Existence and uniqueness of maximizing measures for robust classes of local diffeomorphisms

Krerley Oliveira and Marcelo Viana *

Abstract

We prove existence of maximal entropy measures for an open set of non-expanding local diffeomorphisms on a compact Riemannian manifold. In this context the topological entropy coincides with the logarithm of the degree, and these maximizing measures are eigenmeasures of the transfer operator. When the map is topologically mixing, the maximizing measure is unique and positive on every open set.

1 Introduction

In its most basic form, the variational principle states that the topological entropy of a continuous transformation on a compact space coincides with the supremum of the entropies of the probability measures invariant under the transformation. We call *maximizing measure* any invariant probability for which the supremum is attained. Existence and uniqueness of such measures has been investigated by many authors, in a wide variety of situations. However, the global picture is still very much incomplete.

In this paper we contribute a simple sufficient condition for existence and uniqueness, applicable to a large class of transformations. Some examples we have in mind are the non-uniformly expanding local diffeomorphisms of Alves, Bonatti, Viana [ABV00], which exhibit only positive Lyapunov exponents at "most" points. But our hypothesis, formulated in (1) below, is a condition of the type that Buzzi [Buz99, Buz00] introduced and called entropy-expansivity: we only ask that the derivative do not expand k-dimensional volume too much, for all k less than the dimension of the ambient manifold. We show that this implies existence and, if the transformation is topologically mixing, uniqueness of the maximizing measure.

2 Statement of main result

Let $f : M^d \to M^d$ be a C^1 local diffeomorphism on a compact d-dimensional Riemannian manifold. Let $p \geq 1$ be the degree of f, that is, the number $\# f^{-1}(x)$

*This work was partially supported by Pronex, CNPq, Fapeal, and Faperj, Brazil.

of preimages of any point $x \in M$. Define

$$C_k(f) = \max_{x \in M} \|\Lambda^k Df(x)\|,$$

where Λ^k represents the kth exterior product. We assume that f satisfies

$$\max_{1 \le k \le d-1} \log C_k(f) < \log p. \tag{1}$$

We say that $f : M \to M$ is *topologically mixing* if given any open set U there exists $N \in \mathbb{N}$ such that $f^N(U) = M$. We are going to prove the following

Theorem A. *Assume f satisfies* (1). *Then $h_{top}(f) = \log p$, and any maximal eigenmeasure μ of the transfer operator $\mathcal{L}$ is a maximizing measure. In particular, there exists some maximizing measure for f. If f is topologically mixing then the maximizing measure is unique and positive on open sets.*

The Ruelle-Perron-Frobenius *transfer operator* of $f : M \to M$ is the bounded linear operator $\mathcal{L} : C(M) \to C(M)$ defined on the space $C(M)$ of continuous functions $g : M \to \mathbb{R}$ by

$$\mathcal{L}g(x) = \sum_{y: f(y)=x} g(y).$$

Observe that this is a positive operator. Its dual $\mathcal{L}^* : \mathcal{M}(M) \to \mathcal{M}(M)$ acts on the space of Borel measures of M, by

$$\int g \, d\mathcal{L}^*\nu = \int \mathcal{L}g \, d\nu,$$

preserving the cone of positive measures, and the subset of probability measures.

It is easy to see that the spectra of $\mathcal{L}$ and $\mathcal{L}^*$ are contained in the closed disk of radius p. We call *maximal eigenmeasure* any probability measure μ that satisfies

$$\mathcal{L}^*\mu = p\,\mu.$$

It is well-known that maximal eigenmeasures do exist. A quick proof goes as follows. Define $G : \mathcal{M}_1 \to \mathcal{M}_1$ on the space of probabilities $\mathcal{M}_1$ on M by

$$G(\nu) = \frac{1}{p}\mathcal{L}^*\nu.$$

Then G is continuous relative to the weak* topology on $\mathcal{M}_1$. Since $\mathcal{M}_1$ is a convex compact space, we may use the Tychonoff-Schauder theorem to conclude that there exists some probability μ such $G(\mu) = \mu$. In other words, μ is a maximal eigenmeasure. Observe also that μ is invariant for f. In fact, for every continuous function g we have that $\mathcal{L}(g \circ f)(x) = pg(x)$ and

$$\int (g \circ f) d\mu = \frac{1}{p}\int (g \circ f) \, d\mathcal{L}^*\mu = \frac{1}{p}\int \mathcal{L}(g \circ f) \, d\mu = \int g \, d\mu.$$

The paper is organized as follows. In Section 3 we prove that, under our assumptions, any measure with large entropy has only positive Lyapunov exponents. In Section 5 we prove that measures with positive Lyapunov exponents admits generating partitions with small diameter. This conclusion uses the notion of hyperbolic times, that we recall in Section 4. On its turn, it is used in Section 6 to show that the entropy of such measures is given by a simple formula involving the Jacobian. Using this formula, we prove in Section 7 that the topological entropy is $\log p$ and is attained by any maximal eigenmeasure. Finally, in Section 8 we prove that the maximal measure is unique if the transformation is topologically mixing.

Acknowledgements. We are thankful to Vítor Araújo for a conversation that helped clarify the arguments in the last section.

3 Measures with large entropy

By Oseledets [Os68], if μ is an f-invariant probability measure then for μ-almost every point $x \in M$ there is $k = k(x) \geq 1$, a filtration

$$T_x M = F_x^1 \supset \cdots \supset F_x^k \supset F^{k+1}(x) = \{0\},$$

and numbers $\hat{\lambda}_1(x) > \hat{\lambda}_2(x) > \cdots > \hat{\lambda}_k(x)$ such $Df(x)F_x^i = F_{f(x)}^i$ and

$$\hat{\lambda}_i = \lim_{n\to\infty} \frac{1}{n} \log \|Df^n(x)v\|$$

for every $v \in F_x^i \setminus F_x^{i+1}$ and $i = 1, \ldots, k$. The numbers $\hat{\lambda}_i(x)$ are called *Lyapunov exponents* of f at the point x. The *multiplicity* of $\lambda_i(x)$ is $\dim F_x^i - \dim F_x^{i+1}$. We also write the Lyapunov exponents as

$$\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_d(x),$$

where each number is repeated according to the corresponding multiplicity. Then the *integrated Lyapunov exponents* are the averages

$$\lambda_i(\mu) = \int \lambda_i(x)\, d\mu(x), \quad \text{for } i = 1, \ldots, d.$$

Given a vector space V and a number $k \geq 1$, the kth exterior power of V is the vector space of all alternate k-linear forms defined on the dual of V. We always take V to be finite-dimensional, and then the exterior product $\Lambda^k V$ admits an alternative description, as the linear space spanned by the wedge products $v_1 \wedge \cdots \wedge v_k$ of vectors $v_1, v_2, \ldots, v_k$ in V. Assuming V comes with an inner product, we can endow $\Lambda^k V$ with a inner product such that $\|v_1 \wedge \cdots \wedge v_k\|$ is just the volume of the k-dimensional parallelepiped determined by the vectors $v_1, v_2, \ldots, v_k$ in V.

A linear isomorphism $A : V \to W$ induces another, $\Lambda^k A : \Lambda^k V \to \Lambda^k W$, through

$$\Lambda^k A(v_1 \wedge \cdots \wedge v_k) = Av_1 \wedge \cdots \wedge Av_k.$$

When $V = W$, the eigenvalues of $\Lambda^k A$ are just the products of k distinct eigenvalues of A (where an eigenvalue with multiplicity m is counted as m "distinct" eigenvalues). Correspondingly, there is a simple relation between the Lyapunov spectra of $\Lambda^k Df$ and Df: the Lyapunov exponents of $\Lambda^k Df$ are the sums of k distinct Lyapunov exponents of Df, with the same convention as before concerning multiplicities. Thus,

$$\lambda_{i_1}(x) + \lambda_{i_2}(x) + \cdots + \lambda_{i_k}(x) \leq \log C_k(f)$$

for any $1 \leq i_1 < i_2 < \cdots < i_k \leq d$, and our hypothesis (1) implies that these sums are strictly smaller than $\log p$, for all $k < d$.

Lemma 3.1. *If μ is an invariant probability with some integrated Lyapunov exponent less than*

$$c(f) = \log p - \max_{1 \leq k < d} \log C_k(f), \tag{2}$$

then $h_\mu(f) < \log p$.

Proof. Let μ be an invariant probability, and suppose $\int \lambda_d(x) d\mu < c(f)$. As we have just seen, (1) implies that $\sum_{1 \leq i \leq k} \lambda_i(x) \leq \log C_k(f)$ for all $1 \leq k < d$. Then, using the Ruelle inequality [Rue78],

$$h_\mu(f) \leq \int \sum_{i:\lambda_i(x)>0} \lambda_i(x)\, d\mu < c(f) + \max_{1 \leq k < d} \log C_k(f) \leq \log p.$$

This proves the lemma. $\square$

4 Hyperbolic times

For the next step we need the notion of hyperbolic times, introduced by Alves *et al* [Alv00, ABV00]. Given $c > 0$, we say that $n \in \mathbb{N}$ is a *c-hyperbolic time* for $x \in M$ if

$$\prod_{k=0}^{j-1} \|Df(f^{n-k}(x))^{-1}\| \leq e^{-2cj} \quad \text{for every } 1 \leq j \leq n.$$

In what follows we fix $c = c(f)/10$ and speak, simply, of *hyperbolic times*. We say that f has *positive density of hyperbolic times* for x if the set H_x of integers which are hyperbolic times of f for x satisfies

$$\liminf_n \frac{1}{n} \#(H_x \cap [1, n]) > 0. \tag{3}$$

We quote a few basic properties from [ABV00] (alternatively, see [Ol03]):

Lemma 4.1. *If a point x satisfies*

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \log \|Df(f^i(y))^{-1}\| < -4c < 0,$$

then f has positive density of hyperbolic times for x.

In fact, the density, that is, the $\liminf$ in (3), is bounded below by some positive constant that depends only on f (and our choice of c).

Lemma 4.2. *There exists $\delta_0 > 0$, depending only on f and c, such that given any hyperbolic time $n \geq 1$ for a point $x \in M$, and given any $1 \leq j \leq n$, the inverse branch $f_{x,n}^{-j}$ of f^j that sends $f^n(x)$ to $f^{n-j}(x)$ is defined on the whole ball of radius δ_0 around $f^n(x)$, and satisfies*

$$d(f_{x,n}^{-j}(z), f_{x,n}^{-j}(w)) \leq e^{-jc} d(z,w)$$

for every z, w in that ball.

In view of Lemma 3.1, the next lemma applies to any invariant measure μ with $h_\mu(f) \geq \log p$.

Lemma 4.3. *Given an invariant ergodic measure μ whose Lyapunov exponents are all bigger than $8c$, there exists $N \in \mathbb{N}$ such that f^N has positive density of hyperbolic times for μ-almost every point.*

Proof. Since all Lyapunov exponents of μ are greater than $8c$, for almost every $x \in M$ there exists $n_0(x) \geq 1$ such that

$$\|Df^n(x)w\| \geq e^{6cn}\|w\|, \text{ for all } w \in T_xM \text{ and } n \geq n_0(x).$$

In other words,

$$\|Df^n(x)^{-1}\| \leq e^{-6cn}, \text{ for every } n \geq n_0(x).$$

Define $\alpha_n = \mu(\{x : n_0(x) > n\})$. Since f is a local diffeomorphism, we may also fix a constant $K > 0$ such $\|Df(x)^{-1}\| \leq K$ for all $x \in M$. Then

$$\int_M \log \|Df^n(x)^{-1}\| d\mu \leq -6cn + Kn\alpha_n = -(6c + K\alpha_n)n.$$

Since α_n goes to zero when n goes to infinity, by choosing N big enough we ensure that

$$\int_M \frac{1}{N} \log \|Df^N(x)^{-1}\| d\mu < -4c < 0.$$

Then, since μ is ergodic,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \frac{1}{N} \log \|Df^N(f^{Ni}(y))^{-1}\| = \int_M \frac{1}{N} \log \|Df^N(x)^{-1}\| d\mu < -4c.$$

This means that we may apply Lemma 4.1 to conclude. $\square$

According to the remark following Lemma 4.1, we even have that the density of hyperbolic times is bounded below by some positive constant that depends only on f^N (and our choice of c).

Lemma 4.4. *Let $B \subset M$, $\theta > 0$, and $g : M \to M$ be a local diffeomorphism such that g has density $> 2\theta$ of hyperbolic times for every $x \in B$. Then, given any probability measure ν on B and any $m \geq 1$, there exists $n > m$ such that*

$$\nu\big(\{x \in B : n \text{ is a hyperbolic time of } g \text{ for } x\}\big) > \theta.$$

Proof. Define H to be the set of pairs $(x, n) \in B \times \mathbb{N}$ such that n is a hyperbolic time for x. For each $k \geq 1$, let χ_k be the normalized counting measure on the time interval $[m+1, m+k]$. The hypothesis implies that, given any $x \in B$, we have

$$\chi_k\big(H \cap (\{x\} \times \mathbb{N})\big) > 2\theta$$

for every sufficiently large k. Fix $k \geq 1$ large enough so that this holds for a subset C of points $x \in B$ with $\nu(C) > 1/2$. Then, by Fubini's theorem, $(\nu \times \chi_k)(H) > 2\theta$, and this implies that

$$\nu\big(H \cap (B \times \{n\})\big) > \theta$$

for some $n \in [m+1, m+k]$. This gives the conclusion of the lemma. $\square$

5 Generating partitions

In all that follows the constant $\delta_0 > 0$ is fixed as given by Lemma 4.2. Given a partition α of M, we define

$$\alpha_n = \bigvee_{j=0}^{n-1} f^{-j}(\alpha) \quad \text{for each } n \geq 1.$$

Lemma 5.1. *If μ is an invariant measure such that all its Lyapunov exponents are bigger than $8c$, and α is a partition with diameter less than δ_0, for μ-almost every $x \in M$, the diameter of $\alpha_n(x)$ goes to zero when n goes to ∞. In particular, α is an f-generating partition with respect to μ.*

Proof. By Lemma 4.3 there exists $N \geq 1$ such that f^N has positive density of hyperbolic times for μ-almost every point. Define

$$\gamma_k = \bigvee_{j=0}^{k-1} f^{-jN}(\alpha) \quad \text{for each } k \geq 1.$$

By Lemma 4.2, if k is a hyperbolic time of f^N for x then $\operatorname{diam} \gamma_k(x) \leq e^{-cn}$. In particular, since the sets $\gamma_k(x)$ are non-increasing with k, the diameter of $\gamma_k(x)$ goes to zero when $k \to \infty$. Since $\alpha_{kN}(x) \subset \gamma_k(x)$ and the sequence $\operatorname{diam} \alpha_n(x)$ is non-increasing, this immediatelly gives that the diameter of $\alpha_n(x)$ goes to zero when n goes to infinity, for μ-almost every $x \in M$.

The rest of the argument is very standard. It goes as follows. To prove that α is a generating partition for f with respect to μ, it suffices to show that, given any measurable set E and any $\varepsilon > 0$, there exists $n \geq 1$ and elements A_n^i, $i = 1, \ldots, m(n)$ of α_n such that

$$\mu\big(\bigcup_{i=1}^{m} A_n^i \Delta E\big) < \varepsilon.$$

Consider compact sets $K_1 \subset A$ and $K_2 \subset A^c$ such that $\mu(K_1 \Delta A)$ and $\mu(K_2 \Delta A^c)$ are both less than $\varepsilon/4$. Fix $n \geq 1$ large enough so that $\operatorname{diam} \alpha_n(x)$ is smaller than the distance from K_1 to K_2 outside a set of points x with measure less than $\varepsilon/4$. Let A_n^i, $i = 1, \ldots, m(n)$ be the sets $\alpha_n(x)$ that intersect K_1. Then, they are all disjoint from K_2, and so $\mu(\bigcup_i A_n^i \Delta E)$ is bounded above by

$$\mu(E \setminus \bigcup_i A_n^i) + \mu(\bigcup_i A_n^i \setminus E) \leq \mu(E \setminus K_1) + \mu(E^c \setminus K_2) + \varepsilon/4 \leq \varepsilon.$$

This completes the proof. $\square$

6 Rokhlin's formula

The *Jacobian* of a measure μ with respect to f is the (essentially unique) function $J_\mu f$ satisfying

$$\mu(f(A)) = \int_A J_\mu f \, d\mu.$$

for any measurable set A such that $f|_A$ is injective. In other words, the Jacobian is defined by $J_\mu f = d(f_*\mu)/d\mu$. Jacobians do exist in this context, because f is finite-to-one (countable-to-one would suffice). Using the definition, one can verify that $J_\mu f^n(x) = \prod_{i=0}^{n-1} J_\mu f(f^i(x))$ is a Jacobian for each f^n.

Let $f : M \to M$ be a measurable transformation, μ be an invariant probability. Suppose there exists a finite or countable partition α of M such that

(a) f is locally injective, meaning that it is injective on every atom of α;

(b) α is f-generating with respect to μ, in the sense that $\operatorname{diam} \alpha_n(x) \to 0$ for μ-almost every x.

Proposition 6.1. *If μ is an invariant measure satisfying (a) and (b) as above, then*

$$h_\mu(f) = \int \log J_\mu f \, d\mu,$$

where $J_\mu f$ denotes any Jacobian of f relative to μ.

Let $\alpha_\infty = \bigvee_{j=0}^{\infty} f^{-j}(\alpha)$. Denote $\beta_\infty = \bigvee_{j=1}^{\infty} f^{-j}(\alpha)$ and $\beta_n = \bigvee_{j=1}^{n} f^{-j}(\alpha)$ for each $n \geq 1$. Notice that $\beta_\infty(x) = f^{-1}(\alpha_\infty(f(x)))$. The hypothesis that α is generating implies that $\alpha_\infty(x) = \{x\}$, and so

$$\beta_\infty(x) = \{f^{-1}(f(x))\} \quad \text{for } \mu\text{-almost all } x \in M. \tag{4}$$

The conditional expectation of a function $\varphi : M \to \mathbb{R}$ relative to a partition γ is the essentially unique γ-measurable function $E_\mu(\varphi \mid \gamma)$ such that

$$\int_B E_\mu(\varphi \mid \gamma)\, d\mu = \int_B \varphi\, d\mu \tag{5}$$

for every γ-measurable set B.

Lemma 6.2. *$E_\mu(\varphi \mid \beta_\infty)(x) = \sum_{y \in \beta_\infty(x)} \frac{1}{J_\mu f(y)} \varphi(y)$ for μ-almost every x.*

Proof. It is clear that the function on the right hand side is β_∞-measurable. Let B be any β_∞-measurable set, that is, any measurable set that consists of entire atoms of β_∞. By (4), there exists a measurable set C such that $B = f^{-1}(C)$. Then, since μ is invariant,

$$\begin{aligned}\int_B \sum_{y \in \beta_\infty(x)} \frac{1}{J_\mu f(y)} \varphi(y)\, d\mu(x) &= \int_C \sum_{y \in f^{-1}(z)} \frac{1}{J_\mu f(y)} \varphi(y)\, d\mu(z)\\ &= \sum_{A \in \alpha} \int_{C_A} \frac{1}{J_\mu f(y_A)} \varphi(y_A)\, d\mu(z),\end{aligned}$$

where $C_A = f(B \cap A)$ and $y_A = (f \mid A)^{-1}(z)$. Since every $f \mid A$ is injective, we may use the definition of the Jacobian to rewrite the latter expression as

$$\sum_{A \in \alpha} \int_{B \cap A} \varphi(y)\, d\mu(y) = \int_B \varphi\, d\mu.$$

This proves (5) and the lemma. $\square$

For $* \in \mathbb{N} \cup \{\infty\}$, define the conditional entropy (Definition 4.8 in [Wa82])

$$H_\mu(\alpha \mid \beta_*) = \int \sum_{A \in \alpha} -E_\mu(\chi_A \mid \beta_*) \log E_\mu(\chi_A \mid \beta_*)\, d\mu. \tag{6}$$

Lemma 6.3.

1. *$H_\mu(\alpha \mid \beta_n) = \sum_{A \in \alpha} \sum_{B \in \beta_n} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)}$ for $n \in \mathbb{N}$.*
2. *$H_\mu(\alpha \mid \beta_\infty) = \int \log J_\mu f\, d\mu$.*

Proof. For $n \in \mathbb{N}$, the partition β_n is countable, and so

$$E_\mu(\chi_A \mid \beta_n)(x) = \frac{1}{\mu(\beta_n(x))} \int_{\beta_n(x)} \chi_A\, d\mu = \frac{\mu(A \cap \beta_n(x))}{\mu(\beta_n(x))}$$

for every $A \in \alpha$. It follows that

$$H_\mu(\alpha \mid \beta_n) = \sum_{A \in \alpha} \sum_{B \in \beta_n} \int_B -\frac{\mu(A \cap \beta_n(x))}{\mu(\beta_n(x))} \log \frac{\mu(A \cap \beta_n(x))}{\mu(\beta_n(x))}\, d\mu(x).$$

This gives the first statement. Next, Lemma 6.2 says that

$$E_\mu(\chi_A \mid \beta_\infty) = \psi_A \circ f, \quad \text{where } \psi_A(z) = \sum_{y \in f^{-1}(z)} \frac{1}{J_\mu f(y)} \chi_A(y).$$

Notice that if $z \in f(A)$ then $\psi_A(z) = 1/J_\mu f(y_A)$, where $y_A = (f \mid A)^{-1}(z)$, and if $z \notin f(A)$ then $\psi_A(z) = 0$. Therefore,

$$\begin{aligned} H_\mu(\alpha \mid \beta_\infty) &= \int \sum_{A \in \alpha} -\psi_A(z) \log \psi_A(z)\, d\mu(z) \\ &= \sum_{A \in \alpha} \int_{f(A)} \frac{1}{J_\mu f(y_A)} \log J_\mu f(y_A)\, d\mu(z). \end{aligned}$$

Using the definition of Jacobian, and the assumption that f is injective on A, this gives

$$H_\mu(\alpha \mid \beta_\infty) = \sum_{A \in \alpha} \int_A \log J_\mu f(y)\, d\mu(y) = \int \log J_\mu f(y)\, d\mu(y),$$

as claimed. □

Proof of Proposition 6.1. Since the partition α is generating, $h_\mu(f) = h_\mu(f, \alpha)$. Then,

$$h_\mu(f, \alpha) = \lim_n H_\mu(\alpha \mid \beta_n) = H_\mu(\alpha \mid \beta_\infty),$$

by Theorem 4.14 of [Wa82]. Combined with the second part of Lemma 6.3, this gives $h_\mu(f) = \int \log J_\mu f\, d\mu$, as claimed. □

7 Existence

Here we prove that every maximal eigenmeasure is a maximizing measure. The first step is

Lemma 7.1. *If μ is a maximal eigenmeasure then $J_\mu f$ is constant equal to p.*

Proof. Let A be any measurable set such that $f|_A$ is injective. Take a sequence $\{g_n\} \in C(M)$ such that $g_n \to \chi_A$ at μ-almost every point and $\sup |g_n| \le 2$ for all n. By definition,

$$\mathcal{L} g_n(x) = \sum_{f(y)=x} g_n(y).$$

The last expression converges to $\chi_{f(A)}(x)$ at μ-almost every point. Hence, by the dominated convergence theorem,

$$\int p\, g_n\, d\mu = \int g_n\, d(\mathcal{L}^* \mu) = \int \mathcal{L} g_n\, d\mu \to \mu(f(A)).$$

Since the left hand side also converges to $\int_A p d\mu$, we conclude that

$$\mu(f(A)) = \int_A p d\mu,$$

which proves the lemma. $\square$

Lemma 7.2. *If μ is a maximal eigenmeasure then $h_\mu(f) \geq \log p$.*

Proof. We define the *dynamical ball* $\mathcal{B}_\epsilon(n,x)$ by

$$\mathcal{B}_\epsilon(n,x) = \{y \in M; d(f^i(x), f^i(y)) < \epsilon, \text{ for } i = 0, \dots, n-1\}.$$

If ϵ small enough so that $f^n|_{\mathcal{B}_\epsilon(n,x)}$ is injective, then:

$$1 = \mu(M) \geq \mu(f^n(\mathcal{B}_\epsilon(n,x))) = p^n \mu(\mathcal{B}_\epsilon(n,x)).$$

In particular, we may conclude that

$$-\limsup(1/n) \log \mu(\mathcal{B}_\epsilon(n,x)) \geq \log p$$

for every n and ϵ small. By the Brin-Katok local entropy formula (see [Mañ87])

$$h_\mu(f) = -\int \lim_{\epsilon \to 0} \lim_{n \to \infty} \frac{1}{n} \log \mu(\mathcal{B}_\epsilon(n,x))\, d\mu(x) \geq \log p.$$

This proves the lemma. $\square$

Corollary 7.3. *Every maximal eigenmeasure μ has entropy equal to* $\log p$.

Proof. By Lemma 7.2, the entropy is at least $\log p$. Then, we may apply Lemma 3.1 to conclude that all Lyapunov exponents of μ are positive. It follows, by Lemma 5.1, that μ admits generating partitions with small diameter. Hence, we may apply Proposition 6.1 and Lemma 7.1, to find that $h_\mu(f) = \int \log J_\mu d\mu = \log p$. $\square$

Lemma 7.4. *The topological entropy $h_{top}(f) = \log p$. Moreover, if η is any ergodic maximizing measure then the Jacobian $J_\eta f$ is constant equal to p.*

Proof. Consider any probability η such $h_\eta(f) \geq \log p$. By Lemma 3.1 all Lyapunov exponents of η are bigger than $c(f)$. Then, by Lemma 5.1, there exist generating partitions with arbitrarily small diameter. This ensures we may apply Proposition 6.1 to η. We get that

$$h_\eta(f) = \int \log J_\eta f \, d\eta.$$

Let us write $g_\eta = 1/(J_\eta f)$. The assumption that η is invariant means that

$$\sum_{f(y)=x} g_\eta(y) = 1$$

for η-almost every $x \in M$. From the previous equality, we find

$$0 \le h_\eta(f) - \log p = \int \log \frac{p^{-1}}{g_\eta} \, d\eta = \int \sum_{f(y)=x} g_\eta(y) \log \frac{p^{-1}}{g_\eta(y)} \, d\eta, \tag{7}$$

where the last equality uses $g_\eta = 1/J_\eta f$. Now, we use the following elementary fact from Calculus:

Lemma 7.5. *Let p_i, x_i, $i = 1, 2, \ldots, n$ be positive real numbers such $\sum_{i=1}^n p_i = 1$. Then $\sum_{i=1}^n p_i \log x_i \le \log(\sum_{i=1}^n p_i x_i)$ and the equality holds if and only if the x_i are all equal.*

Taking $p_i = g_\eta(y)$ and $x_i = p^{-1}/g_\eta(y)$, we obtain

$$\sum_{f(y)=x} g_\eta(y) \log \frac{p^{-1}}{g_\eta(y)} \le \log\Big(\sum_{f(y)=x} g_\eta(y) \frac{p^{-1}}{g_\eta(y)} \Big) = \log\Big(\sum_{f(y)=x} p^{-1} \Big) = 0$$

at η-almost every point. Since the integral is non-negative, by (7), the equality must hold η-almost everywhere, and $h_\eta(f) - \log p = 0$. Since η is arbitrary, this proves that $\log p = h_{top}(f)$.

¿From the last part of Lemma 7.5, we get that the values of $\log p^{-1}/g_\eta(y)$ are the same for all $y \in f^{-1}(x)$. In other words, for η-almost every $x \in M$ there exists a number $c(x)$ such that $p^{-1}/g_\eta(y) = c(x)$ for every $y \in f^{-1}(x)$. Then

$$\frac{1}{c(x)} = \sum_{y \in f^{-1}(x)} \frac{p^{-1}}{c(x)} = \sum_{y \in f^{-1}(x)} g_\eta(y) = 1$$

for η-almost every x. This means, precisely, that $J_\eta f(y) = p$ for every y on the pre-image of a full η-measure set. $\square$

8 Uniqueness

In this section we assume f is topologically mixing, and conclude that the maximizing measure is unique and supported on the whole ambient M. It suffices to consider ergodic measures, because the ergodic components of maximizing measures are also maximizing measures.

Lemma 8.1. *Any ergodic maximizing measure μ is supported on the whole M.*

Proof. Suppose $\mu(U) = 0$ for some non-empty open set U. By the mixing assumption, there exists $N \ge 1$ such $f^N(U) = M$. Partitioning U into subsets $U_1, \ldots, U_k$ such that ever $f^N | U_j$ is injective, we get that

$$\mu(f^N(U_j)) = \int_{U_j} J_\mu f^N \, d\mu = 0$$

for $j = 1, \ldots, k$. Recall Lemma 7.4. This implies that $M = \mu((f^N(U)) = 0$, which is a contradiction. $\square$

This has the following useful consequence: given any $\delta > 0$ there exist $b = b(\delta) > 0$ such that

$$\mu(B(x,\delta)) \geq b \quad \text{for all } x \in M. \tag{8}$$

Indeed, if there were points such that the balls of radius δ around them have arbitrarily small measures then, considering an accumulation point, one would get a ball with zero measure, and that would contradict Lemma 8.1.

Now let μ_1 and μ_2 be any two ergodic maximizing measures. Our goal is to prove that the two measures coincide. As a first step we prove that they are equivalent. For this, we fix any (finite) partition $\mathcal{P}$ of M into subsets P such that P has non-empty interior, and the boundary ∂P has zero measure for both μ_1 and μ_2. Fixing $\delta > 0$ small so that every $P \in \mathcal{P}$ contains some ball of radius δ, and applying (8) to both measures, we conclude that there exists $B > 0$ such that

$$\mu_1(P) \leq B\mu_2(P) \text{ and } \mu_2(P) \leq B\mu_1(P) \quad \text{for all } P \in \mathcal{P}. \tag{9}$$

Now let g be an inverse branch of any iterate f^n, $n \geq 1$. Using Lemma 7.4, we get that $\mu_i(P) = p^n \mu_i(g(P))$ for $i = 1, 2$. It follows that (9) remains valid for the images $g(P)$:

$$\mu_1(g(P)) \leq B\mu_2(g(P)) \text{ and } \mu_2(g(P)) \leq B\mu_1(g(P)) \tag{10}$$

for every $P \in \mathcal{P}$ and every inverse branch g of f^n, for any $n \geq 1$. We denote by $\mathcal{Q}$ the family of all such images $g(P)$.

Lemma 8.2. *Given any measurable set $E \subset M$ and any $\varepsilon > 0$ there exists a family $\mathcal{E}$ of pairwise disjoint elements of $\mathcal{Q}$ such that*

$$\mu_i\big(E \setminus \bigcup_{\mathcal{E}} g(P)\big) = 0 \quad \textit{and} \quad \mu_i\big(\bigcup_{\mathcal{E}} g(P) \setminus E\big) \leq \varepsilon \textit{ for } i = 1, 2.$$

Proof. By Lemma 3.1, all Lyapunov exponents of μ_i are larger than $c(f)$. Hence, by Lemma 4.3 and the remark following it, there exists $N \geq 1$ and $\theta > 0$ such that μ_i-almost every point has density $> 2\theta$ of hyperbolic times.

Let U_1 be an open set and K_1 be a compact set such that $K_1 \subset E \subset U_1$ and $\mu_i(U_1 \setminus E) \leq \varepsilon$ for $i = 1, 2$ and $\mu_i(K_1) \geq (1/2)\mu(E)$. Using Lemma 4.4 with $B = K_1$ and $\nu = \mu_i/\mu_i(K_1)$, we may find $n_1 \geq 1$ such that $e^{-cn_1} < d(K_1, U_1^c)$ and the subset L_1 of points $x \in K_1$ for which n_1 is a hyperbolic time satisfies $\mu_i(L_1) \geq \theta\mu_i(K_1) \geq (\theta/2)\mu_i(E)$. Let $\mathcal{E}_1$ the family of all $g(P)$ that intersect L_1, with $P \in \mathcal{P}$ and g an inverse branch of f^{n_1}. Notice that the elements of $\mathcal{E}_1$ are pairwise disjoint, because the elements of P are pairwise disjoint. Moreover, by Lemma 4.2, their diameter is less than e^{-cn_1}. Thus, the union E_1 of all the elements of $\mathcal{E}_1$ is contained in U_1. By construction, it satisfies

$$\mu_i(E_1 \cap E) \geq \mu_i(L_1) \geq \theta\mu_i(K_1) \geq (\theta/2)\mu_i(E).$$

Next, consider the open set $U_2 = U_1 \setminus \overline{E}_1$ and let $K_2 \subset E \setminus \overline{E}_1$ be a compact set such that $\mu_i(K_2) \geq (1/2)\mu_i(E \setminus E_1)$. Observe $\mu_i(\overline{E}_1 \setminus E_1) = 0$ because the boundaries of the atoms of $\mathcal{P}$ have zero measure and that is preserved by the inverse branches, since μ_i is invariant. Reasoning as before, we may find $n_2 > n_1$ such that $e^{-cn_2} < d(K_2, U_2^c)$ and a set $L_2 \subset K_2$ such that $\mu_i(L_2) \geq \theta\mu_i(K_2)$ and n_2 is a hyperbolic time for every $x \in L_2$. Denote by $\mathcal{E}_2$ the family of inverse images $g(P)$ that intersect L_2, with $P \in \mathcal{P}$ and g an inverse branch of f^{n_2}. As before, the elements of $\mathcal{E}_2$ are pairwise disjoint, and their diameters are smaller than e^{-cn_2}. The latter ensures that their union E_2 is contained in U_2. Consequently, the elements of the union $\mathcal{E}_1 \cup \mathcal{E}_2$ are also pairwise disjoint. Moreover,

$$\mu_i\big(E_2 \cap [E \setminus E_1]\big) \geq \mu_i(L_2) \geq \theta\mu_i(K_2) \geq (\theta/2)\mu_i(E \setminus E_1).$$

Repeating this procedure, we construct families $\mathcal{E}_k$, $k \geq 1$ of elements of $\mathcal{Q}$ such that their elements are all pairwise disjoint and contained in U_1, and

$$\mu_i\big(E_{k+1} \cap [E \setminus (E_1 \cup \cdots \cup E_k)]\big) \geq (\theta/2)\mu_i\big(E \setminus (E_1 \cup \cdots \cup E_k)\big) \qquad (11)$$

for all $k \geq 1$, where $E_j = \cup_{\mathcal{E}_j} g(P)$. Thus, $\mu_i\big(\bigcup_{k=1}^\infty E_k \setminus E\big) \leq \mu_i(U_1 \setminus E) \leq \varepsilon$ for $i = 1, 2$, and (11) implies that

$$\mu_i(E \setminus \bigcup_{k=1}^{\infty} E_k) = 0.$$

This completes the proof of the lemma, with $\mathcal{E} = \bigcup_{k=1}^\infty \mathcal{E}_k$. $\square$

Remark 8.3. The lemma remains true if one asks that $\mu_i\big(E \setminus \bigcup_{\mathcal{E}} g(P)\big) = 0$ for both $i = 1, 2$. This follows from a variation of the previous construction, considering each one of the two measures alternately: for each $k \geq 1$ consider $i \equiv k \mod 2$; then ask that $\mu_i(K_k) \geq (1/2)\mu_i(E \setminus [E_1 \cup \cdots \cup E_k])$, and choose n_k such that $\mu_i(L_k) \geq \theta\mu(K_k)$. The same kind of argument applies with any number of probability measures $\mu_1, \ldots, \mu_r$. These extensions will not be used here.

Combining (10) with Lemma 8.2, we get that, for any measurable set $E \subset M$,

$$\mu_1(E) \leq \varepsilon + \sum_{\mathcal{E}} \mu_1(g(P)) \leq \varepsilon + B \sum_{\mathcal{E}} \mu_2(g(P)) = \varepsilon + B\mu_2(E).$$

As $\varepsilon > 0$ is arbitrary, we get that $\mu_1(E) \leq B\mu_2(E)$. A symmetric argument gives that $\mu_2(E) \leq B\mu_1(E)$ for any measurable set E. This implies that $\mu_1 = h\mu_2$ where the Radon-Nikodym derivative h satisfies $B^{-1} \leq h \leq B$. Since μ_1 and μ_2 are invariant measures,

$$\mu_1 = f_*\mu_1 = (h \circ f) f_*\mu_2 = (h \circ f)\mu_2.$$

As the Radon-Nikodym derivative is essentially unique, we get that $h = h \circ f$ at μ_2-almost every point. By ergodicity, it follows that h is constant almost everywhere. Since the μ_i are both probabilities, we get that $h = 1$ and so $\mu_1 = \mu_2$. This proves uniqueness of the maximizing measure.

References

[Alv00] J. F. Alves. SRB measures for non-hyperbolic systems with multidimensional expansion. *Ann. Sci. École Norm. Sup.*, 33:1–32, 2000.

[AAS03] J. F. Alves, V. Araújo, and B. Saussol. On the uniform hyperbolicity of some nonuniformly hyperbolic systems. *Proc. Amer. Math. Soc.*, 131:1303–1309, 2003.

[ABV00] J. F. Alves, C. Bonatti, and M. Viana. SRB measures for partially hyperbolic systems whose central direction is mostly expanding. *Invent. Math.*, 140:351–398, 2000.

[AMO04] A. Arbieto, C. Matheus and K. Oliveira. Equilibrium states for random non-uniformly expanding maps *Nonlinearity*, 17:581–593, 2004.

[Bo75] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lect. Notes in Math.* Springer Verlag, 1975.

[Buz99] J. Buzzi. Markov extensions for multi-dimensional dynamical systems. *Israel J. Math.*, 112:357–380, 1999.

[Buz00] J. Buzzi. On entropy-expanding maps. Preprint École Polythecnique.

[Cao03] Y. Cao. Non-zero Lyapunov exponents and uniform hyperbolicity. *Nonlinearity*, 16:1473–1480, 2003.

[Mañ87] R. Mañé. *Ergodic theory and differentiable dynamics*. Springer Verlag, 1987.

[Ol03] K. Oliveira. Equilibrium states for non-uniformly expanding maps. *Ergodic Theory & Dynamical Systems*, 23:1891-1906, 2003.

[Os68] V. I. Oseledets. A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems *Trans. Moscow Math. Society*, 19: 197-231, 1968.

[Rok67] V.A. Rokhlin. Lectures on the entropy theory of measure-preserving transformations. *Russ. Math. Surveys*, 22-5:1–52, 1967. Transl. from Uspekhi Mat. Nauk. 22 - 5 (1967), 3–56.

[Rue78] D. Ruelle. An inequality for the entropy of differentiable maps. *Bull. Braz. Math. Soc.*, 9:83–87, 1978.

[Rue89] D. Ruelle. The thermodynamical formalism for expanding maps. *Comm. Math. Phys.*, 125:239–262, 1989.

[Wa82] P. Walters. *An introduction to ergodic theory.* Springer Verlag, 1982

Krerley Oliveira (krerley@mat.ufal.br)
Departamento de Matemática - UFAL, Campus A.C. Simões, s/n 57072-090
Maceió, Alagoas - Brazil

Marcelo Viana (viana@impa.br)
IMPA, Est. D. Castorina 110
22460-320 Rio de Janeiro, RJ, Brazil