. The **natural extension** of f is the map $\hat{f} : \hat{M} \to \hat{M}$ defined by

$$\hat{f}(\dots, x_{-n}, \dots, x_{-1}, x_0) = (\dots, x_{-n}, \dots, x_{-1}, x_0, f(x_0));$$

$\hat{f}$ is a hyperbolic homeomorphism ([27, Definition 1.3]) and satisfies $\pi \circ \hat{f} = f \circ \pi$.

For any $\hat{x} = (x_{-n})_n$ in $\hat{M}$, define the local stable set and local unstable set by

$$W^s_{\mathrm{loc}}(\hat{x}) = \{\hat{y} \in \hat{M} : \hat{d}(\hat{f}^n(\hat{x}), \hat{f}^n(\hat{y})) < \rho \text{ for every } n \geq 0\},$$
$$W^u_{\mathrm{loc}}(\hat{x}) = \{\hat{y} \in \hat{M} : \hat{d}(\hat{f}^{-n}(\hat{x}), \hat{f}^{-n}(\hat{y})) < \rho \text{ for every } n \geq 0\}.$$

Assuming τ is small enough, $W^s_{\mathrm{loc}}(\hat{x})$ coincides with the fiber $\pi^{-1}(x_0)$ and π maps $W^u_{\mathrm{loc}}(\hat{x})$ homeomorphically to

$$U_{\hat{x}} = \pi(W^u_{\mathrm{loc}}(\hat{x})),$$

with

$$B(x_0, 9\rho/10) \subset U_{\hat{x}} \subset B(x_0, \rho).$$

Moreover, each $\hat{V}_{\hat{x}} = \pi^{-1}(U_{\hat{x}})$ may be identified to the product

$$U_{\hat{x}} \times \pi^{-1}(x_0) \approx W^u_{\mathrm{loc}}(\hat{x}) \times W^s_{\mathrm{loc}}(\hat{x}) \tag{1}$$

through a homeomorphism, so that π becomes the projection to the first coordinate.

Let $\hat{\mu}$ be the lift of μ to $\hat{M}$, that is, the unique $\hat{f}$-invariant measure that projects down to μ under π. Then $\hat{\mu}$ is ergodic and supported on the whole of $\hat{M}$. Moreover, it has local product structure (see [10, Section 2.1]): the restriction of $\hat{\mu}$ to each $\hat{V}_{\hat{x}}$ may be written as

$$\hat{\mu} \mid \hat{V}_{\hat{x}} = \phi(\hat{\mu}^u \times \hat{\mu}^s), \tag{2}$$

where $\phi : \hat{V}_{\hat{x}} \to (0,\infty)$ is a continuous function, $\hat{\mu}^u = \mu \mid U_{\hat{x}}$ and $\hat{\mu}^s$ is a probability measure on $W^s_{\mathrm{loc}}(\hat{x})$. This means that $\hat{\mu} \mid \hat{V}_{\hat{x}}$ is equivalent to a product $\hat{\mu}^u \times \hat{\mu}^s$, with ϕ as the Radon–Nikodym density.

The **projective cocycle** defined by a continuous map $A : M \to \mathrm{SL}(2)$ over the transformation f is the map $F_A : M \times \mathbb{PR}^2 \to M \times \mathbb{PR}^2$,

$$F_A(x,v) = (f(x), A(x)v).$$

Denote $A^n(x) = A(f^{n-1}(x)) \cdots A(x)$ for every $n \geq 1$. By [15, 18], there exists $\lambda(A) \geq 0$, called the **Lyapunov exponent**, such that

$$\lim_n \frac{1}{n} \log \|A^n(x)\| = \lim_n \frac{1}{n} \log \|A^n(x)^{-1}\| = \lambda(A) \quad \text{for } \mu\text{-almost every } x \in M. \tag{3}$$

We say that A is u-**bunched** if there exists $\theta > 0$ such that A is θ-Hölder continuous and

$$\|A(x)\| \|A(x)^{-1}\| \, \sigma^{-\theta} < 1 \quad \text{for every } x \in M. \tag{4}$$

See [9, Definitions 1.11 and 2.2] and [1, Definition 2.2 and Remark 2.3].

Let $\hat{A} : \hat{M} \to \mathrm{SL}(2)$ be defined by $\hat{A} = A \circ \pi$. It is clear that $\hat{A}$ is θ-Hölder continuous with respect to $\hat{d}$ if A is θ-Hölder continuous with respect to d, because $\hat{d}(\hat{x},\hat{y}) \geq d(x_0,y_0)$ for any $\hat{x},\hat{y} \in \hat{M}$. Assuming that A is u-bunched, the cocycle $\hat{F}_A$ defined by $\hat{A}$ over $\hat{f}$ admits invariant u-holonomies (see [9, Section 1.4] and [1, Section 3]), namely,

$$h^u_{\hat{x},\hat{y}} = \lim_n \hat{A}^n(\hat{f}^{-n}(\hat{y}))\hat{A}^n(\hat{f}^{-n}(\hat{x}))^{-1} \quad \text{for any } \hat{y} \in W^u_{\mathrm{loc}}(\hat{x}).$$

As $\hat{A}$ is constant on local stable sets, $\hat{F}_A$ also admits trivial invariant s-holonomies:

$$h^s_{\hat{x},\hat{y}} = \mathrm{id} \quad \text{for any } \hat{y} \in W^s_{\mathrm{loc}}(\hat{x}).$$

A probability measure $\hat{m}$ on $\hat{M} \times \mathbb{PR}^2$ is said to be u-*invariant* if it admits a disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ such that

$$(h^u_{\hat{x},\hat{y}})_* \hat{m}_{\hat{x}} = \hat{m}_{\hat{y}} \quad \text{for any } \hat{y} \in W^u_{\mathrm{loc}}(\hat{x}). \tag{5}$$

Similarly, we say that $\hat{m}$ is s**-invariant** if it admits a disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ such that

$$\hat{m}_{\hat{x}} = \hat{m}_{\hat{y}} \quad \text{for any } \hat{y} \in W^s_{\text{loc}}(\hat{x}), \tag{6}$$

up to identifying $\{\hat{x}\} \times \mathbb{PR}^2$ with $\{\hat{y}\} \times \mathbb{PR}^2$ in the natural way. A u-invariant (respectively s-invariant) probability measure $\hat{m}$ is called a u**-state** (respectively, an s-state) if it is also invariant under $\hat{F}_A$. We call $\hat{m}$ an su**-state** (see [2, Section 4]) if it is both a u-state and an s-state.

Theorem A: *If A is u-bunched and has no su-states, then $\lambda(A) > 0$ and A is a continuity point for the function $B \mapsto \lambda(B)$ in the space of continuous maps $B : M \to \mathrm{SL}(2)$ equipped with the C^0 topology. In particular, the Lyapunov exponent $\lambda(\cdot)$ is bounded away from zero in a C^0-neighborhood of A.*

Example 2.1: Let $M = \mathbb{R}/\mathbb{Z}$ and $f : M \to M$, $f(x) = kx \mod \mathbb{Z}$, for some integer $k \geq 2$. Note that the natural extension $\hat{f} : \hat{M} \to \hat{M}$ is the usual (k-fold) solenoid. Let μ be the Lebesgue measure on M. Let $A : M \to \mathrm{SL}(2)$ be given by $A(x) = A_0 R_x$, where $A_0 \in \mathrm{SL}(2)$ is a hyperbolic matrix and R_x is the rotation by angle $2\pi x$; A is 1-Hölder continuous and, in view of the definition (4), it is u-bunched provided $k > \|A_0\| \|A_0^{-1}\|$.

We claim that A has no su-states if k is large enough. Indeed, suppose that $\hat{F}_A$ has some su-state $\hat{m}$. Then, by [2, Proposition 4.8], $\hat{m}$ admits a continuous disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ which is simultaneously s-invariant, u-invariant and $\hat{F}_A$-invariant. By s-invariance, we may write the disintegration as $\{\hat{m}_x : x \in M\}$ instead. Continuity and invariance under the dynamics imply that $(A_0)_* \hat{m}_0 = \hat{m}_0$. Since A_0 is hyperbolic, this means that m_0 is either a Dirac mass or a convex combination of two Dirac masses. Thus, by holonomy invariance, either every $\hat{m}_x$ is a Dirac mass or every $\hat{m}_x$ is supported on exactly 2 points.

In the first case, $\xi(x) = \operatorname{supp} \hat{m}_x$ defines a map $\xi : M \to \mathbb{PR}^2$ which is continuous and invariant under the cocycle:

$$\xi(f(x)) = A_0 R_x \xi(x) \quad \text{for every } x \in M.$$

It follows that the degree $\deg \xi$ satisfies $k \deg \xi = \deg \xi + 2$ (the term 2 comes from the fact that $M \to \mathbb{PR}^2$, $x \mapsto R_x v$ has degree 2 for any v). This is impossible when $k \geq 4$, and so this first case can be disposed of. This argument goes back to Herman [16]. In the second case, $\xi(x) = \operatorname{supp} \hat{m}_x$ defines a continuous invariant section with values in the space $\mathbb{P}\mathbb{R}^{2,2}$ of pairs of distinct points in $\mathbb{P}\mathbb{R}^2$. This can be reduced to the previous case by considering the 2-to-1 covering map $M \to M$, $z \mapsto 2z \mod \mathbb{Z}$ (notice that f is its own lift under this covering map). Thus, this second case can also be disposed of. This proves our claim that A has no su-states.

By Theorem A, it follows that $\lambda(B) > 0$ for every continuous $B : M \to \mathrm{SL}(2)$ in a C^0-neighborhood of A. Incidentally, this shows that [8, Corollary 12.34] is not correct: indeed, the "proof" assumes the Bochi–Mañé theorem in the non-invertible case.

Now let $f : M \to M$ be a $C^{1+\epsilon}$ (that is, C^1 with Hölder continuous derivative) expanding map on a compact manifold M and $A : M \to \mathrm{SL}(2)$ be a $C^{1+\epsilon}$ function. All the other objects, μ, F_A, π, $\hat{M}$, $\hat{f}$, $\hat{\mu}$, π, $\hat{A}$ and $\hat{F}_A$ are as before. An **invariant section** is a continuous map $\hat{\xi} : \hat{M} \to \mathbb{P}\mathbb{R}^2$ or $\hat{\xi} : \hat{M} \to \mathbb{P}\mathbb{R}^{2,2}$ such that

$$\hat{A}(\hat{x})\hat{\xi}(\hat{x}) = \hat{\xi}(\hat{f}(\hat{x})) \quad \text{for every } \hat{x} \in \hat{M}.$$

Theorem B: *If A admits no invariant section in $\mathbb{P}\mathbb{R}^2$ or $\mathbb{P}\mathbb{R}^{2,2}$, then it is a continuity point for the function $B \mapsto \lambda(B)$ in the space of continuous maps $B : M \to \mathrm{SL}(2)$ equipped with the C^0 topology. Moreover, $\lambda(A) > 0$ if and only if there exists some periodic point $p \in M$ such that $A^{\mathrm{per}(p)}(p)$ is a hyperbolic matrix. In that case, $\lambda(\cdot)$ is bounded from zero for all continuous cocycles in a C^0-neighborhood of A.*

This applies, in particular, in the setting of Young [31] and thus Theorem B contains a much stronger version of a main result in there: the Lyapunov exponent is C^0-continuous at every C^2 cocycle in the isotopy class; moreover, it is non-zero if and only if the cocycle is hyperbolic on some periodic orbit. It is clear that the latter condition is open, and the arguments in [10, Section 9] show that it is also dense.

All the cocycles we consider are of the form

$$\hat{F}_B(\hat{x}, v) = (\hat{f}(\hat{x}), B \circ \pi(\hat{x})v)$$

for some continuous $B : M \to \mathrm{SL}(2)$ and so they all have (trivial) s-holonomy. Thus the notion of s-invariant measure, as defined in (6), makes sense for such cocycles. In Section 3 we study certain properties of such measures. We do not assume the cocycle to be u-bunched, and so the conclusions apply for both theorems. In Section 4 we deduce Theorem A.

The key point to keep in mind is that, even though A is taken to be u-bunched—and so exhibits unstable holonomies—that need not be the case for C^0-nearby cocycles. Instead, we argue with s-states, taking advantage of the fact that (trivial) stable holonomies are always defined. The assumption that A has no su-states—which makes sense because A itself has unstable holonomies—implies that its s-state is unique, and thus continuity follows.

The proof of Theorem B is similar but necessarily more delicate: since we do not assume u-bunching at all, unstable holonomies may not exist for A, and thus the notion of su-state may not make sense. Indeed, most of the proof consists in bypassing this difficulty.

The first step is to explain what we mean by a u-invariant measure and a u-state. See Section 5. Next, we introduce a suitable version of the Invariance Principle of [2, 9, 21], where a Pesin unstable lamination is used to define an unstable holonomy on some full-measure subset. This is done in Section 6 and uses ideas of Tahzibi and Yang [26]. In Section 7, we check that the assumptions ensure that there are no su-states for A. This is rather subtle, since the Pesin lamination is only measurable. In Section 8 we wrap up the proof.

3. s-invariant measures

As before, let $f : M \to M$ be a topologically mixing uniformly expanding map on a compact metric space, and μ be the equilibrium state of some Hölder continuous potential. Let $\hat{f} : \hat{M} \to \hat{M}$ be the natural extension of f and $\hat{\mu}$ be the lift of μ to $\hat{M}$. Let $F_A : M \times \mathbb{PR}^2 \to M \times \mathbb{PR}^2$ be the projective cocycle defined by a continuous map $A : M \to \mathrm{SL}(2)$. Moreover, let $\hat{A} : \hat{M} \to \mathrm{SL}(2)$ and $\hat{F}_A : \hat{M} \times \mathbb{PR}^2 \to \hat{M} \times \mathbb{PR}^2$ be given by $\hat{A} = A \circ \pi$ and $\hat{F}_A(\hat{x}, v) = (\hat{f}(\hat{x}), \hat{A}(\hat{x})v))$ respectively. Here we do *not* assume A to be u-bunched.

Let $\mathcal{M}^s$ be the space of probability measures on $\hat{M} \times \mathbb{PR}^2$ that project down to $\hat{\mu}$ and are s-invariant. Let $\mathcal{M}$ be the space of probability measures on $M \times \mathbb{PR}^2$ that project down to μ. Both spaces are equipped with the weak* topology. Consider the map $\Psi : \mathcal{M} \to \mathcal{M}^s$ defined as follows: given any $m \in \mathcal{M}$ and a disintegration $\{m_x : x \in M\}$ along the fibers $\{x\} \times \mathbb{PR}^2$, let $\hat{m} = \Psi(m)$ be the measure on $\hat{M} \times \mathbb{PR}^2$ that projects down to $\hat{\mu}$ and whose conditional probabilities $\hat{m}_{\hat{x}}$ along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ are given by (up to the canonical identification of the fibers)

$$\hat{m}_{\hat{x}} = m_{\pi(\hat{x})}. \tag{7}$$

It is clear from the definition that $\hat{m} \in \mathcal{M}^s$. Moreover, if $\{m'_x : x \in M\}$ is another disintegration of m then, by essential uniqueness of the disintegration, $m_x = m'_x$ for μ-almost every x. Recalling that $\hat{\mu}$ is the lift of μ, this implies that $m_{\pi(\hat{x})} = m'_{\pi(\hat{x})}$ for $\hat{\mu}$-almost every $\hat{x}$. Thus $\hat{m}$ does not depend on the choice of the disintegration. This shows that Ψ is well-defined. We are going to prove:

Proposition 3.1: *$\Psi : \mathcal{M} \to \mathcal{M}^s$ is an affine homeomorphism.*

Proof. It is clear from (7) that Ψ is affine. To prove that it is a homeomorphism, we use the fact that $\hat{\mu}$ has local product structure, as follows.

For each $\hat{p} \in \hat{M}$, let $p = \pi(\hat{p})$ and consider the neighborhoods $U_{\hat{p}} = \pi(W^u_{\mathrm{loc}}(\hat{p}))$ of p in M and $\hat{V}_{\hat{p}} = \pi^{-1}(U_{\hat{p}})$ of $\hat{p}$ in $\hat{M}$. Recall (1)–(2). By local product structure,

$$\hat{\mu} \mid \hat{V}_{\hat{p}} = \phi(\hat{\mu}^u \times \hat{\mu}^s),$$

where $\phi : \hat{V}_{\hat{p}} \to (0, \infty)$ is a continuous function, $\hat{\mu}^u = \mu \mid U_{\hat{p}}$ and $\hat{\mu}^s$ is a probability measure on $W^s_{\mathrm{loc}}(\hat{p}) = \pi^{-1}(\hat{p})$.

Lemma 3.2: *For any $m \in \mathcal{M}$, the measure $\hat{m} = \Psi(m)$ satisfies*

$$\hat{m} \mid \hat{V}_{\hat{p}} \times \mathbb{PR}^2 = \tilde{\phi}(m \mid U_{\hat{p}} \times \mathbb{PR}^2) \times \hat{\mu}^s \quad \textit{for any } \hat{p} \in \hat{M},$$

where $\tilde{\phi} : \hat{V}_{\hat{p}} \to \mathbb{R}$ is defined by $\tilde{\phi}(x, \xi, v) = \phi(x, \xi)$.

Proof. Given any bounded measurable function $g : \hat{V}_{\hat{p}} \times \mathbb{PR}^2 \to \mathbb{R}$,

$$\begin{aligned}\int_{\hat{V}_{\hat{p}} \times \mathbb{PR}^2} g \, d\hat{m} &= \int_{\hat{V}_{\hat{p}}} \int_{\mathbb{PR}^2} g(\hat{x}, v) \, d\hat{m}_{\hat{x}}(v) \, d\hat{\mu}(\hat{x}) \\ &= \int_{W^s_{\mathrm{loc}}(\hat{p})} \int_{W^u_{\mathrm{loc}}(\hat{p})} \int_{\mathbb{PR}^2} g(x, \xi, v) \, d\hat{m}_{(x,\xi)}(v) \, \phi(x, \xi) \, d\hat{\mu}^u(x) \, d\hat{\mu}^s(\xi).\end{aligned}$$

Since $\hat{m}_{(x,\xi)} = m_x$ for every $x \in M$, by definition, it follows that

$$\begin{aligned}\int_{\hat{V}_{\hat{p}} \times \mathbb{PR}^2} g \, d\hat{m} &= \int_{W^s_{\mathrm{loc}}(\hat{p})} \int_{W^u_{\mathrm{loc}}(\hat{p})} \int_{\mathbb{PR}^2} g(x, \xi, v) \phi(x, \xi) \, dm_x(v) \, d\hat{\mu}^u(x) \, d\hat{\mu}^s(\xi) \\ &= \int_{W^s_{\mathrm{loc}}(\hat{p})} \int_{W^u_{\mathrm{loc}}(\hat{p}) \times \mathbb{PR}^2} g(x, \xi, v) \phi(x, \xi) \, dm(x, v) \, d\hat{\mu}^s(\xi).\end{aligned}$$

This proves the claim. ∎

Let us prove that Ψ is continuous, that is, that given any sequence $(m_n)_n$ converging to some m in $\mathcal{M}$, we have

$$\int_{\hat{M}\times\mathbb{PR}^2} g\,d\Psi(m_n) \to \int_{\hat{M}\times\mathbb{PR}^2} g\,d\Psi(m) \tag{8}$$

for every continuous function $g : \hat{M}\times\mathbb{PR}^2 \to \mathbb{R}$. It is no restriction to suppose that the support of g is contained in $\hat{V}_{\hat{p}}$ for some $\hat{p}\in\hat{M}$, for every continuous function is a finite sum of such functions. Then, by Lemma 3.2,

$$\int_{\hat{M}\times\mathbb{PR}^2} g\,d\Psi(m_n) = \int_{W^u_{\mathrm{loc}}(\hat{p})\times\mathbb{PR}^2}\int_{W^s_{\mathrm{loc}}(\hat{p})} g(x,\xi,v)\phi(x,\xi)\,d\hat{\mu}^s(\xi)\,dm_n(x,v).$$

Our hypotheses ensure that

$$G(x,v) = \int_{W^s_{\mathrm{loc}}(\hat{p})} g(x,\xi,v)\phi(x,\xi)\,d\hat{\mu}^s(\xi)$$

defines a continuous function. Hence, the assumption that $(m_n)\to m$ implies that

$$\begin{aligned}\int_{\hat{M}\times\mathbb{PR}^2} g\,d\Psi(m_n) &= \int_{W^u_{\mathrm{loc}}(\hat{p})\times\mathbb{PR}^2} G(x,v)\,dm_n(x,v)\\ &\to \int_{W^u_{\mathrm{loc}}(\hat{p})\times\mathbb{PR}^2} G(x,v)\,dm(x,v) = \int_{\hat{M}\times\mathbb{PR}^2} g\,d\Psi(m),\end{aligned}$$

as claimed. Since $\mathcal{M}$ is compact, we are left to proving that Ψ is a bijection.

Surjectivity is clear: given $\hat{m}\in\mathcal{M}^s$, take m to be the probability measure on $M\times\mathbb{PR}^2$ that projects down to μ and whose conditional probabilities along the vertical fibers $\{x\}\times\mathbb{PR}^2$ are given by $m_x = \hat{m}_{\hat{x}}$ for any $\hat{x}\in\pi^{-1}(x)$. This is well defined, by (6), and it is clear from the definition that $\Psi(m)=\hat{m}$. Injectivity is a consequence of Lemma 3.2. Indeed, if $\Psi(m_1)=\Psi(m_2)$ then

$$\int_{X\times V}\int_{W^s_{\mathrm{loc}}(\hat{p})} \phi(x,\xi)\,d\hat{\mu}^s(\xi)\,dm_1(x,v) = \int_{X\times V}\int_{W^s_{\mathrm{loc}}(\hat{p})} \phi(x,\xi)\,d\hat{\mu}^s(\xi)\,dm_2(x,v)$$

for any $\hat{p}\in\hat{M}$ and any $X\times V\subset U_{\hat{p}}\times\mathbb{PR}^2$. This implies that

$$Hm_1 \mid U_{\hat{p}} = Hm_2 \mid U_{\hat{p}}, \quad \text{where } H(x) = \int_{W^s_{\mathrm{loc}}(\hat{p})} \phi(x,\xi)\,d\hat{\mu}^s(\xi).$$

Noting that H is positive, it follows that the restrictions of m_1 and m_2 to $U_{\hat{p}}$ coincide, for every $\hat{p}\in\hat{M}$. Thus $m_1=m_2$. ∎

COROLLARY 3.3: *$\mathcal{M}^s$ is non-empty, convex and compact. Moreover, it contains some s-state.*

Proof. The first part follows directly from Proposition 3.1, since $\mathcal{M}$ is non-empty, convex and compact.

If $\hat{m}$ projects down to $\hat{\mu}$, then so does $(\hat{F}_A^{-1})_*\hat{m}$, because $\hat{F}_A^{-1}$ is given by $\hat{f}^{-1}$ in the first coordinate, and $\hat{f}^{-1}$ preserves $\hat{\mu}$. It is also clear that if $\hat{m}$ is s-invariant, then so is $\hat{F}_{A,*}^{-1}\hat{m}$: its conditional measures are given by

$$(\hat{F}_{A,*}^{-1}\hat{m})_{\hat{x}} = \hat{A}(\hat{x})_*^{-1}\hat{m}_{\hat{f}(\hat{x})} = A(\pi(\hat{x}))_*^{-1}\hat{m}_{\hat{f}(\hat{x})},$$

and, keeping in mind that $\pi(\hat{f}(\hat{x})) = f(\pi(\hat{x}))$, the assumption ensures that the right hand side depends only on $\pi(\hat{x})$. This proves that $\mathcal{M}^s$ is invariant under $\hat{F}_A^{-1}$.

Now let $\hat{m}_0$ be an arbitrary element of $\mathcal{M}^s$. By invariance and compactness, the sequence $n^{-1}\sum_{j=0}^{n-1}\hat{F}_{A,*}^{-j}\hat{m}_0$ has some accumulation point in $\mathcal{M}^s$. It is straightforward to see that every accumulation point is $\hat{F}_A$-invariant and, thus, an s-state. ∎

Let $(B_n)_n$ be a sequence of maps converging uniformly to some B in the space of continuous maps $M \to \mathrm{SL}(2)$, and $(\hat{m}_n)_n$ be a sequence of probability measures on $\hat{M}$ converging in the weak* topology to some probability measure m.

Corollary 3.4: *If $\hat{m}_n$ is an s-state of B_n for every n, then $\hat{m}$ is an s-state of B.*

Proof. It follows from Corollary 3.3 that $\hat{m} \in \mathcal{M}^s$. It is clear that $\hat{m}$ is $\hat{F}_B$-invariant, because m_n is $\hat{F}_{B_n}$-invariant for every n and $\hat{F}_{B_n}$ converges uniformly to $\hat{F}_B$. ∎

4. Proof of Theorem A

If $\lambda(A) = 0$ then, by the Invariance Principle ([2, Theorem D], [9, Théorème 1]), every $\hat{F}_A$-invariant probability measure $\hat{m}$ that projects down to $\hat{\mu}$ is an su-state. Thus, the hypothesis that A has no su-states implies that $\lambda(A) > 0$.

We are left to proving that A is a continuity point for the Lyapunov exponent. Define (here v denotes both a direction and any non-zero vector in that direction)

$$\phi_B : \hat{M} \times \mathbb{PR}^2 \to \mathbb{R}, \quad \phi_B(\hat{x}, v) = \log \frac{\|\hat{B}(\hat{x})v\|}{\|v\|}.$$

The next statement stems from a classical observation of Furstenberg [14]; see, for instance, [28, Section 6.1].

Proposition 4.1: *Every continuous $B : M \to \mathrm{SL}(2)$ admits some s-state $\hat{m}_B$ such that*

$$-\lambda(B) = \int_{\hat{M}\times\mathbb{PR}^2} \phi_B \, d\hat{m}.$$

Proof. First, suppose that $\lambda(B) = 0$. For every $(\hat{x}, v) \in \hat{M} \times \mathbb{PR}^2$ and $n \geq 1$,

$$\sum_{j=0}^{n-1} \phi_B(\hat{F}_B^j(\hat{x}, v)) = \log \frac{\|\hat{B}^n(\hat{x})v\|}{\|v\|} \in [-\log \|\hat{B}^n(\hat{x})^{-1}\|, \log \|\hat{B}^n(\hat{x})\|].$$

We also have that, for $\hat{\mu}$-almost every $\hat{x} \in M$,

$$\lim_n \frac{1}{n} \log \|\hat{B}^n(\hat{x})\| = \lim_n \frac{1}{n} \log \|\hat{B}^n(\hat{x})^{-1}\| = \lambda(B).$$

Thus, for any $\hat{F}_B$-invariant measure $\hat{m}$ that projects down to $\hat{\mu}$,

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \phi_B(\hat{F}_B^j(\hat{x}, v)) = 0 \quad \text{for } \hat{m}\text{-almost every } (\hat{x}, v).$$

By the ergodic theorem, this implies that

$$\int_{\hat{M}\times\mathbb{PR}^2} \phi_B \, d\hat{m} = 0 = \lambda(B).$$

By Corollary 3.3, we may choose $\hat{m}$ to be an s-state, in which case it satisfies the conclusion of the lemma.

Now suppose that $\lambda(B) > 0$. By the theorem of Oseledets [24], there exists an $\hat{F}_B$-invariant splitting $\mathbb{R}^2 = E^u_{\hat{x}} \oplus E^s_{\hat{x}}$ defined at $\hat{\mu}$-almost every point $\hat{x}$ and such that

$$\begin{aligned} &\lim_{n\to\pm\infty} \frac{1}{n} \log \|\hat{B}^n(\hat{x})v\| = \lambda(B) && \text{for non-zero } v \in E^u_{\hat{x}} \quad \text{and} \\ &\lim_{n\to\pm\infty} \frac{1}{n} \log \|\hat{B}^n(\hat{x})v\| = -\lambda(B) && \text{for non-zero } v \in E^s_{\hat{x}}. \end{aligned} \tag{9}$$

Let $\hat{m}$ be the probability measure on $\hat{M} \times \mathbb{PR}^2$ that projects down to $\hat{\mu}$ and whose conditional probabilities along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ are given by the Dirac masses at $E^s_{\hat{x}}$. Then $\hat{m}$ is an s-state: it is clear that it is $\hat{F}_B$-invariant; to check that it is s-invariant, just note that the subspace $E^s_{\hat{x}}$ depends only on the forward iterates, and so it is constant on each $\pi^{-1}(x)$. Moreover, by the ergodic theorem and (9),

$$\begin{aligned}\int_{\hat{M}\times\mathbb{PR}^2} \phi_B \, d\hat{m} &= \int_{\hat{M}\times\mathbb{PR}^2} \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \phi_B \circ \hat{F}_B^j \, d\hat{m} \\ &= \int_{\hat{M}} \int_{\mathbb{PR}^2} \lim_n \frac{1}{n} \log \frac{\|\hat{B}^n(\hat{x})v\|}{\|v\|} \, d\delta_{E^s_{\hat{x}}}(v) \, d\hat{\mu}(\hat{x}) \\ &= \int_{\hat{M}} -\lambda(B) \, d\hat{\mu}(\hat{x}) = -\lambda(B).\end{aligned}$$

This completes the proof. ∎

Lemma 4.2: *If A has no su-states, then it has exactly one s-state.*

Proof. Existence is contained in Proposition 4.1. To prove uniqueness, we argue as follows. Let $\hat{m}$ be any s-state. As observed before, the hypothesis implies that $\lambda(A) > 0$. Let $\mathbb{R}^2 = E^u_{\hat{x}} \oplus E^s_{\hat{x}}$ be the Oseledets invariant splitting, defined at $\hat{\mu}$-almost every point $\hat{x}$. Let $\hat{m}^u$ and $\hat{m}^s$ be the probability measures on $\hat{M} \times \mathbb{PR}^2$ that project down to $\hat{\mu}$ and whose conditional probabilities along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ are the Dirac masses at $E^u_{\hat{x}}$ and $E^s_{\hat{x}}$, respectively. Then $\hat{m}^u$ is a u-state, $\hat{m}^s$ is an s-state and every $\hat{F}_A$-invariant probability measure is a convex combination of $\hat{m}^u$ and $\hat{m}^s$ (compare [2, Lemma 6.1]). Then, keeping in mind that $\hat{\mu}$ is ergodic, there is $\alpha \in [0, 1]$ such that $\hat{m} = \alpha\hat{m}^u + (1 - \alpha)\hat{m}^s$. If $\alpha > 0$, we may write

$$\hat{m}^u = \frac{1}{\alpha}\hat{m} + \Big(1 - \frac{1}{\alpha}\Big)\hat{m}^s$$

and, as $\hat{m}$ and $\hat{m}^s$ are s-states, it follows that $\hat{m}^u$ is an s-state. Since $\hat{m}^u$ is also a u-state, this contradicts the hypothesis. Thus $\alpha = 0$, that is, $m = m^s$. ∎

Theorem A is an easy consequence. Indeed, we already know that $\lambda(A) > 0$. Consider any sequence $A_k : M \to \mathrm{SL}(2)$, $k \in \mathbb{N}$ converging to A in the C^0 topology. By Proposition 4.1, for each k one can find some s-state $\hat{m}_k$ for A_k such that

$$-\lambda(A_k) = \int_{\hat{M}\times\mathbb{PR}^2} \phi_{A_k} \, d\hat{m}_k.$$

Up to restricting to a subsequence, we may suppose that $(\hat{m}_k)_k$ converges to some probability measure $\hat{m}$ in the weak* topology. By Corollary 3.4, $\hat{m}$ is an s-state for A. Moreover, since ϕ_{A_k} converges uniformly to ϕ_A,

$$\lim_k -\lambda(A_k) = \int_{\hat{M}\times\mathbb{PR}^2} \phi_A \, d\hat{m}. \tag{10}$$

By Proposition 4.1 and Lemma 4.2, the right-hand side is equal to $\lambda(A)$. This proves continuity of the Lyapunov exponent in the C^0 topology.

Remark 4.3: The converse to Lemma 4.2 is true when $\lambda(A) > 0$.

5. u-states without u-bunching

Next we prove Theorem B. Initially, suppose that $0 \leq \lambda(A) < \log \sigma$. Then the cocycle is "nonuniformly u-bunched," in a sense that was exploited before, in [27, Sections 2.1 and 2.2]. Using those methods, one gets that (compare [27, Proposition 2.5]) unstable holonomy maps

$$h^u_{\hat{x},\hat{y}} : \{\hat{x}\} \times \mathbb{PR}^2 \to \{\hat{y}\} \times \mathbb{PR}^2, \quad h^u_{\hat{x},\hat{y}} = \lim_n \hat{A}^n(\hat{f}^{-n}(\hat{y}))\hat{A}^n(\hat{f}^{-n}(\hat{x}))^{-1}$$

exist for $\hat{\mu}$-almost every $\hat{x}$ and any $\hat{y} \in W^u_{\mathrm{loc}}(\hat{x})$. Then we define a probability measure $\hat{m}$ on $\hat{M} \times \mathbb{PR}^2$ to be **u-invariant** if it admits a disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ such that

$$(h^u_{\hat{x},\hat{y}})_* \hat{m}_{\hat{x}} = \hat{m}_{\hat{y}} \quad \text{for } \hat{\mu}\text{-almost every } \hat{x} \text{ and any } \hat{y} \in W^u_{\mathrm{loc}}(\hat{x}). \tag{11}$$

As before, a **u-state** is an $\hat{F}_A$-invariant probability measure which is u-invariant.

When $\lambda(A) \geq \log \sigma$ we have to restrict ourselves to the subclass of $\hat{F}_A$-invariant probability measures whose center Lyapunov exponent is strictly less than σ. More precisely, we consider only $\hat{F}_A$-invariant probability measures $\hat{m}$ such that

$$\lim_n \frac{1}{n} \log \|D\hat{A}^n(\hat{x})v\| \leq c < \log \sigma \quad \text{for } \hat{m}\text{-almost every } (\hat{x}, v) \in \hat{M} \times \mathbb{PR}^2, \tag{12}$$

where $D\hat{A}(\hat{x})v$ denotes the derivative of the projective map $\hat{A}(\hat{x}) : \mathbb{PR}^2 \to \mathbb{PR}^2$.

Remark 5.1: The following elementary bound will be useful:

$$\|\hat{A}(\hat{x})\|^{-1}\|\hat{A}(\hat{x})^{-1}\|^{-1} \leq \frac{\|D\hat{A}(\hat{x})v\|}{\|v\|} \leq \|\hat{A}(\hat{x})\|\|\hat{A}(\hat{x})^{-1}\| \quad \text{for every } \hat{x}.$$

In the next result we use the fact that the natural extension of f admits a $C^{1+\epsilon}$ realization: there exist a $C^{1+\epsilon}$ embedding $g : U \to U$ defined on some open subset U of an Euclidean space, and a topological embedding $\iota : \hat{M} \to U$ with $g(\iota(\hat{M})) = \iota(\hat{M})$ and $g \circ \iota = \iota \circ \hat{f}$. A proof is given in Appendix A. Identifying $\hat{M}$ with $\iota(\hat{M})$ we may view $\hat{f}$ as a restriction of g, and so we may apply Pesin theory to it.

PROPOSITION 5.2: *If $\hat{m}$ satisfies* (12)*, then for $(\hat{x}, v)$ in a full $\hat{m}$-measure subset Λ of $\hat{M} \times \mathbb{PR}^2$ there exists a C^1 function $\psi_{\hat{x},v} : W^u_{\mathrm{loc}}(\hat{x}) \to \mathbb{PR}^2$ depending measurably on $(\hat{x}, v)$ such that $\psi_{\hat{x},v}(\hat{x}) = v$ and the graphs*

$$\mathcal{W}^u_{\mathrm{loc}}(\hat{x}, v) = \{(\hat{y}, \psi_{\hat{x},v}(\hat{y})) : \hat{y} \in W^u_{\mathrm{loc}}(\hat{x})\}$$

satisfy

(a) $\hat{F}^{-1}(\mathcal{W}^u_{\mathrm{loc}}(\hat{x}, v)) \subset \mathcal{W}^u_{\mathrm{loc}}(\hat{F}^{-1}(\hat{x}, v))$ *for every* $(\hat{x}, v) \in \Lambda$;
(b) $d(\hat{F}^{-n}(\hat{x}, v), \hat{F}^{-n}(\hat{y}, w)) \to 0$ *exponentially fast for any* $(\hat{y}, w) \in \mathcal{W}^u_{\mathrm{loc}}(\hat{x}, v)$.

Proof. The assumption ensures that there exists $\hat{m}$-almost everywhere an Oseledets strong-unstable subspace $\hat{E}^u_{\hat{x},v} \subset T_{\hat{x}}U \times \mathbb{R}^2$ that is a graph over the unstable direction $E_{\hat{x}} \subset T_{\hat{x}}U$ of g. Then, by Pesin theory, there exists $\hat{m}$-almost everywhere a C^1 embedded disk $\widetilde{W}^u_{\mathrm{loc}}(\hat{x}, v)$ tangent to $\hat{E}^u_{\hat{x},v}$ and such that

$$\hat{F}^{-n}(\hat{y}, w) \in \widetilde{W}^u_{\mathrm{loc}}(\hat{F}^{-n}(\hat{x}, v)) \quad \text{and} \quad d(\hat{F}^{-n}(\hat{x}, v), \hat{F}^{-n}(\hat{y}, w)) \leq \sigma^{-n}$$

for every $n \geq 0$ and $(\hat{y}, w) \in \widetilde{W}^u_{\mathrm{loc}}(\hat{x}, v)$. This also implies that $\widetilde{W}^u_{\mathrm{loc}}(\hat{x}, v)$ is a C^1 graph over a neighborhood of $\hat{x}$ inside $W^u(\hat{x})$. While the radius $r(\hat{x})$ of this neighborhood need not be bounded from zero, in principle, Pesin theory also gives that it decreases sub-exponentially along orbits:

$$\lim_n \frac{1}{n} \log r(\hat{f}^{-n}(\hat{x})) = 0.$$

On the other hand, the size of $\hat{f}^{-n}(W^u_{\mathrm{loc}}(\hat{x}))$ decreases exponentially fast (faster than σ^{-n}). Thus, the projection of $\widetilde{W}^u_{\mathrm{loc}}(\hat{F}^{-n}(\hat{x}, v))$ contains $\hat{f}^{-n}(W^u_{\mathrm{loc}}(\hat{x}))$ for any large n. Then $\hat{F}^n(\widetilde{W}^u_{\mathrm{loc}}(\hat{F}^{-n}(\hat{x}, v)))$ is a C^1 graph whose projection contains $W^u_{\mathrm{loc}}(\hat{x})$. Take $\mathcal{W}^u_{\mathrm{loc}}(\hat{x}, v)$ to be the part of that graph that lies over $W^u_{\mathrm{loc}}(\hat{x})$. It is clear from the construction that conditions (a) and (b) in the statement are satisfied. ∎

This means that unstable holonomy maps are still defined on suitable full measure subsets of the fibers, namely $h^u_{\hat{x},\hat{y}} : \Lambda_{\hat{x}} \to \{\hat{y}\} \times \mathbb{PR}^2$ where

$$\Lambda_{\hat{x}} = \Lambda \cap (\{\hat{x}\} \times \mathbb{PR}^2).$$

Note that $\hat{m}_{\hat{x}}(\Lambda_{\hat{x}}) = 1$ for $\hat{\mu}$-almost every $\hat{x}$, because $\hat{m}(\Lambda) = 1$. Then we can extend the definition of u-invariance to this setting: we say that $\hat{m}$ is ***u*-invariant** if it admits a disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ along the fibers $\{\hat{x}\} \times \mathbb{PR}^2$ satisfying (11). We can also extend the notion of ***u*-state**: in the present setting it means that $\hat{m}$ is $\hat{F}_A$-invariant, satisfies (12) and is u-invariant.

6. A new u-invariance principle

Here we prove the following form of the Invariance Principle (Ledrappier [21], Bonatti, Gómez-Mont and Viana [9], Avila and Viana [2]) where the main novelty is that no u-bunching is assumed:

THEOREM 6.1: *Every $\hat{F}_A$-invariant probability measure $\hat{m}$ satisfying*

$$\lim_n \frac{1}{n} \log \|D\hat{A}^n(\hat{x})v\| \leq 0 \quad \text{for } \hat{m}\text{-almost every } (\hat{x}, v) \in \hat{M} \times \mathbb{PR}^2, \tag{13}$$

is a u-state.

We are going to extend to our setting an approach introduced by Tahzibi and Yang [26] for bunched cocycles. This is based on the notion of partial entropy, which may be defined as follows (see [20, 30] for more information).

Let $\mathcal{R}$ be a Markov partition of $\hat{f}$, in the sense of Bowen [11, Section 3.C], with diameter small enough that the element $\mathcal{R}(\hat{x})$ that contains $\hat{x}$ is a subset of $\hat{V}_{\hat{x}}$ for every $\hat{x} \in \hat{M}$. (Actually, elements of $\mathcal{R}$ may intersect along their boundaries but, since the boundaries are nowhere dense and have zero $\hat{\mu}$-measure, we may ignore the trajectories through them.) Let $\xi^u(\hat{x}) \subset \hat{M}$ be the connected component of $\mathcal{R}(\hat{x}) \cap W^u_{\mathrm{loc}}(\hat{x})$ that contains $\hat{x}$. Take Λ as in Proposition 5.2. For $v \in \mathbb{PR}^2$ such that $(\hat{x}, v) \in \Lambda$, let $\Xi^u(\hat{x}, v)$ be the connected component of $(\mathcal{R}(x) \times \mathbb{PR}^2) \cap \mathcal{W}^u_{\mathrm{loc}}(\hat{x}, v)$ that contains $(\hat{x}, v)$.

It follows directly from the Markov property of $\mathcal{R}$ that the family ξ^u is an **adapted partition** for $(\hat{f}, \hat{\mu})$: its elements are pairwise disjoint and, for $\hat{\mu}$-almost every $\hat{x}$,

- $\hat{f}^{-1}(\xi^u(\hat{x})) \subset \xi^u(\hat{f}^{-1}(\hat{x}))$ (thus $\xi^u \prec \hat{f}^{-1}\xi^u$), and
- $\xi^u(\hat{x})$ contains a neighborhood of $\hat{x}$ inside $W^u(\hat{x})$.

Analogously, Ξ^u is an adapted partition for $(\hat{F}, \hat{m})$. The corresponding **partial entropies** are defined by

$$h_{\hat{\mu}}(\hat{f}, W^u) = H_{\hat{\mu}}(\hat{f}^{-1}\xi^u \mid \xi^u) \quad \text{and} \quad h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u) = H_{\hat{m}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u). \tag{14}$$

Recall (from Rokhlin [25]) that if ξ and η are measurable partitions of a Lebesgue space (Z, ζ) satisfying $\xi \prec \eta$, then

$$H_\zeta(\eta) = \int -\log \zeta(\eta(x))\, d\zeta(x) \quad \text{and} \quad H_\zeta(\eta \mid \xi) = \int H_{\zeta_z}(\eta)\, d\zeta(z)$$

where $\{\zeta_z : z \in Z\}$ is a disintegration of ζ with respect to ξ.

6.1. c-INVARIANT MEASURES. Let $\{\hat\mu^u_{\hat x} : \hat x \in \hat M\}$ and $\{\hat m^u_{\hat x,v} : (\hat x, v) \in \hat M \times \mathbb{P}\mathbb{R}^2\}$ be disintegrations of, respectively, $\hat\mu$ relative to ξ^u and $\hat m$ relative to Ξ^u. Let $p : \hat M \times \mathbb{P}\mathbb{R}^2 \to \hat M$ be the canonical projection. We call $\hat m$ **c-invariant** if

$$(h^c_{\hat x,v,w})_* \hat m^u_{\hat x,v} = \hat m^u_{\hat x,w} \quad \text{for } \hat m\text{-almost every } (\hat x, v) \text{ and } (\hat x, w), \tag{15}$$

where $h^c_{\hat x,v,w} : \Xi^u(\hat x, v) \to \Xi^u(\hat x, w)$ is the bijection defined by

$$p \circ h^c_{\hat x,v,w} = p.$$

Equivalently, the measure $\hat m$ is c-invariant if

$$p_*(\hat m^u_{\hat x,v}) = \hat\mu^u_{\hat x} \quad \text{for } \hat m\text{-almost every } (\hat x, v). \tag{16}$$

PROPOSITION 6.2: *The measure $\hat m$ is u-invariant if and only if it is c-invariant.*

Proof. Let us start with a model: let ν be a probability measure on a product $X \times Y$ of two separable metric spaces, and let $\{\nu^1_y : y \in Y\}$ and $\{\nu^2_x : x \in X\}$ be disintegrations of ν relative to the partition into horizontals $X \times \{y\}$ and the partition into verticals $\{x\} \times Y$, respectively. We call ν **v-invariant** (respectively, **h-invariant**) if the disintegrations may be chosen such that ν^1_y is independent of y (respectively, ν^2_x is independent of x).

LEMMA 6.3: *ν is v-invariant if and only if it is h-invariant.*

Proof. Suppose that ν is v-invariant and let ν^1 be such that $\nu^1_y = \nu^1$ for every y. Let ν^2 be the quotient of ν relative to the horizontal partition or, equivalently, the projection of ν to the second coordinate. Then, by the definition of disintegration,

$$\nu = \nu^1 \times \nu^2.$$

This implies that ν^1 is the projection of ν to the first coordinate and $\nu^2_x = \nu_2$ defines a disintegration of ν relative to the vertical partition. In particular, ν is h-invariant. The proof that h-invariance implies v-invariance is identical. ∎

To deduce the proposition we only have to reduce its setting to that of Lemma 6.3. Consider the partitions Ξ^c and Ξ^{cu} of $\hat M \times \mathbb{P}\mathbb{R}^2$ defined by

$$\Xi^c(\hat x, v) = p^{-1}(\hat x) \quad \text{and} \quad \Xi^{cu}(\hat x, v) = p^{-1}(\xi^u(\hat x)).$$

Let $\{\hat m^c_{\hat x,v} : (\hat x, v) \in \hat M \times \mathbb{P}\mathbb{R}^2\}$ and $\{\hat m^{cu}_{\hat x,v} : (\hat x, v) \in \hat M \times \mathbb{P}\mathbb{R}^2\}$ be disintegrations of $\hat m$ relative to Ξ^c and Ξ^{cu}, respectively. Both Ξ^c and Ξ^u refine Ξ^{cu}. So, by essential uniqueness of the disintegration,

(i) $\{\hat{m}^u_{\hat{y},w} : (\hat{y}, w) \in \Xi^{cu}(\hat{x}, v)\}$ is a disintegration of $\hat{m}^{cu}_{\hat{x},v}$ with respect to the partition $\Xi^u \mid \Xi^{cu}(\hat{x}, v)$, and
(ii) $\{\hat{m}^c_{\hat{y},w} : (\hat{y}, w) \in \Xi^{cu}(\hat{x}, v)\}$ is a disintegration of $\hat{m}^{cu}_{\hat{x},v}$ with respect to the partition $\Xi^c \mid \Xi^{cu}(\hat{x}, v)$,

for $\hat{m}$-almost every $(\hat{x}, v)$. This will be used a few times in the following.

Now consider the map

$$\Psi_{\hat{x},v} : \Xi^{cu}(\hat{x}, v) \to \xi^u(\hat{x}) \times \mathbb{PR}^2, \quad \Phi_{\hat{x}}(\hat{y}, w) = (\hat{y}, z)$$

where z is such that $(\hat{x}, z)$ is the point where $\Xi^u(\hat{y}, w)$ intersects $\Xi^c(\hat{x}, v)$. Since Λ has full $\hat{m}$-measure, $\Psi_{\hat{x},v}$ is defined $\hat{m}^{cu}_{\hat{x},v}$-almost everywhere for $\hat{m}$-almost every $(\hat{x}, v)$. Clearly, it is an invertible measurable map sending atoms of $\Xi^u \mid \Xi^{cu}(\hat{x}, v)$ to horizontals $\xi^u(\hat{x}) \times \{z\}$ and atoms of $\Xi^c \mid \Xi^{cu}(\hat{x}, v)$ to verticals $\{\hat{y}\} \times \mathbb{PR}^2$.

Identifying $\Xi^{cu}(\hat{x}, v)$ to $\xi^u(\hat{x}) \times \mathbb{PR}^2$ through $\Psi_{\hat{x},v}$, (i) and (ii) above correspond to disintegrations of $\hat{m}_{\hat{x},v}$ relative to the horizontal partition and the vertical partition, respectively. Moreover, s-invariance and u-invariance translate to v-invariance and h-invariance, respectively. Thus the claim follows from Lemma 6.3. ∎

6.2. A criterion for c-invariance. Note that $h_{\hat{\mu}}(\hat{f}) \le h_{\hat{m}}(\hat{F}_A)$, because $(\hat{f}, \hat{\mu})$ is a factor of $(\hat{F}_A, \hat{m})$. For the partial entropies the inequality goes in the opposite direction:

Proposition 6.4: *$h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u) \le h_{\hat{\mu}}(\hat{f}, W^u)$ and the equality holds if and only if $\hat{m}$ is c-invariant.*

Proof. Keep in mind that $\xi^u \prec \hat{f}^{-1}\xi^u$ and $\Xi^u \prec \hat{F}_A^{-1}\Xi^u$. By definition,

$$(17)\qquad \begin{aligned} h_{\hat{\mu}}(\hat{f}, W^u) = H_{\hat{\mu}}(\hat{f}^{-1}\xi^u \mid \xi^u) &= \int H_{\hat{\mu}^u_{\hat{x}}}(\hat{f}^{-1}\xi^u)\, d\hat{\mu}(\hat{x}) \quad \text{where} \\ H_{\hat{\mu}^u_{\hat{x}}}(\hat{f}^{-1}\xi^u) &= \int -\log \hat{\mu}^u_{\hat{x}}((\hat{f}^{-1}\xi^u)(\hat{y}))\, d\hat{\mu}^u_{\hat{x}}(\hat{y}), \end{aligned}$$

and similarly for $h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u)$ and Ξ^u.

Lemma 6.5: *For $\hat{m}$-almost every $(\hat{x}, v) \in \hat{M} \times \mathbb{PR}^2$,*

(a) *$H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u) \le H_{\hat{\mu}^u_{\hat{x}}}(\hat{f}^{-1}\xi^u)$, and*
(b) *the equality holds if and only if $\hat{m}^u_{\hat{x},v}((\hat{F}_A^{-1}\Xi^u)(\hat{y}, w)) = \hat{\mu}^u_{\hat{x}}((\hat{f}^{-1}\xi^u)(\hat{y}))$ for $\hat{m}^{cu}_{\hat{x},v}$-almost every $(\hat{y}, w) \in \Xi^{cu}(\hat{x}, v)$.*

Proof. Since $\Xi^{cu}(\hat{x}, v) = p^{-1}(\xi^u(\hat{x}))$, $(\hat{F}_A^{-1}\Xi^{cu})(\hat{x}, v) = p^{-1}((\hat{f}^{-1}\xi^u)(\hat{x}))$ and $\hat{\mu} = p_*\hat{m}$, essential uniqueness of disintegrations gives that $\hat{\mu}^u_{\hat{x}} = p_*(\hat{m}^{cu}_{\hat{x},v})$ for $\hat{m}$-almost every $(\hat{x}, v)$. Thus

$$\begin{aligned} H_{\hat{\mu}^u_{\hat{x}}}(\hat{f}^{-1}\xi^u) &= \int -\log \hat{\mu}^u_{\hat{x}}((\hat{f}^{-1}\xi^u)(\hat{y}))\, d\hat{\mu}^u_{\hat{x}}(\hat{y}) \\ &= \int -\log \hat{m}^{cu}_{\hat{x},v}((\hat{F}_A^{-1}\Xi^{cu})(\hat{y}, w))\, d\hat{m}^{cu}_{\hat{x},v}(\hat{y}, w) = H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^{cu}) \end{aligned}$$

for $\hat{m}$-almost every $(\hat{x}, v)$. Moreover, using the relation $\hat{F}_A^{-1}\Xi^{cu} \vee \Xi^u = \hat{F}_A^{-1}\Xi^u$,

$$H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^{cu}) \geq H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^{cu} \mid \Xi^u) = H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u).$$

This proves claim (a). Moreover, the equality holds if and only if the partitions $\hat{F}_A^{-1}\Xi^{cu}$ and Ξ^u are independent relative to $\hat{m}^{cu}_{\hat{x},v}$, that is,

$$\hat{m}^u_{\hat{x},v}((\hat{F}_A^{-1}\Xi^{cu})(\hat{y}, w)) = \hat{m}^{cu}_{\hat{x},v}((\hat{F}_A^{-1}\Xi^{cu})(\hat{y}, w)) \quad \text{for } \hat{m}^{cu}_{\hat{x},v}\text{-almost every } (\hat{y}, w).$$

By the previous observations, this is equivalent to

$$\hat{m}^u_{\hat{x},v}((\hat{F}_A^{-1}\Xi^u)(\hat{y}, w)) = \hat{\mu}^u_{\hat{x}}((\hat{f}^{-1}\xi^u)(\hat{y})) \quad \text{for } \hat{m}^{cu}_{\hat{x},v}\text{-almost every } (\hat{y}, w),$$

as claimed in (b). ∎

Similarly to (17), we have $H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u) = \int H_{\hat{m}^u_{\hat{y},w}}(\hat{F}_A^{-1}\Xi^u)\, d\hat{m}^{cu}_{\hat{x},v}(\hat{y}, w)$. So, integrating the inequality in part (a) of the lemma,

$$\begin{aligned} H_{\hat{m}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u) &= \int H_{\hat{m}^u_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^u)\, d\hat{m}(\hat{x}, v) = \int H_{\hat{m}^{cu}_{\hat{x},v}}(\hat{F}_A^{-1}\Xi^u \mid \Xi^u)\, d\hat{m}(\hat{x}, v) \\ &\leq \int H_{\mu^u_{\hat{x}}}(\hat{f}^{-1}\xi^u)\, d\hat{\mu}(\hat{x}) = H_{\hat{\mu}}(\hat{f}^{-1}\xi \mid \xi^u). \end{aligned}$$

This proves the inequality stated in Proposition 6.4. Moreover, the equality holds if and only if $\hat{m}^u_{\hat{x},v}((\hat{F}_A^{-1}\Xi^u)(\hat{x}, v)) = \hat{\mu}^u_{\hat{x}}((\hat{f}^{-1}\xi^u)(\hat{x}))$ for $\hat{m}$-almost every $(\hat{x}, v)$. In other words, the equality holds if and only if $p_*\hat{m}^u_{\hat{x},v} = \hat{\mu}^u_{\hat{x}}$ restricted to the σ-algebra generated by $\hat{F}_A^{-1}\Xi^u$.

Replacing $\hat{F}_A$ by any iterate $\hat{F}_A^n$, and noting that

$$h_{\hat{m}}(\hat{F}_A^n, \mathcal{W}^u) = n h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u) \quad \text{and} \quad h_{\hat{\mu}}(\hat{f}^n, W^u) = n h_{\hat{\mu}}(\hat{f}, W^u),$$

we get that the equality holds if and only if $p_*\hat{m}^u_{\hat{x},v} = \hat{\mu}^u_{\hat{x}}$ restricted to the σ-algebra generated by $\hat{F}_A^{-n}\Xi^u$. Since $\bigcup_n \hat{F}_A^{-n}\Xi^u$ generates the Borel σ-algebra of every $\Xi^u(\hat{x}, v)$, this is the same as saying that $p_*\hat{m}^u_{\hat{x},v} = \hat{\mu}^u_{\hat{x}}$ for $\hat{m}$-almost every $(\hat{x}, v)$, that is, that $\hat{m}$ is c-invariant. ∎

6.3. Proof of Theorem 6.1. The hypothesis (13) ensures that the Lyapunov exponents of $\hat{m}$ along the center fibers $\{\hat{x}\} \times \mathbb{PR}^2$ are non-positive. Then

$$h_{\hat{m}}(\hat{F}_A) = h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u)$$

(see [22, Corollary 5.3]). Similarly, $h_{\hat{\mu}}(\hat{f}) = h_{\hat{\mu}}(\hat{f}, W^u)$. Since $h_{\hat{m}}(\hat{F}_A) \geq h_{\hat{\mu}}(\hat{f})$, because $(\hat{f}, \hat{\mu})$ is a factor of $(\hat{F}_A, \hat{m})$, it follows that

$$h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u) \geq h_{\hat{\mu}}(\hat{f}, W^u).$$

In view of Proposition 6.4, this implies that

$$h_{\hat{m}}(\hat{F}_A, \mathcal{W}^u) = h_{\hat{\mu}}(\hat{f}, W^u).$$

Then, by Propositions 6.2 and 6.4, the measure $\hat{m}$ is u-invariant, as claimed.

7. Invariant sections and *su*-states

We say that an $\hat{F}_A$-invariant probability measure $\hat{m}$ is an ***su*-state** if it is both an s-state and a u-state. Here we prove:

Theorem 7.1: *Assume that A admits no invariant section in $\mathbb{PR}^2$ or $\mathbb{PR}^{2,2}$, and there exists some periodic point p of f such that $A^{\mathrm{per}(p)}(p)$ is hyperbolic. Then A has no su-states.*

Assume, by contradiction, that $\hat{F}_A$ does admit some su-state $\hat{m}$. Suppose for a while that $\hat{m}$ admits a continuous disintegration $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ along the vertical fibers $\{\hat{x}\} \times \mathbb{PR}^2$. The fact that $\hat{m}$ is $\hat{F}_A$-invariant means that

$$A(\hat{x})_* \hat{m}_{\hat{x}} = \hat{m}_{\hat{f}(\hat{x})}$$

for $\hat{m}$-almost every $\hat{x}$. Then, by continuity, this must hold for every $\hat{x}$.

Let $\hat{p}$ be the fixed point of $\hat{f}$ in $\pi^{-1}(p)$ and $\kappa = \mathrm{per}(p)$ be its period. Then $\hat{A}^\kappa(\hat{p}) = A^\kappa(p)$ is hyperbolic. The fact that $\hat{A}^\kappa(\hat{p})_* \hat{m}_{\hat{p}} = \hat{m}_{\hat{p}}$ implies that $\hat{m}_{\hat{p}}$ is a convex combination of not more than two Dirac masses. Then, by su-invariance, the same is true about $\hat{m}_{\hat{x}}$ for every $\hat{x}$. Thus $\xi(\hat{x}) = \operatorname{supp} \hat{m}_{\hat{x}}$ defines an invariant section for $\hat{F}_A$, which is in contradiction with the hypotheses.

In general, disintegrations are only measurable. In what follows we explain how to bypass that and turn the previous outline into an actual proof of Theorem 7.1.

7.1. Dirac disintegrations. By the definition of su-state, there are disintegrations $\{\hat{m}^1_{\hat{x}} : \hat{x} \in \hat{M}\}$ and $\{\hat{m}^2_{\hat{x}} : \hat{x} \in \hat{M}\}$ of $\hat{m}$ and there exists a full $\hat{\mu}$-measure subset $U_{\hat{p}}$ of the neighborhood $\hat{V}_{\hat{p}} \approx W^u_{\mathrm{loc}}(\hat{p}) \times W^s_{\mathrm{loc}}(\hat{p})$ such that

(i) $(h^u_{\hat{x},\hat{y}})_* \hat{m}^1_{\hat{x}} = \hat{m}^1_{\hat{y}}$ for every $\hat{x}, \hat{y} \in U_{\hat{p}}$ with $\hat{y} \in W^u_{\mathrm{loc}}(\hat{x})$ (u-invariance of $\hat{m}^1$);
(ii) $\hat{m}^2_{\hat{y}} = \hat{m}^2_{\hat{z}}$ for every $\hat{y}, \hat{z} \in U_{\hat{p}}$ with $\hat{z} \in W^s_{\mathrm{loc}}(\hat{y})$ (s-invariance of $\hat{m}^2$);
(iii) $\hat{m}^1_{\hat{x}} = \hat{m}^2_{\hat{x}}$ for every $\hat{x} \in U_{\hat{p}}$ (essential uniqueness of disintegrations).

Also, we may choose $U_{\hat{p}}$ so that $\hat{m}^1_{\hat{x}}(\Lambda_{\hat{x}}) = 1$ (recall that $\Lambda_{\hat{x}} = \Lambda \cap (\{\hat{x}\} \times \mathbb{PR}^2)$) for every $\hat{x} \in U_{\hat{p}}$.

Since the Pesin unstable manifolds $\mathcal{W}^u(\hat{z}, u)$ vary measurably with the point, we may find compact sets $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda$ such that $\hat{m}(\Lambda_j) \to 1$ and $\mathcal{W}^u(\hat{z}, u)$ varies continuously on every Λ_j. We may choose these compact sets in such a way that $\hat{F}_A(\Lambda_j) \subset \Lambda_{j+1}$ for every $j \geq 1$. Up to reducing $U_{\hat{p}}$ if necessary, $\hat{m}^1(\Lambda_{j,\hat{x}}) \to 1$ for every $\hat{x} \in U_{\hat{p}}$.

Fix any $\hat{x} \in U_{\hat{p}}$ such that $\hat{\mu}^u_{\hat{x}}(\xi^u(\hat{x}) \setminus U_{\hat{p}}) = 0$. Then define $\hat{m}_{\hat{x}} = \hat{m}^1_{\hat{x}}$ and

(a) $\hat{m}_{\hat{y}} = (h^u_{\hat{x},\hat{y}})_* \hat{m}_{\hat{x}}$ for every $\hat{y} \in \xi^u(\hat{x})$;
(b) $\hat{m}_{\hat{z}} = \hat{m}_{\hat{y}}$ for every $\hat{z} \in W^s_{\mathrm{loc}}(\hat{y}) \cap \hat{V}_{\hat{p}}$ with $\hat{y} \in \xi^u(\hat{x})$.

By (i)–(iii), we have that $\hat{m}_{\hat{y}} = \hat{m}^1_{\hat{y}} = \hat{m}^2_{\hat{y}}$ for every $\hat{y} \in \xi^u(\hat{x}) \cap U_{\hat{p}}$ and $\hat{m}_{\hat{z}} = \hat{m}^2_{\hat{z}}$ for every $\hat{z} \in W^s_{\mathrm{loc}}(\hat{y}) \cap \hat{V}_{\hat{p}}$ with $\hat{y} \in \xi^u(\hat{x}) \cap U_{\hat{p}}$. By the choice of $\hat{x}$ and the fact that $\hat{\mu}$ has local product structure, the latter corresponds to a full $\hat{\mu}$-measure subset of points $\hat{z} \in \hat{V}_{\hat{p}}$. In particular, $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{V}_{\hat{p}}\}$ is a disintegration of $\hat{m}$ on $\hat{V}_{\hat{p}}$.

Let us collect some immediate consequences of the definition of the measures $\hat{m}_{\hat{z}}$. For $\hat{x}, \hat{y}, \hat{z}$ as in (a)–(b) above, denote $h^{su}_{\hat{x},\hat{z}} = h^s_{\hat{y},\hat{z}} \circ h^u_{\hat{x},\hat{y}}$ with $h^s_{\hat{y},\hat{z}} : \{\hat{y}\} \times \mathbb{PR}^2 \to \{\hat{z}\} \times \mathbb{PR}^2$ given by the identity. For $j \geq 1$, denote

$$\alpha_j = \hat{m}_{\hat{x}}(\Lambda_{j,\hat{x}});$$

keep in mind that $\alpha_j \to 1$.

Lemma 7.2: *For each $j \geq 1$,*

(a) $\tilde{\Lambda}_{j,\hat{z}} = h^{su}_{\hat{x},\hat{z}}(\Lambda_{j,\hat{x}})$ *is compact and varies continuously with $\hat{z} \in \hat{V}_{\hat{p}}$ in the Hausdorff topology;*
(b) *the measure $\hat{m}_z \mid \tilde{\Lambda}_{j,\hat{z}}$ varies continuously with $\hat{z} \in \hat{V}_{\hat{p}}$ in the weak* topology;*
(c) $\hat{m}_{\hat{z}}(\tilde{\Lambda}_{j,\hat{z}}) = \alpha_j$ *for every $\hat{z} \in \hat{V}_{\hat{p}}$.*

Since the matrix $\hat{A}^\kappa(\hat{p})$ is hyperbolic, its action on the projective space $\mathbb{PR}^2$ is a North pole–South pole map, that is, a Morse–Smale diffeomorphism with one attractor a and one repeller r. We are going to prove:

PROPOSITION 7.3: *The support of $\hat{m}_{\hat{p}}$ is contained in $\{a, r\}$.*

Proof. Since $\{\hat{m}_{\hat{z}} : \hat{z} \in \hat{V}_{\hat{p}}\}$ is a disintegration and $\hat{m}$ is $\hat{F}_A$-invariant,

$$(\hat{F}_A^\kappa)_* \hat{m}_{\hat{z}} = \hat{m}_{\hat{f}^\kappa(\hat{z})} \quad \text{for } \hat{\mu}\text{-almost every } \hat{z} \in \hat{V}_{\hat{p}} \cap \hat{f}^{-\kappa}(\hat{V}_{\hat{p}}). \tag{18}$$

The identity may not hold for $\hat{z} = \hat{p}$, but we are going to show that $\hat{m}_{\hat{p}}$ is at least "almost $\hat{F}_A$-invariant," in a suitable sense:

LEMMA 7.4: *$\hat{m}_{\hat{p}}(\hat{F}_A^{-l\kappa}(K)) \geq \hat{m}_{\hat{p}}(K)$ for any compact set $K \subset \tilde{\Lambda}_{j,\hat{p}}$ and every $l \geq 1$ and $j \geq 1$.*

Proof. Fix K, l and j. For any $\hat{q}$ close to $\hat{p}$, define

$$h_{\hat{q}} = h^s_{\hat{z}, \hat{F}_A^{l\kappa}(\hat{q})} \circ h^u_{\hat{y},\hat{z}} \circ h^s_{\hat{p},\hat{y}}$$

where $\hat{y}$ and $\hat{z}$ are the points where $W^u_{\text{loc}}(\hat{x})$ intersects $W^s_{\text{loc}}(\hat{p})$ and $W^s_{\text{loc}}(\hat{f}^{l\kappa}(\hat{q}))$, respectively. Keep in mind that the two s-holonomies are given by the identity. Also, $K \subset \tilde{\Lambda}_{j,\hat{p}}$ ensures that $h_{\hat{q}}$ is continuous restricted to K. Define $K_{\hat{q}} = h_{\hat{q}}(K)$. Then $K_{\hat{q}}$ is a compact subset of $\{\hat{f}^{l\kappa}(\hat{q})\} \times \mathbb{PR}^2$ such that

$$\hat{m}_{\hat{f}^{l\kappa}(\hat{q})}(K_{\hat{q}}) = \hat{m}_{\hat{p}}(K).$$

The point $\hat{q}$ may be chosen arbitrarily close to $\hat{p}$ because $\hat{\mu}$ has full support. Making $\hat{q} \to \hat{p}$, the point $\hat{f}^{l\kappa}(\hat{q})$ also goes to $\hat{p}$, and then $\hat{y}$ converges to $\hat{z}$ (which is fixed). Thus $K_{\hat{q}} \to K$ as $\hat{q} \to \hat{p}$.

Using (18), we may choose $\hat{q}$ such that $(\hat{F}_A^{l\kappa})_* \hat{m}_{\hat{q}} = \hat{m}_{\hat{f}^{l\kappa}(\hat{q})}$ and $\hat{q}$ is close enough to $\hat{p}$ that $\hat{f}^{n\kappa}(\hat{q}) \in \hat{V}_{\hat{p}}$ for $0 \leq n \leq l$. It follows that

$$\hat{m}_{\hat{q}}(\hat{F}_A^{-l\kappa}(K_{\hat{q}})) = \hat{m}_{\hat{f}^{l\kappa}(\hat{q})}(K_{\hat{q}}) = \hat{m}_{\hat{p}}(K).$$

Lemma 7.2(c) gives that $\hat{m}_{\hat{q}}(\tilde{\Lambda}_{k,\hat{q}}) = \alpha_k$ for every $k \geq 1$. Thus

$$(\hat{m}_{\hat{q}} \mid \tilde{\Lambda}_{k,\hat{q}})(\hat{F}_A^{-l\kappa}(K_{\hat{q}})) = \hat{m}_{\hat{q}}(\tilde{\Lambda}_{k,\hat{q}} \cap \hat{F}_A^{-l\kappa}(K_{\hat{q}})) \geq \hat{m}_{\hat{p}}(K) + \alpha_k - 1. \tag{19}$$

By parts (a) and (b) of Lemma 7.2 the compact set $\tilde{\Lambda}_{k,\hat{q}}$ and the measure $\hat{m}_{\hat{z}} \mid \tilde{\Lambda}_{k,\hat{q}}$ depend continuously on $\hat{q}$. We know that the same is true for the sets $\hat{F}_A^{-l\kappa}(K_{\hat{q}})$. Thus, making $\hat{q} \to \hat{p}$ in (19), we get that

$$(\hat{m}_{\hat{p}} \mid \tilde{\Lambda}_{k,\hat{p}})(\hat{F}_A^{-l\kappa}(K)) \geq \hat{m}_{\hat{p}}(K) + \alpha_k - 1 \tag{20}$$

(use the general fact that if $\nu_i \to \nu$ in the weak* topology and $C_i \to C$ in the Hausdorff topology, then $\nu(C) \geq \limsup_i \nu_i(C_i)$). Clearly, the left-hand side of (20) is less than or equal to $\hat{m}_{\hat{p}}(\hat{F}_A^{-l\kappa}(K))$. So, making $k \to \infty$ we get the claim. ∎

We are ready to complete the proof of Proposition 7.3. Suppose that $\hat{m}_{\hat{p}}$ is not supported inside $\{a, r\}$. Then, since the $\tilde{\Lambda}_{k,\hat{p}}$ are a non-decreasing sequence whose union has full $\hat{m}_{\hat{p}}$-measure, for every large $k \geq 1$ the measure $\hat{m}_{\hat{p}} \mid \tilde{\Lambda}_{k,\hat{p}}$ is not supported on $\{a, r\}$. Then we can find a compact set $K \subset \tilde{\Lambda}_{k,\hat{p}}$ contained in a fundamental domain of $\hat{A}^\kappa(\hat{p})$ with positive $\hat{m}_{\hat{p}}$-measure. By Lemma 7.4, it follows that $\hat{m}_{\hat{p}}(\hat{F}_A^{-l\kappa}(K)) \geq \hat{m}_{\hat{p}}(K) > 0$ for every $l \geq 0$. Since these sets are pairwise disjoint, it follows that $\hat{m}_{\hat{p}}$ is an infinite measure, which is a contradiction. ∎

7.2. Proof of Theorem 7.1. By Proposition 7.3, $\hat{m}_{\hat{p}}$ is a convex combination of not more than two Dirac masses. Then, in view of the definition of this disintegration, the same is true about $\hat{m}_{\hat{z}}$ for every $\hat{z} \in \hat{V}_{\hat{p}}$. Then $\hat{\xi}(\hat{z}) = \operatorname{supp} \hat{m}_{\hat{z}}$ defines a continuous map on $\hat{V}_{\hat{p}}$ with values on $\mathbb{PR}^2$ or $\mathbb{PR}^{2,2}$ and such that $\hat{A}(\hat{z})\hat{\xi}(\hat{z}) = \hat{\xi}(\hat{f}(\hat{z}))$ for every $\hat{z} \in \hat{V}_{\hat{p}} \cap \hat{f}^{-1}(\hat{V}_{\hat{p}})$.

The same argument shows that for any point $\hat{y} \in \hat{M}$ there exists a continuous disintegration $\{\hat{m}_{\hat{y},\hat{z}} : \hat{z} \in \hat{V}_{\hat{y}}\}$ of the su-state restricted to $\hat{V}_{\hat{y}}$. Since disintegrations are essentially unique and the neighborhoods $\hat{V}_{\hat{y}}$ overlap on positive $\hat{\mu}$-measure subsets, all these conditional measures $\hat{m}_{\hat{y},\hat{z}}$ must be supported on the same number, 1 or 2, of points. Thus, the map ξ in the previous paragraph extends to a continuous invariant section on the whole $\hat{M}$, which contradicts the assumptions of Theorem B.

8. Proof of Theorem B

If $\lambda(A) = 0$ then, trivially, A is a continuity point. Now assume that $\lambda(A) > 0$. Then (see Kalinin [17, Theorem 1.4]) there exists some periodic point p of f such that $A^{\mathrm{per}(p)}(p)$ is hyperbolic. Thus we may use Theorem 7.1 to conclude that there are no su-states. Now the proof of continuity of the Lyapunov exponents is entirely analogous to Section 4.

The same arguments also prove the converse: if the cocycle is hyperbolic at some periodic point then, again by Theorem 7.1, there are no su-states and thus the exponent cannot vanish. The proof of Theorem B is complete.

Appendix A. Smooth natural extensions

We show that the natural extension of any C^k local diffeomorphism $f : M \to M$ on a compact manifold admits a C^k realization.

Since M is compact and f is locally injective, we may find families of open sets $\{U_i, V_i : i = 1, \dots, N\}$ such that: $\{U_1, \dots, U_N\}$ covers M; every V_i contains the closure of U_i; and every $f \mid V_i$ is injective. Take smooth functions $h_i : M \to [0, 1]$ such that $h_i \mid U_i \equiv 1$ and $h_i \mid V_i^c \equiv 0$. Define

$$h(x) = (h_1(x), \dots, h_N(x))$$

for $x \in M$. Then $h : M \to [0, 1]^N$ is such that $h(x) \neq h(y)$ for any pair (x, y) with $x \neq y$ and $f(x) = f(y)$. Since f is locally injective, the set of such pairs is a compact subset of M^2. Hence, there is $\delta > 0$ such that $\|h(x) - h(y)\| \geq \delta$ for any (x, y) with $x \neq y$ and $f(x) = f(y)$.

Let $\phi : M \to \mathbb{R}^m$ be a Whitney embedding of M into some Euclidean space, and $\psi : M \times D \to \mathbb{R}^m$ be a tubular neighborhood: D denotes the open unit ball in $\mathbb{R}^{m-\dim M}$ and ψ is a smooth embedding with $\psi(x, 0) = \phi(x)$. Identify $M \times D$ with its image $U = \psi(M \times D)$ through ψ. Fix $\lambda < \delta/4N$ and define

$$g : M \times D \to M \times D, \quad g(x, v) = (f(x), h(x)/2N + \lambda v).$$

It is clear that g is well defined and a C^k local diffeomorphism, and the image $g(M \times D)$ is relatively compact in $M \times D$.

Suppose that $g(x, v) = g(y, w)$. Then $f(x) = f(y)$ and

$$h(x) - h(y) = 2N\lambda(w - v).$$

In particular, $\|h(x) - h(y)\| \leq 4N\lambda < \delta$. By the definition of δ, this implies that $x = y$. Then the previous identities imply that $v = w$. This proves that g is injective and, consequently, an embedding.

For each $\hat{x} = (x_{-n})_n \in \hat{M}$ and $n \geq 1$ the set $g^n(\{x_{-n}\} \times D)$ is a disk $D_n(\hat{x})$ of radius λ^n inside $\{x_0\} \times D$. These disks are nested and each $D_{n+1}(\hat{x})$ is relatively compact in $D_n(\hat{x})$. Thus, the intersection consists of exactly one point, which we denote as $\iota(\hat{x})$. By construction, the map $\iota : \hat{M} \to M \times D$ defined in this way satisfies $g \circ \iota = \iota \circ \hat{f}$. Moreover, the image $\iota(M)$ coincides with $\bigcap_n g^n(M \times D)$ and so it satisfies

$$g(\iota(\hat{M})) = \iota(\hat{M}).$$

References

[1] A. Avila, J. Santamaria and M. Viana, *Holonomy invariance: rough regularity and applications to Lyapunov exponents*, Astérisque **358** (2013), 13–74, 2013.

[2] A. Avila and M. Viana, *Extremal Lyapunov exponents: an invariance principle and applications*, Inventiones Mathematicae **181** (2010), 115–189.

[3] L. Backes, A. Brown and C. Butler, *Continuity of Lyapunov exponents for cocycles with invariant holonomies*, Journal of Modern Dynamics **12** (2018), 223–260.

[4] J. Bochi, *Genericity of zero Lyapunov exponents*. PhD thesis, IMPA, 2000, preprint, www.preprint.impa.br.

[5] J. Bochi, *Genericity of zero Lyapunov exponents*, Ergodic Theory and Dynamical Systems **22** (2002), 1667–1696.

[6] J. Bochi, *C^1-generic symplectic diffeomorphisms: partial hyperbolicity and zero centre Lyapunov exponents*, Journal of the Institute of Mathematics of Jussieu **9** (2010), 49–93.

[7] J. Bochi and M. Viana, *The Lyapunov exponents of generic volume-preserving and symplectic maps*, Annals of Mathematics **161** (2005), 1423–1485.

[8] C. Bonatti, L. J. Díaz and M. Viana, *Dynamics Beyond Uniform Hyperbolicity*, Encyclopaedia of Mathematical Sciences, Vol. 102, Springer-Verlag, Berlin, 2005.

[9] C. Bonatti, X. Gómez-Mont and M. Viana, *Généricité d'exposants de Lyapunov non-nuls pour des produits déterministes de matrices*, Annales de l'Institut Henri Poincaré. Analyse Non Linéaire **20** (2003), 579–624.

[10] C. Bonatti and M. Viana, *Lyapunov exponents with multiplicity 1 for deterministic products of matrices*, Ergodic Theory and Dynamical Systems **24** (2004), 1295–1330.

[11] R. Bowen, *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms*, Lecture Notes in Mathematics, Vol. 470, Springer Verlag, Berlin–New York, 1975.

[12] C. Butler, *Discontinuity of Lyapunov exponents near fiber bunched cocycles*, Ergodic Theory and Dynamical Systems **38** (2018), 523–539.

[13] P. Duarte and S. Klein, *Lyapunov Exponents of Linear Cocycles*, Atlantis Studies in Dynamical Systems, Vol. 3, Atlantis Press, Paris, 2016.

[14] H. Furstenberg, *Non-commuting random products*, Transactions of the American Mathematical Society **108** (1963), 377–428.

[15] H. Furstenberg and H. Kesten, *Products of random matrices*, Annals of Mathematical Statistics **31** (1960), 457–469.

[16] M. Herman, *Une méthode nouvelle pour minorer les exposants de Lyapunov et quelques exemples montrant le caractère local d'un théorème d'Arnold et de Moser sur le tore de dimension* 2, Commentarii Mathematici Helvetici **58** (1983), 453–502.

[17] B. Kalinin, *Livšic theorem for matrix cocycles*, Annals of Mathematics **173** (2011), 1025–1042.

[18] J. Kingman, *The ergodic theory of subadditive stochastic processes*, Journal of the Royal Statistical Society **30** (1968), 499–510.

[19] O. Knill, *The upper Lyapunov exponent of* $\mathrm{SL}(2,\mathbf{R})$ *cocycles: discontinuity and the problem of positivity*, in *Lyapunov Exponents (Oberwolfach, 1990)*, Lecture Notes in Mathematics, Vol. 1486, Springer-Verlag, Berlin, 1991, pp. 86–97.

[20] F. Ledrappier, *Propriétés ergodiques des mesures de Sinaï*, Institut des Hautes Études Scientifiques. Publications Mathématiques **59** (1984), 163–188.

[21] F. Ledrappier, *Positivity of the exponent for stationary sequences of matrices*, in *Lyapunov Exponents (Bremen, 1984)*, Lecture Notes in Mathematics, Vol. 1186, Springer-Verlag, Berlin, 1986, pp. 56–73.

[22] F. Ledrappier and L.-S. Young, *The metric entropy of diffeomorphisms. I. Characterization of measures satisfying Pesin's entropy formula*, Annals of Mathematics **122** (1985), 509–539.

[23] R. Mañé, *Oseledec's theorem from the generic viewpoint*, in *Proceedings of the International Congress of Mathematicians, Vols. 1, 2 (Warsaw, 1983)*, PWN, Warsaw, 1984, pp. 1269–1276.

[24] V. I. Oseledets, *A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems*, Transactions of the Moscow Mathematical Society **19** (1968), 197–231.

[25] V. A. Rokhlin, *Lectures on the entropy theory of measure-preserving transformations*, Russian Mathematical Surveys **22**(1967), 1–52; translated from Uspekhi Matematičeskih Nauk **22** (1967), 3–56.

[26] A. Tahzibi and J. Yang, *Invariance principle and rigidity of high entropy measures*, Transactions of the American Mathematical Society, to appear.

[27] M. Viana, *Almost all cocycles over any hyperbolic system have nonvanishing Lyapunov exponents*, Annals of Mathematics **167** (2008), 643–680.

[28] M. Viana, *Lectures on Lyapunov Exponents*, Cambridge Studies in Advanced Mathematics, Vol. 145, Cambridge University Press, Cambridge, 2014.

[29] M. Viana and K. Oliveira, *Foundations of Ergodic Theory*, Cambridge Studies in Advanced Mathematics, Vol. 151, Cambridge University Press, Cambridge, 2015.

[30] J. Yang, *Entropy along expanding foliations*, arXiv:1601.05504.

[31] L.-S. Young, *Some open sets of nonuniformly hyperbolic cocycles*, Ergodic Theory and Dynamical Systems **13** (1993), 409–415.