LYAPUNOV EXPONENTS OF NON-LINEAR COCYCLES

ARTUR AVILA AND MARCELO VIANA

1. Introduction

Let $(\hat{M}, \hat{\mathcal{B}}, \hat{\mu})$ be a probability space and $\hat{f} : \hat{M} \to \hat{M}$ be a measurable transformation preserving the probability $\hat{\mu}$. Let $\hat{P} : \hat{\mathcal{E}} \to \hat{M}$ be a fiber bundle with fibers $\hat{\mathcal{E}}_x$ diffeomorphic to some Riemannian manifold N. A *non-linear cocycle* over $\hat{f}$ is a measurable transformation $\hat{F} : \hat{\mathcal{E}} \to \hat{\mathcal{E}}$ such that $\hat{P} \circ \hat{F} = \hat{f} \circ \hat{P}$ and every $\hat{F}_{\hat{x}} : \hat{\mathcal{E}}_x \to \hat{\mathcal{E}}_{\hat{f}(x)}$ is a diffeomorphism.

We always assume that the norms of the derivative $D\hat{F}_{\hat{x}}(\xi)$ and its inverse are uniformly bounded. Then the functions

$$(\hat{x}, \xi) \mapsto \log \|D\hat{F}_{\hat{x}}(\xi)\| \quad \text{and} \quad (\hat{x}, \xi) \mapsto \log \|D\hat{F}_{\hat{x}}(\xi)^{-1}\| \tag{1}$$

are integrable, relative to any probability measure on $\hat{\mathcal{E}}$. The *extremal Lyapunov exponents* of $\hat{F}$ at a point $(\hat{x}, \xi) \in \hat{\mathcal{E}}$ are

$$\lambda_+(\hat{F}, \hat{x}, \xi) = \lim_{n\to\infty} \frac{1}{n} \log \|D\hat{F}^n_{\hat{x}}(\xi)\| .$$
$$\lambda_-(\hat{F}, \hat{x}, \xi) = \lim_{n\to\infty} \frac{1}{n} \log \|D\hat{F}^n_{\hat{x}}(\xi)^{-1}\|^{-1} .$$

The limits exist $\hat{m}$-almost everywhere, with respect to any $\hat{F}$-invariant probability $\hat{m}$ on $\hat{\mathcal{E}}$, by sub-additivity (Kingman [4]). Notice that

$$\lambda_-(\hat{F}, \hat{x}, \xi) \le \lambda_+(\hat{F}, \hat{x}, \xi),$$

because $\|D\hat{F}^n_{\hat{x}}(\xi)\| \|D\hat{F}^n_{\hat{x}}(\xi)^{-1}\| \ge 1$. Denote

$$\lambda_\pm = \lambda_\pm(\hat{F}, \hat{m}) = \int \lambda_\pm(\hat{F}, \hat{x}, \xi)\, d\hat{m}(\hat{x}, \xi).$$

If $(\hat{F}, \hat{m})$ is ergodic then $\lambda_\pm(\hat{F}, \hat{x}, \xi) = \lambda_\pm$ for $\hat{m}$-almost every $(\hat{x}, \xi)$. Throughout, we shall only be interested in measures $\hat{m}$ that project down to μ under $\hat{P}$.

Date: January 31, 2007.

2. A non-linear invariance result

In this section we state and prove a non-linear extension of a theorem of Ledrappier [5] about Lyapunov exponents of linear cocycles, which is crucial for what follows. The proof is close to the arguments in [5].

Take $(\hat{M}, \hat{\mathcal{B}}, \mu)$ to be a Lebesgue space, that is, a complete separable probability space. See Rokhlin [6, §2–§3]. Then any probability $\hat{m}$ on $\hat{\mathcal{E}}$ such that $\hat{P}_* \hat{m} = \hat{\mu}$ admits a family $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ of probabilities such that $\hat{x} \mapsto \hat{m}_{\hat{x}}$ is $\hat{\mathcal{B}}$-measurable, every $\hat{m}_{\hat{x}}$ is supported inside the fiber $\hat{\mathcal{E}}_{\hat{x}}$ and

$$\hat{m}(E) = \int \hat{m}_{\hat{x}}(E)\, d\hat{\mu}(\hat{x})$$

for any measurable set $E \subset \hat{\mathcal{E}}$. Moreover, such a family is essentially unique. We call it the *disintegration* of $\hat{m}$ and refer to the $\hat{m}_{\hat{x}}$ as its *conditional probabilities* along the fibers.

Throughout this section we assume that $\hat{f}$ is invertible. A σ-algebra $\mathcal{B}_0 \subset \hat{\mathcal{B}}$ is *generating* if its iterates $\hat{f}^n(\mathcal{B}_0)$, $n \in \mathbb{Z}$ generate the whole $\hat{\mathcal{B}}$ mod 0. The main result in this section is

Theorem 2.1. *Suppose $\lambda_-(\hat{F}, \hat{x}, \xi) \geq 0$ for $\hat{m}$-almost every $(\hat{x}, \xi)$. Let $\mathcal{B}_0 \subset \hat{\mathcal{B}}$ be a generating σ-algebra relative to which both $\hat{f}$ and $\hat{x} \mapsto \hat{F}_{\hat{x}}$ are measurable* mod 0. *Then $\hat{x} \mapsto \hat{m}_{\hat{x}}$ is also $\mathcal{B}_0$-measurable* mod 0.

The same holds true if we assume, instead, that $\lambda_+(\hat{F}, \hat{x}, \xi) \leq 0$ for $\hat{m}$-almost every (x, ξ). To see this, notice that $\hat{F}$ is invertible and, clearly, has the same invariant probabilities as its inverse. Since

$$\lambda_+(\hat{F}, \hat{x}, \xi) + \lambda_-(\hat{F}^{-1}, \hat{x}, \xi) = 0,$$

the new assumption means that $\lambda_-(\hat{F}^{-1}, \hat{x}, \xi) \geq 0$ for $\hat{m}$-almost every (x, ξ). Thus, we may apply Theorem 2.1 to the inverse cocycle, to obtain the same conclusion as before under this new assumption.

Example 2.2. Given any (non-invertible) measure-preserving transformation $f : M \to M$ in a probability space $(M, \mathcal{B}, \mu)$, define $\hat{M}$ to be the space of all sequences $(x_n)_{n \leq 0}$ in M such that $f(x_n) = x_{n+1}$ for all $n < 0$, and consider the *natural extension* of f,

$$\hat{f} : \hat{M} \to \hat{M}, \quad \hat{f}(\ldots, x_n, \ldots, x_0) = (\ldots, x_n, \ldots, x_0, f(x_0)).$$

Then $\hat{f}$ is invertible and $\pi \circ \hat{f} = f \circ \pi$, where $\pi : \hat{M} \to M$ is the projection to the zeroth term. Denote $\mathcal{B}_0 = \pi^{-1}(\mathcal{B})$ and let $\hat{\mathcal{B}}$ be the σ-algebra on $\hat{M}$ generated by the iterates $\hat{f}^n(\mathcal{B}_0)$, $n \geq 0$. Then $\hat{f}$ is measurable with respect to $\mathcal{B}_0$ and to $\hat{\mathcal{B}}$. Let μ_0 be the probability measure defined on $\mathcal{B}_0$ by $\pi_*\mu_0 = \mu$. There is a unique $\hat{f}$-invariant probability $\hat{\mu}$ on $(\hat{M}, \hat{\mathcal{B}})$ such that $\pi_*\hat{\mu} = \mu$: it is characterized by

$$E\big(\hat{\mu} \mid \hat{f}^n(\mathcal{B}_0)\big) = \hat{f}^n_*\mu_0 \quad \text{for every } n \geq 0. \tag{2}$$

To any non-linear cocycle $F : \mathcal{E} \to \mathcal{E}$ over f, defined on a fiber bundle $P : \mathcal{E} \to M$, we may associate the non-linear cocycle $\hat{F} : \hat{\mathcal{E}} \to \hat{\mathcal{E}}$ over $\hat{f}$ defined by $\hat{\mathcal{E}}_{\hat{x}} = \mathcal{E}_{\pi(\hat{x})}$ and $\hat{F}_{\hat{x}} = F_{\pi(\hat{x})}$. Their extremal Lyapunov exponents are related by

$$\lambda_\pm(\hat{F}, \hat{x}, \xi) = \lambda_\pm(F, \pi(\hat{x}), \xi).$$

Clearly, $\hat{x} \mapsto \hat{F}_{\hat{x}}$ is $\mathcal{B}_0$-measurable. We denote by $\pi \times \mathrm{id}$ the natural projection from $\hat{\mathcal{E}}$ to $\mathcal{E}$ (this terminology is motivated by the case when $\hat{\mathcal{E}} = \hat{M} \times N$ and $\mathcal{E} = M \times N$). Given any F-invariant probability m, there is exactly one $\hat{F}$-invariant probability $\hat{m}$ with $(\pi \times \mathrm{id})_*\hat{m} = m$: it is characterized by

$$E\big(\hat{x} \mapsto \hat{m}_{\hat{x}} \mid \hat{f}^n(\mathcal{B}_0)\big) = [\hat{x} \mapsto (\hat{F}^n_{\hat{x}})_* m_{\pi(\hat{x})}] \quad \text{for every } n \geq 0 \tag{3}$$

(see Lemma 2.5 below), where $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ and $\{m_x : x \in M\}$ are the disintegrations of $\hat{m}$ and m, respectively. If $P_*m = \mu$ then $\hat{P}_*\hat{m} = \hat{\mu}$.

Let us begin the proof of Theorem 2.1. For the time being we consider the special case of natural extensions described in Example 2.2: we shall soon argue that the statement of the theorem can always be reduced to that case. Let

$$(F_x^{-1})_* m_{f(x)} = J(x, \xi) m_x + \eta_x \tag{4}$$

be the Lebesgue decomposition of $(F_x^{-1})_* m_{f(x)}$ relative to m_x (the function $J(x, \cdot)$ is integrable for m_x and the measure η_x is singular with respect to m_x). We call $J : \mathcal{E} \to [0, \infty)$ the *fibered Jacobian*, and define the *fibered entropy* to be

$$h = h(\hat{F}, \hat{m}) = \int -\log J \, dm. \tag{5}$$

The definition (4) implies $\int_{\{J>0\}} J \, dm = \int J \, dm \leq 1$. Then, by Jensen's inequality,

$$\int_{\{J>0\}} -\log J \, dm \geq 0. \tag{6}$$

The definition (5) means that h is the sume of this integral with the term $(+\infty) \cdot m(\{J = 0\})$ with the usual convention that the latter vanishes if $m(\{J = 0\}) = 0$. Thus, h is always well-defined and non-negative. Besides, we shall see later that $\{J = 0\}$ always has zero measure in our context.

Let κ be the dimension of N and $\lambda^0(\hat{F}, \hat{x}, \xi) = \min\{0, \lambda_-(\hat{F}, \hat{x}, \xi)\}$.

Proposition 2.3. *We have* $0 \le h \le -\kappa \int \lambda^0(\hat{F}, \hat{x}, \xi)\, d\hat{m}(\hat{x}, \xi)$.

Proposition 2.4. *If* $h = 0$ *then* $\hat{x} \mapsto \hat{m}_{\hat{x}}$ *is* $\mathcal{B}_0$*-measurable* $\bmod 0$.

Proof of Theorem 2.1. Propositions 2.3 and 2.4 immediately lead to Theorem 2.1 in the natural extension case: $\lambda_-(\hat{F}, \hat{x}, \xi) \ge 0$ means that $\lambda^0(\hat{F}, \hat{x}, \xi)$ vanishes identically; then Proposition 2.3 yields $h = 0$ and, by Proposition 2.4, it follows that $\hat{x} \mapsto \hat{m}_{\hat{x}}$ is $\mathcal{B}_0$-measurable $\bmod 0$.

So, now we only have to show that the general case in the theorem can always be viewed as a natural extension. By Rokhlin [6], one may find a Lebesgue space $(M, \mathcal{B}, \mu)$ and a projection $\pi : \hat{M} \to M$ such that $\mathcal{B} = \pi_* \mathcal{B}_0$ and $\mu = \pi_* \hat{\mu}$. In other words, $B \in \mathcal{B}$ if and only if $\pi^{-1}(B) \in \mathcal{B}_0$ and then $\mu(B) = \hat{\mu}(\pi^{-1}(B))$. Since $\hat{f}$ is $\mathcal{B}_0$-measurable $\bmod 0$, there exists a $\mathcal{B}$-measurable $\bmod 0$ transformation $f : M \to M$ such that $\pi \circ \hat{f} = f \circ \pi$. This transformation is usually non-invertible, but it preserves the measure μ:

$$f_* \mu = f_* \pi_* \hat{\mu} = \pi_* \hat{f}_* \mu = \pi_* \hat{\mu} = \mu.$$

Since $\hat{F}$ is $\mathcal{B}_0$-measurable, it may be written as $\hat{F} = F \circ (\pi \times \mathrm{id})$, where $F : \mathcal{E} \to \mathcal{E}$ is a $\mathcal{B}$-measurable non-linear cocycle over f. Notice that $DF_{\pi(\hat{x})} = D\hat{F}_{\hat{x}}$ for every $\hat{x} \in \hat{M}$. In particular,

$$\lambda(F, \pi(\hat{x}), \xi) = \lambda(\hat{F}, \hat{x}, \xi)$$

for almost every $(\hat{x}, \xi)$, where $\lambda(F, x, \xi)$ is the Lyapunov exponent of the cocycle F at a point (x, ξ). Then $m = (\pi \times \mathrm{id})_* \hat{m}$ is an F-invariant probability and $P_* m = \mu$. Applying the previous arguments to F we get that the disintegration $\hat{x} \mapsto \hat{m}_{\tilde{x}}$ if $\mathcal{B}$-measurable $\bmod 0$. $\square$

2.1. **Entropy zero means deterministic.** As a first step we prove Proposition 2.4. Let $\{\hat{m}_{\hat{x}} : \hat{x} \in \hat{M}\}$ and $\{m_x : x \in M\}$ be the disintegrations of $\hat{m}$ and m, respectively.

Lemma 2.5. *For* $\hat{\mu}$*-almost every* $\hat{x} \in \hat{M}$,

$$\hat{m}_{\hat{x}} = \lim_{n \to \infty} (F^n_{x(n)})_* m_{x(n)} \qquad \textit{with } x(n) = \pi(\hat{f}^{-n}(\hat{x})).$$

Proof. Let m_0 be the probability defined on $\mathcal{B}_0$ by $\pi_* m_0 = m$. The disintegration of m_0 is just $\hat{x} \mapsto m_{\pi(\hat{x})}$. The relation $\pi_* \hat{m} = m$ implies that $\hat{m} \mid \mathcal{B}_0 = m_0$ or, in other words, $E(\hat{x} \mapsto \hat{m}_{\hat{x}} \mid \mathcal{B}_0) = [\hat{x} \mapsto m_{\pi(\hat{x})}]$. Next, the relation $\hat{F}_* \hat{m} = \hat{m}$ implies that

$$E(\hat{x} \mapsto \hat{m}_{\hat{x}} \mid \hat{f}^n(\mathcal{B}_0)) = E(\hat{x} \mapsto (\hat{F}^n_{\hat{x}(n)})_* \hat{m}_{\hat{x}(n)} \mid \mathcal{B}_0),$$

with $\hat{x}(n) = \hat{f}^{-n}(\hat{x})$, and so

$$E(\hat{x} \mapsto \hat{m}_{\hat{x}} \mid \hat{f}^{n}(\mathcal{B}_0)) = [\hat{x} \mapsto (F^n_{x(n)})_* m_{x(n)}].$$

Any of these expressions defines a martingale of probability measures, relative to the sequence of σ-algebras $\hat{f}^{n}(\mathcal{B}_0)$. Since $\mathcal{B}_0$ is generating and the sequence $\hat{f}^{n}(\mathcal{B}_0)$ is increasing, the limit of the left hand side is

$$[\hat{x} \mapsto \hat{m}_{\hat{x}}] = E(\hat{x} \mapsto \hat{m}_{\hat{x}} \mid \hat{\mathcal{B}}).$$

It follows that $(F^n_{x(n)})_* m_{x(n)}$ converges and the limit coincides with $\hat{m}_{\hat{x}}$ at $\hat{\mu}$-almost every point. $\square$

Lemma 2.6. *If $h = 0$ then $(F_x)_* m_x = m_{f(x)}$ for μ-almost every $x \in M$.*

Proof. The definition (4) implies that $\int J(x,\xi)\, dm_x(\xi) \le 1$ for μ-every x. So, by Jensen's inequality, $\int -\log J(x,\xi)\, dm_x(\xi) \ge 0$ for μ-every x. Moreover, the equalities hold if and only if $J(x,\xi) = 1$ for m_x-almost every ξ. This implies that $h \ge 0$, and $h = 0$ if and only if $J(x,\xi) = 1$ for m_x-almost every ξ and μ-almost x. In particular, $h = 0$ implies $m_{f(x)} = (F_x)_* m_x$ for μ-almost x, as claimed. $\square$

Lemma 2.6 implies $(F^n_{x(n)})_* m_{x(n)} = m_{x(0)}$ for every $n \ge 0$ and $\hat{\mu}$-almost every $\hat{x}$. Then Lemma 2.5 yields $\hat{m}_{\hat{x}} = m_{x(0)}$ for $\hat{\mu}$-almost every $\hat{x}$. Since $x(0) = \pi(\hat{x})$, this implies that $\hat{x} \mapsto \hat{m}_{\hat{x}}$ is $\mathcal{B}_0$-measurable, and so the proof of Proposition 2.4 is complete.

2.2. **Entropy is smaller than exponents.**

Now we prove Proposition 2.3. Let $\{m_\alpha\}$ be the ergodic decomposition of m, and $d\alpha$ denote the corresponding quotient measure:

$$\varphi\, dm = \int \left(\int \varphi\, dm_\alpha \right) d\alpha$$

for any integrable function. If $\lambda_-(F,x,\xi) \ge 0$ at m-almost every point then the same is true at m_α-almost every point, for $d\alpha$-almost every ergodic component. Assuming the proposition holds for ergodic measures, it follows that

$$0 \le \int -\log J\, dm_\alpha \le \kappa \int \lambda^0(F,x,\xi)\, dm_\alpha$$

for $d\alpha$-almost every α. Integrating with respect to $d\alpha$, we obtain that

$$0 \le h \le \kappa \int \lambda^0(F,x,\xi)\, dm$$

as claimed. Hence, it is no restriction to assume that m is ergodic, and we do so in all that follows. Then $\lambda^0(F,x,\xi)$ is constant m-almost everywhere; let $\lambda^0 = \lambda^0(F,m)$ denote this constant.

The proof of the proposition will follow from upper and lower estimates on the measures of balls that we are going to state next. Define

$$\Delta^l(y,\eta) = \min\{1, \|DF_y^l(\eta)^{-1}\|^{-1}\} \quad \text{and} \quad \Delta^{l,n}(x,\xi) = \prod_{j=0}^{n-1} \Delta^l(F^{lj}(x,\xi))$$

for any $l \geq 1$ and $n \geq 1$ and let $L = \sup |\log \|(DF)^{-1}\|^{-1}|$. It is clear from the definitions that

$$-L \leq \frac{1}{nl} \log \Delta^{l,n}(x,\xi) \leq 0 \tag{7}$$

for every $l \geq 1$, $n \geq 1$, and (x,ξ).

Lemma 2.7. *Given $\varepsilon > 0$ there exists $k \geq 1$ and some F^k-ergodic component m_0 of the measure m such that*

$$\lim_{n\to\infty} \frac{1}{nk} \log \Delta^{k,n}(x,\xi) \geq \lambda^0 - 3\varepsilon \quad \textit{for } m_0\textit{-almost every } (x,\xi).$$

Proof. The limit is the Birkhoff time average of the (bounded) function $k^{-1} \log \Delta^k$, and so it exists at m-almost. Choose $k \geq 1$ large enough so that

$$\frac{1}{k} \log \|DF_x^k(\xi)^{-1}\|^{-1} \in [\lambda_- - \varepsilon, \lambda_- + \varepsilon] \tag{8}$$

for a subset E of points (x,ξ) with $m(E) > 1 - \varepsilon/L$. It follows that $m_0(E) > 1 - \varepsilon/L$ for some F^k-ergodic component m_0 of the measure m. Fix m_0 from now on. There are three cases to consider, depending on the sign of the exponent. If λ_- is negative then $\lambda_- = \lambda^0$. It is no restriction to assume from the start that $\varepsilon < |\lambda_-|$. Then (8) implies

$$\frac{1}{k} \log \Delta^k(x,\xi) = \frac{1}{k} \log \|DF_x^k(\xi)^{-1}\|^{-1} \geq \lambda^0 - \varepsilon,$$

for every $(x,\xi) \in E$. In general this expression is bounded by L, as observed previously. By ergodicity, m_0-typical orbits of F^k spend a fraction $> 1 - \varepsilon/L$ of the time inside E. Thus, for m_0-almost every (x,ξ) and every large n,

$$\frac{1}{nk} \Delta^{k,n}(x,\xi) \geq (1 - \varepsilon/L)(\lambda^0 - \varepsilon) - (\varepsilon/L)L \geq \lambda^0 - 3\varepsilon$$

(for the last step, notice that $|\lambda^0| \leq |\lambda_-| \leq L$). This settles the case $\lambda_- < 0$. If λ_- is zero then $\lambda_- = \lambda^0 = 0$. Then (8) implies

$$\frac{1}{k} \log \Delta^k(x,\xi) = \min\{0, \frac{1}{k} \log \|DF_x^k(\xi)^{-1}\|^{-1}\} \geq -\varepsilon$$

for every $(x,\xi) \in E$. Thus, arguing as before,

$$\frac{1}{nk} \Delta^{k,n}(x,\xi) \geq -(1 - \varepsilon/L)\varepsilon - (\varepsilon/L)L \geq -3\varepsilon$$

for m_0-almost every (x, ξ) and every n sufficiently large. This settles the lemma $\lambda_- = 0$. If λ_- is positive then $\lambda^0 = 0$. It is no restriction to assume from the beginning that $\varepsilon < \lambda_-$. Then (8) implies

$$\frac{1}{k} \log \Delta^k(x, \xi) = \min\{0, \frac{1}{k} \log \|DF_x^k(\xi)^{-1}\|^{-1}\} = 0$$

for every $(x, \xi) \in E$. Thus, arguing as before,

$$\frac{1}{nk} \Delta^{k,n}(x, \xi) \geq -(\varepsilon/L)L \geq -3\varepsilon$$

for m_0-almost every (x, ξ) and every n sufficiently large. This completes the proof of the lemma. □

Remark 2.8. There exists some divisor $s \geq 1$ of k such that

$$m = \frac{1}{s}\left(m_0 + F_*(m_0) + \cdots + F_*^{s-1}(m_0)\right) \quad \text{and} \quad F_*^s(m_0) = m_0.$$

Indeed, by the F-ergodicity of m, if A is any F^k-invariant set with positive measure then $m(A) \geq 1/k$. Fix A with minimal positive measure. Then every $A \cap F^{-i}(A)$ has either zero or full measure in A. Let $s \geq 1$ be the smallest integer such that $A \cap F^{-s}(A)$ has full measure in A. Note that s divides k. Using F-ergodicity once more, $A \cup F^{-1}(A) \cup \cdots \cup F^{-s+1}(A)$ has full measure in $\mathcal{E}$ and so $m(A) = 1/s$. The normalized restrictions

$$m_i = s(m \mid F^{i-k}(A)), \quad i = 0, \ldots, s-1$$

are the F^k-ergodic components of m. Clearly, $m_i = F_*^i(m_0)$ for every $0 \leq i \leq s-1$ and $F_*^s(m_0) = m_0$. Up to replacing m_0 by an iterate, we may take it to coincide with the ergodic component in the previous lemma.

Let $\varepsilon > 0$ be fixed and k and m_0 be as in Lemma 2.7. The proofs of the next two propositions will be given in Section 2.3 and Section 2.5, respectively:

Proposition 2.9. *For m_0-almost every (x, ξ),*

$$\limsup_{n\to\infty} \frac{1}{nk} \log m_{f^{nk}(x)}(B(F_x^{nk}(\xi), e^{-nk\varepsilon}\Delta^{k,n}(x, \xi))) \geq \kappa(\lambda^0 - 5\varepsilon).$$

Proposition 2.10. *For m_0-almost every (x, ξ),*

$$\limsup_{n\to\infty} \frac{1}{nk} \log m_{f^{nk}(x)}(B(F_x^{nk}(\xi), e^{-nk\varepsilon}\Delta^{k,n}(x, \xi))) \leq -h + 3\varepsilon.$$

It follows that $\kappa(\lambda^0 - 5\varepsilon) \leq -h + 3\varepsilon$ and, since ε is arbitrary, this implies $h \leq -\kappa\lambda^0$. We have already seen in Lemma 2.6 that $h \geq 0$. So, we have reduced the proof of Proposition 2.3, and of Theorem 2.1, to proving Propositions 2.9 and 2.10.

2.3. **Lower estimate.** Here we prove Proposition 2.9. We begin with the following consequence of the Besicovitch covering lemma (see [3, Theorem 1.1]):

Lemma 2.11. *Given any $\theta > 0$ the set*

$$Z_\theta = \{(y, \eta) : m_y(B(\eta, \rho)) \leq \rho^{(1+\theta)\kappa} \textit{ for arbitrarily small } \rho > 0\}$$

has zero m-measure.

Proof. Write the manifold N as a union of increasing compact subsets N_k. For each y consider a covering of Z_θ intersected with $\{y\} \times N_k$ (viewed as a subset of $\mathcal{E}_y$) by balls $B(\eta_j, \rho_j)$ as in the definition of Z_θ, with $\rho_j \leq \rho$ for some small ρ. Then

$$\begin{aligned} m(Z_\theta \cap (\mathcal{E}_k)) &\leq \int \sum_j m_y(B(\eta_j, \rho_j))\, d\mu(y) \\ &\leq C(\kappa)\rho^{\theta\kappa} \int \sum_j |B(\eta_j, \rho_j)|\, d\mu(y) \end{aligned}$$

where $|\cdot|$ denotes Riemannian volume relative to some fixed Riemannian metric on N. By the Besicovitch covering lemma we may take the covering such that every point belongs to at most $C(\kappa)$ of these balls. It follows that

$$m(Z_\theta \cap (\bigcup_{y \in M} \{y\} \times N_k)) \leq C(\kappa)\rho^{\theta\kappa} |N_k|\,.$$

Making $\rho \to 0$ and then $k \to \infty$ we get $m(Z_\theta) = 0$, as claimed. $\square$

Denote $r_n = e^{-nk\varepsilon}\Delta^{k,n}(x, \xi)$. We are going to prove Proposition 2.9 by contradiction. Suppose there exists a constant $\theta > 0$ and a positive m-measure set E such that for every $(x, \xi) \in E$ there exists $\bar{n}(x, \xi) \geq 1$ such that

$$\log m_{f^{nk}(x)}(B(F_x^{nk}(\xi), r_n)) \leq \kappa(\lambda^0 - 5\varepsilon)(1 + \theta)nk \tag{9}$$

for all $n \geq \bar{n}(x, \xi)$. By Lemma 2.7, we may choose $\bar{n}(x, \xi)$ so that

$$\log r_n \geq nk(\lambda^0 - 5\varepsilon) \tag{10}$$

for all $n \geq \bar{n}(x, \xi)$. Define

$$Z_{\theta,r} = \{(y, \eta) : m_y(B(\eta, \rho)) \leq \rho^{(1+\theta)\kappa} \text{ for some } 0 < \rho \leq r\}.$$

These sets are monotone decreasing when $r \to 0$, and their intersection is the set Z_θ. The upper bound in (7) implies that $r_n \to 0$ and the inequalities (9) and (10) imply $F^{nk}(x, \xi) \in Z_\theta(r_n)$ for every $n \geq \bar{n}(x, \xi)$. This shows that almost every (x, ξ) eventually enters $Z_{\theta,r}$ for every fixed $r > 0$. In particular, we may find $n_r \geq 1$ and $E_r \subset E$ such that $m(E_r) \geq m(E)/2$ and $F^{nk}(x,\xi) \in Z_{\theta,r}$ for every $n \geq n_r$. Then

$$m(Z_{\theta,r}) = m(F^{-nk}(Z_{\theta,r}) \geq m(E_r) \geq m(E)/2.$$

Since this applies for arbitrary r, we conclude that $m(Z_\theta) > 0$, which contradicts Lemma 2.11. This proves Proposition 2.9.

Remark 2.12. For arbitrary $l \geq 1$, instead of (10) we may use

$$\log e^{-nl\varepsilon} \Delta^{l,n}(x,\xi) \geq -nl(L+\varepsilon) \geq -nl(L+\varepsilon).$$

Then the same arguments as before yield, for m-almost every point,

$$\limsup_{n\to\infty} \frac{1}{nl} \log m_{f^{nl}(x)}(B(F_x^{nl}(\xi), e^{-nl\varepsilon}\Delta^{l,n}(x,\xi))) \geq -\kappa(L+\varepsilon). \tag{11}$$

2.4. **Auxiliary statements.** Here we recall some general facts from measure theory. Let X be a compact metric space locally Lipeomorphic to $\mathbb{R}^\kappa$, $\kappa \geq 1$ and let μ and ν be Borel probability measures on X.

Lemma 2.13. *There exists a function $g \in L^1(\nu)$ and a Borel measure η totally singular with respect to ν such that $\mu = f\nu + \eta$. Moreover,*

$$g(x) = \lim_{\delta\to 0} \frac{\mu(B(x,\delta))}{\nu(B(x,\delta))} \quad \textit{for } \nu\textit{-almost every } x \in X.$$

See Rudin [7, Theorem 7.14] for a related statement involving the Lebesgue measure. The proof in the general case is analogous, just using the Besicovich covering lemma ([3, Theorem 1.1]) in the place of the Vitali covering lemma ([7, Theorem 7.3]).

Lemma 2.14. *There exists $c(X) > 0$ such that $\int g^*\, d\nu \leq c(X)$, where*

$$g_*(x) = \sup\Big\{\frac{\mu(B(x,\delta))}{\nu(B(x,\delta))} : \delta > 0\Big\}.$$

We consider $X = \mathcal{E}_x \simeq N$ and $\mu = (F_x^{-l})_* m_{f^l(x)}$ and $\nu = m_x$ for every $x \in M$ and $l \geq 1$. Lemma 2.13 gives that

$$(F_x^{-l})_* m_{f^l(x)} = J^l(x,\xi) m_x + \eta_x^l$$

for some measure η_x^l totally singular with respect to m_x, and

$$J^l(x,\xi) = \lim_{\delta\to 0} \frac{m_{f^l(x)}(F_x^l(B(\xi,\delta)))}{m_x(B(\xi,\delta))} \quad \text{for } m_x\text{-almost every } \xi. \tag{12}$$

By Lemma 2.14 there exists $c(N) > 0$ such that

$$\int \log J_*^l(x,\xi)\, dm_x(\xi) \leq c(N) < \infty, \tag{13}$$

for every $x \in M$ and $l \geq 1$, where

$$J_*^l(x,\xi) = \sup\left\{\frac{m_{f^l(x)}(F_x^l(B(\xi,\delta)))}{m_x(B(\xi,\delta))} : \delta > 0\right\}.$$

Corollary 2.15. *For m_x-almost every ξ there exists $\delta_1(x,\xi,\varepsilon,l) > 0$ such that*

$$\frac{m_{f^l(x)}\big(B(F_x^l(\xi), e^{-l\varepsilon}\Delta^l(x,\xi)\delta)\big)}{m_x(B(\xi,\delta))} \leq \begin{cases} e^{l\varepsilon} J^l(x,\xi) & \text{if } J^l(x,\xi) > 0 \\ e^{-\tau_l} & \text{if } J^l(x,\xi) = 0. \end{cases}$$

for every $0 < \delta \leq \delta_1(x,\xi,\varepsilon,l)$, where $\tau_l = 4\kappa L/m(\{J^l = 0\})$.

Proof. The relation (12) implies that

$$\frac{m_{f^l(x)}\big(F_x^l(B(\xi,\delta))\big)}{m_x(B(\xi,\delta))} \leq \begin{cases} e^{l\varepsilon} J^l(x,\xi) & \text{if } J^l(x,\xi) > 0 \\ e^{-\tau_l} & \text{if } J^l(x,\xi) = 0 \end{cases}$$

if δ is small enough. Secondly, since every F_x is a diffeomorphism,

$$\begin{aligned} F_x^l(B(\xi,\delta)) &\supset B(F_x^l(\xi), e^{-l\varepsilon}\|DF_x^l(\xi)^{-1}\|^{-1}\delta) \\ &\supset B(F_x^l(\xi), e^{-l\varepsilon}\Delta^l(x,\xi)\delta) \end{aligned}$$

if δ is small enough. The claim follows from these two observations. $\square$

Corollary 2.16. *Suppose $J > 0$ at m-almost every point. Then*

$$\int \log J^k \, dm_0 = -kh.$$

Proof. We begin by proving that, for $l \geq 1$ and m-almost every (x,ξ),

$$J^l(x,\xi) = \prod_{j=0}^{l-1} J(F^j(x,\xi)). \tag{14}$$

Indeed, the relation (12) may be rewritten as

$$J^l(x,\xi) = \lim_{\delta\to 0}\prod_{j=0}^{l-1}\frac{m_{f^{j+1}(x)}(F_x^{j+1}(B(\xi,\delta)))}{m_{f^j(x)}(F_x^j(B(\xi,\delta)))}$$

where the jth factor converges to $J(F^j(x,\xi))$ for $(F_x^{-j})_* m_{f^j(x)}$-almost every ξ. Now, the assumption means that $J(x,\xi) > 0$ for μ-almost every x and m_x-almost every ξ. Then m_x is absolutely continuous with respect to $(F_x^{-1})_* m_{f(x)}$.

Let us consider x in the full μ-measure set points such that m_x is absolutely continuous with respect to $(F_x^{-j})_* m_{f^j(x)}$ for every $j \geq 1$. Then each factor in the previous formula converges to $J(F^j(x,\xi))$ for

m_x-almost every ξ. This proves that (14) is true for any $l \geq 1$ and m_x-almost every ξ. Next, from (14) we get

$$\int \log J^k \, dm_0 = \sum_{m=0}^{k-1} \int \log J \circ F^m \, dm_0 = \sum_{i=1}^{k/s} \sum_{j=0}^{s-1} \log J \, dF_*^i(m_0).$$

Using Remark 2.8, we find that the right hand side is equal to

$$\sum_{i=1}^{k/s} s \int \log J \, dm = k \int \log J \, dm = -kh.$$

This proves the statement. □

The same arguments apply for arbitrary $l \geq 1$: the corollary remains true if one replaces m_0 by any F^l-ergodic component of m (we shall not use this fact). On the other hand, (14) is usually false if one omits the assumption $J > 0$.

2.5. **Upper estimate.** Now we prove Proposition 2.10. Let $l \geq 1$ be arbitrary, for the time being. For each $p \geq 1$ define $W_{l,p}$ to be the set of points (x, ξ) such that

$$e^{-jl\varepsilon} \Delta^{l,j}(x, \xi) \leq \delta_1(x, \xi, \varepsilon, l) \quad \text{for every } j \geq p.$$

The upper bound in (7) implies that $m(W_{l,p})$ goes to 1 as $p \to \infty$. Let p be fixed, for the time being. Write

$$\frac{m_{f^{nl}(x)}\big(B(F_x^{nl}(\xi), e^{-nl\varepsilon}\Delta^{l,n}(x,\xi))\big)}{m_{f^{pl}(x)}\big(B(F_x^{pl}(\xi), e^{-pl\varepsilon}\Delta^{l,p}(x,\xi))\big)} = \prod_{j=p}^{n-1} q_{l,j}(x, \xi), \tag{15}$$

for every $n \geq p$ and $(x, \xi) \in \mathcal{E}$, where

$$q_{l,j}(x, \xi) = \frac{m_{f^{l(j+1)}(x)}\big(B(F_x^{l(j+1)}(\xi), e^{-l(j+1)\varepsilon}\Delta^{l,j+1}(x,\xi))\big)}{m_{f^{lj}(x)}\big(B(F_x^{lj}(\xi), e^{-lj\varepsilon}\Delta^{l,j}(x,\xi))\big)}.$$

Denote $(x_j, \xi_j) = F^{lj}(x, \xi)$ for each $j \geq 1$. If $(x, \xi) \in W_{l,p}$ then, by Corollary 2.15,

$$\log q_{l,j}(x, \xi) \leq \begin{cases} l\varepsilon + \log J^l(x_j, \xi_j) & \text{if } J^l(x_j, \xi_j) > 0 \\ -\tau_l & \text{if } J^l(x_j, \xi_j) = 0. \end{cases}$$

If $(x, \xi) \notin W_{l,p}$ we use the trivial bound $\log q_{l,j}(x, \xi) \leq \log J_*^l(x_j, \xi_j)$ instead. In any event,

$$\log q_{l,j}(x, \xi) \leq l\phi_{l,p}(x_j, \xi_j) \tag{16}$$

where

$$\phi_{l,p} = \varepsilon + \frac{1}{l} \log J^l \mathcal{X}_{\{J^l > 0\} \cap W_{l,p}} - \tau_l \mathcal{X}_{\{J^l = 0\} \cap W_{l,p}} + \frac{1}{l} \log J_*^l \mathcal{X}_{W_{l,p}^c}. \tag{17}$$

In the proof of the next couple of lemmas we consider the case $l = 1$. Afterwards, we use these relations in the case $l = k$ to complete the proof of Proposition 2.10.

Lemma 2.17. *We have $J(x, \xi) > 0$ for m-almost every (x, ξ).*

Proof. Suppose $m(\{J = 0\}) > 0$. Then, right from the start, we may choose p large enough so that

$$m(\{J = 0\} \cap W_{1,p}) \geq m(\{J = 0\})/2 > 0.$$

In view of (6) and (13), we may also suppose that

$$\int_{\{J>0\}\cap W_{1,p}} \log J \, dm \leq \varepsilon \quad \text{and} \quad \int_{W_{1,p}^c} \log J_* \, dm \leq \varepsilon,$$

up to increasing p. Recall $\tau_1 = 4\kappa L/m(\{J = 0\})$. Then, from (17),

$$\int \phi_{1,p} \, dm \leq 3\varepsilon - 2\kappa L < -\kappa(L + \varepsilon) \tag{18}$$

(for the last inequality, let ε be small enough with respect to L). Since m is F-ergodic, the relation (16) implies that

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{j=p}^{n-1} \log q_{1,j}(x, \xi) \leq \lim_{n\to\infty} \frac{1}{n} \sum_{j=p}^{n-1} \phi_{1,p}(x_j, \xi_j) = \int \phi_{1,p} \, dm.$$

In view of (15), this just means that

$$\limsup_{n\to\infty} \frac{1}{n} \log m_{f^n(x)}(B(F_x^n(\xi), e^{-n\varepsilon} \Delta^{1,n}(x, \xi))) \leq \int \phi_{1,p} \, dm.$$

According to (18), this inequality contradicts (11). This contradiction proves that $\{J = 0\}$ has zero measure, as claimed. □

Lemma 2.18. *The fibered entropy h is finite.*

Proof. Suppose $h > 2\kappa L$. Then, by (6) and (13), we may choose p right from the start so that

$$\int_{\{J>0\}\cap W_{1,p}} \log J \, dm \leq -2L \quad \text{and} \quad \int_{W_{l,p}^c} \log J_* \, dm \leq \varepsilon.$$

Then, using also the previous lemma, (17) yields

$$\int \phi_{1,p} \, dm \leq 2\varepsilon - 2\kappa L < -\kappa(L + \varepsilon). \tag{19}$$

Now, arguing precisely as in Lemma 2.17, we conclude that

$$\limsup_{n\to\infty} \frac{1}{n} \log m_{f^n(x)}(B(F_x^n(\xi), e^{-n\varepsilon} \Delta^{1,n}(x, \xi))) < -\kappa(L + \varepsilon),$$

which contradicts (11). This proves that $h \leq 2\kappa L < \infty$, as claimed. □

We are ready to conclude the proof of Proposition 2.10. Lemma 2.17 ensures we are in a position to apply Corollary 2.16. The relation (14) implies that $J^l > 0$ at m-almost every point, for every $l \geq 1$. We are going to use this fact, and the relations (16) and (17), in the case $l = k$. Fix $p \geq 1$ large enough so that

$$\int_{\{J^k>0\}\cap W_{k,p}} \log J^k \, dm \leq k(-h+\varepsilon) \quad \text{and} \quad \int_{W_{l,p}^c} \log J_*^k \, dm \leq k\varepsilon.$$

Then, from the definition (17),

$$\int \phi_{k,p} \, dm_0 \leq 3\varepsilon - h. \tag{20}$$

Since m_0 is ergodic for F^k,

$$\limsup_{n\to\infty} \frac{1}{nk} \sum_{j=p}^{n-1} \log q_{k,j}(x,\xi) \leq \lim_{n\to\infty} \frac{1}{n} \sum_{j=p}^{n-1} \phi_{k,p}(x_j,\xi_j) = \int \phi_{k,p} \, dm_0$$

m_0-almost everywhere. In view of (15), this means that

$$\limsup_{n\to\infty} \frac{1}{nk} \log m_{f^{nk}(x)}(B(F_x^{nk}(\xi), e^{-nk\varepsilon}\Delta^{k,n}(x,\xi))) \leq \int \phi_{k,p} \, dm_0$$

for m_0-almost every (x,ξ). Combined with (20), this implies the conclusion of Proposition 2.10.

References

[1] C. Bonatti and M. Viana. SRB measures for partially hyperbolic systems whose central direction is mostly contracting. *Israel J. Math.*, 115:157–193, 2000.

[2] K. Burns and A. Wilkinson. On the ergodicity of partially hyperbolic systems. Preprint Northwestern University 2005.

[3] M. de Guzman. *Differentiation of integrals in* $\mathbb{R}^n$, volume 481 of *Lect. Notes in Math.* Springer Verlag, 1975.

[4] J. Kingman. The ergodic theorem of subadditive stochastic processes. *J. Royal Statist. Soc.*, 30:499–510, 1968.

[5] F. Ledrappier. Positivity of the exponent for stationary sequences of matrices. In *Lyapunov exponents (Bremen, 1984)*, volume 1186 of *Lect. Notes Math.*, pages 56–73. Springer, 1986.

[6] V. A. Rokhlin. On the fundamental ideas of measure theory. *A. M. S. Transl.*, 10:1–52, 1962. Transl. from Mat. Sbornik 25 (1949), 107–150.

[7] W. Rudin. *Real and complex analysis.* McGraw-Hill, 3 edition, 1987.

[8] M. Viana. Almost all cocycles over any hyperbolic system have non-vanishing Lyapunov exponents. *Annals of Math.*

[9] J.-C. Yoccoz. Travaux de Herman sur les tores invariants. *Astérisque*, 206:Exp. No. 754, 4, 311–344, 1992. Séminaire Bourbaki, Vol. 1991/92.