GROUND STATES ARE GENERICALLY A PERIODIC ORBIT

GONZALO CONTRERAS

Abstract. We prove that for an expanding transformation the maximizing measures of a generic Lipschitz function are supported on a single periodic orbit.

1. Introduction

Let X be a compact metric space and $T : X \to X$ an expanding map. This means that there are numbers $d \in \mathbb{Z}^+$, $0 < \lambda < 1$ such that for every point $x \in X$ there is a neighborhood U_x of x in X and continuous branches S_i, $i = 1, \dots, \ell_x \le d$ of the inverse of T such that $T^{-1}(U_x) = \cup_{i=1}^{\ell_x} S_i(U_x)$, $T \circ S_i = I_{U_x}$ $\forall i$, and

$$d\big(S_i(y), S_i(z)\big) \le \lambda\, d(y,z) \qquad \forall y,\, z \in U_x.$$

Given a continuous function $F : X \to \mathbb{R}$, a *maximizing measure* or a *ground state* is a T-invariant Borel probability measure μ which maximizes the integral of F among all T-invariant Borel probabilities:

$$\int F\, d\mu = \sup \Big\{ \int F\, d\nu \;\Big|\; \nu \in \mathcal{M}(T) \Big\},$$

where

$$\mathcal{M}(T) = \big\{\ T\text{-invariant Borel probabilities in } X\ \big\}.$$

They are called ground states because they correspond to the usual variational principle in ergodic theory

$$\mu_F := \arg\max \Big\{ h_\mu(T) + \int F\, d\mu \;\Big|\; \mu \in \mathcal{M}(T) \Big\},$$

without the entropy term $h_\mu(T)$. The measure μ_F is called the *equilibrium state* for F.

They are also the candidates for zero temperature limits of equilibrium states. This is, limits of the form $\lim\limits_{\beta \to +\infty} \mu_{\beta F}$. Here β is interpreted as the inverse of the temperature. It is known [8, Proposition 29] that if the limit of a sequence $\{\mu_{\beta_k F}\}_k$ with $\beta_k \to \infty$ exists, then it has to be a maximizing measure with maximal entropy among the maximizing

Gonzalo Contreras was Partially supported by conacyt, Mexico, grant 178838.

measures. Brémont [4] proves that the limit $\lim_{\beta\to+\infty}\mu_{\beta F}$ exists if F is locally constant. Chazottes, Gambaudo and Ugalde [7] give a characterization of the limit and a new proof of Brémont's result. Lepaideur [13] proves the convergence for general Hölder functions F.

For generic Hölder or Lipschitz functions F, the maximizing measure is unique. This is proven in [8] and it is presented in a general version in [12]. The ideas came from an analogous result for lagrangian systems by Mañé [14]. After Jenkinson lecture notes [14] the study of maximizing measures for a fixed dynamical system became known as Ergodic Optimization. Surveys of the subject are presented by Jenkinson [12] and Baraviera, Leplaideur, Lopes [1].

1.1. **Theorem** (Contreras, Lopes, Thieullen [8], see also[12]). *Let $T : X \to X$ be a continuous map of a compact metric space. Let E be a topological vector space which is densely and continuously embedded in $C^0(X,\mathbb{R})$. Write*

$$\mathcal{U}(E) := \left\{ F \in E \;\middle|\; \text{there is a unique } F\text{-maximizing measure} \right\}.$$

Then $\mathcal{U}(E)$ is a countable intersection of open and dense sets.

If moreover E is a Baire space, then $\mathcal{U}(E)$ is dense in E.

The main conjecture in Ergodic Optimization during the last decade have been wether the maximizing measure for generic Hölder or Lipschitz functions F is supported on a periodic orbit. For lagrangian systems an analogous statement is known as Mañé's conjecture.

On the space $\mathrm{Lip}(X,\mathbb{R})$ of Lipschitz functions on X we use the norm

$$\|f\| := \sup_{x\in X}|f(x)| + \sup_{x\neq y}\frac{|f(x)-f(y)|}{d(x,y)}. \tag{1}$$

We denote the second term in (1) as $\mathrm{Lip}(f)$.

Recall that using a Markov partition (cf. [17, §7.29]) the symbolic model for an expanding map is given by a one-sided subshift of finite type. Here we prove

Theorem A. *If X is a compact metric space and $T : X \hookleftarrow$ is an expanding map then there is an open and dense set $\mathcal{O} \subset \mathrm{Lip}(X,\mathbb{R})$ such that for all $F \in \mathcal{O}$ there is a single F-maximizing measure and it is supported on a periodic orbit.*

Corollary B. *For an open and dense set $\mathcal{O}$ of Lipschitz functions F on X the zero temperature limit $\lim_{\beta\to+\infty}\mu_{\beta F}$ exists and it is supported on a single periodic orbit.*

On the negative side, for expanding transformations Bousch [3, Proposition 9, p. 306] proves that for generic continuous functions the maximizing measure is not supported on a periodic orbit. The case of hyperbolic sets is presented by Jenkinson in [12, Theorem 4.2].

There have been several approaches to the conjecture from which we will use some of their techniques. Write

$$\mathcal{P}(E) := \big\{\, F \in E \;\big|\; \text{the } F\text{-maximizing measure is supported on a periodic orbit } \big\}.$$

Contreras, Lopes, Thieullen [8] prove that $\mathcal{P}(E)$ is open for $E = C^{\alpha}(X,\mathbb{R})$ the space of α-Hölder continuous functions and in the α-Hölder topology it is open and dense in $E = C^{!\alpha}(X,\mathbb{R})$, the space of functions $F : X \to \mathbb{R}$ such that

$$\forall \eta > 0 \quad \exists \varepsilon > 0 \qquad d(x,y) < \varepsilon \implies |F(x) - F(y)| < \eta\, d(x,y)^{\alpha}.$$

The main technique is the introduction of a sub-action $u : X \to \mathbb{R}$ to transform the function F to a cohomologous function $G = F + u \circ T - u$ such that $G \le a = \int G\, d\mu^G$, where μ^G is a maximizing measure for G and F. The sub-action is defined similarly, and plays the same role, as a sub-solution of the Hamilton-Jacobi equation for Lagrangian systems. In fact analogous constructions to the weak KAM theory can be translated to this setting. In proposition 2.1 we construct a sub-action following the original method by Fathi [9] to construct weak KAM solutions. This method was used in ergodic optimization by Bousch in [2]. In fact many results from Lagrangians systems can be translated to the ergodic optimization setting, see for example [10].

Bousch proves that $\mathcal{P}(E)$ is dense for Walters functions. Yuan and Hunt [19] prove that if a fixed measure is maximizing for an open set of functions F in the Lipschitz topology, then it is supported on a periodic orbit. Their method of perturbation is the basis of the present work. Quas and Siefken [16] work in a one-sided shift. They prove that $\mathcal{P}(E)$ contains an open and dense set if E is the space of super-continuous functions. They present an elegant version of the method of Yuan and Hunt. We need to modify it for Lipschitz functions and pseudo-orbits with finitely many jumps in Proposition 2.4.

Another ingredient of the proof is the following theorem. As a weak version of the conjecture, Morris [15] proves

1.2. **Theorem** (Morris [15])**.** *Let X be a compact metric space and $T : X \hookleftarrow$ an expanding map. There is a residual set $\mathcal{G} \subset \mathrm{Lip}(X,\mathbb{R})$ such that if $F \in \mathcal{G}$ then there is a unique F-maximizing measure and it has zero metric entropy.*

The original version of Theorem 1.2 is for Hölder functions in a shift of finite type. In appendix A we describe the modifications from the proof in [15] needed to obtain Theorem 1.2.

2. Preliminars

Let $e_0 > 0$ and $0 < \lambda < 1$ be such that for every $x \in X$ the branches of the inverse of T are well defined, injective and are λ-contractions on the ball of radius e_0 centered at x.

Given $F \in \mathrm{Lip}(X, \mathbb{R})$, the Lax operator for F is $\mathcal{L}_F : \mathrm{Lip}(X, \mathbb{R}) \hookleftarrow$

$$\mathcal{L}_F(u)(x) = \max_{y \in T^{-1}(x)} \big\{ \alpha + F(x) + u(x) \big\},$$

where

$$\alpha = -\max_{\mu \in \mathcal{M}(T)} \int F \, d\mu.$$

Denote the set of maximizing measures by

$$\mathcal{M}(F) := \left\{ \mu \in \mathcal{M}(T) \;\middle|\; \int F \, d\mu = -\alpha(F) \right\}.$$

A *calibrated sub-action* for F is a fixed point of the Lax operator $\mathcal{L}_F$. If $\mathcal{L}_F(u) = u$, writing

$$\overline{F} := F + \alpha + u - u \circ T \tag{2}$$

it is easy to see that

(i) $-\alpha(\overline{F}) = \max_{\mu \in \mathcal{M}(T)} \int \overline{F} \, d\mu = 0.$

(ii) $\overline{F} \le 0.$

(iii) $\mathcal{M}(F) = \mathcal{M}(\overline{F}) = \{\, T\text{-invariant measures supported on } [\overline{F} = 0] \,\}$

2.1. **Proposition.** *There exists a Lipschitz calibrated sub-action.*

Proof: We claim that

$$\mathrm{Lip}(\mathcal{L}_F(u)) \le \lambda \big(\mathrm{Lip}(u) + \mathrm{Lip}(F) \big).$$

Indeed, given x, $y \in X$, let $\overline{y} \in T^{-1}(y)$ be such that $\mathcal{L}_F(u)(y) = \alpha + F(\overline{y}) + u(\overline{y})$. Let S be the branch of the inverse of T such that $S(y) = \overline{y}$. Let $\overline{x} := S(x)$. Then

$$\begin{aligned}\mathcal{L}_F u(x) - \mathcal{L}_F u(y) &\le \alpha + F(\overline{x}) + u(\overline{x}) - \alpha - F(\overline{y}) - u(\overline{y}) \\ &\le \big(F(\overline{x}) - F(\overline{y})\big) + \big(u(\overline{x}) - u(\overline{y})\big) \\ &\le \big(\operatorname{Lip}(F) + \operatorname{Lip}(u)\big)\, d(\overline{x}, \overline{y}) \\ &\le \big(\operatorname{Lip}(F) + \operatorname{Lip}(u)\big)\lambda\, d(x, y).\end{aligned}$$

The other inequality is similar.

Thus $\mathcal{L}_F$ leaves invariant the space

$$\mathbb{E} := \left\{ u \in \operatorname{Lip}(X, \mathbb{R}) \;\middle|\; \operatorname{Lip}(u) \le \frac{\lambda \operatorname{Lip}(F)}{1-\lambda} \right\}.$$

The quotient space $\mathbb{E}/\{\text{constants}\}$ is compact and convex and on it $\mathcal{L}_F$ is continuous. By Schauder Theorem [11, Theorem 18.10, p. 197] $\mathcal{L}_F$ has a fixed point in $\mathbb{E}$.

In fact $\mathcal{L}_F$ is non-expanding in the supremum norm and a simpler fixed point applies [11, Theorem 3.1, p. 28].

□

If u is a calibrated sub-action, every point $z \in X$ has a *calibrating pre-orbit*, $(z_k)_{k\le 0}$ such that $T^i(z_{-i}) = z_0 = z$ and

$$u(z_{k+1}) = u(z_k) + \alpha + F(z_k), \qquad \forall k \le -1. \tag{3}$$

Or equivalently, since $T(z_k) = z_{k+1}$,

$$\overline{F}(z_k) = 0, \quad \forall k \le -1.$$

2.2. **Remark.** If ν is a maximizing measure then by (iii) its support is contained in $[\overline{F} = 0]$. If $z \in \operatorname{supp}(\nu)$ then there is a pre-orbit of z which is included in $\operatorname{supp}(\nu)$. Since $\operatorname{supp}(\nu) \subset [\overline{F} = 0]$ the pre-orbit must calibrate u. In proposition 2.4 we will obtain a periodic orbit $\mathcal{O}(y)$ such that the α-limit of every calibrating pre-orbit is $\mathcal{O}(y)$. This implies that every maximizing measure has support on $\mathcal{O}(y)$.

The iteration of equality (3) gives

$$\forall k \le -1, \quad u(z_0) = u(z_{-k}) + k\alpha + \sum_{i=-k}^{-1} F(z_i)$$

for any calibrating pre-orbit.

We say that a sequence $(x_n)_{n\in\mathbb{N}} \subset X$ is a *δ-pseudo-orbit* if $d(x_{n+1}, T(x_n)) \le \delta$, $\forall n \in \mathbb{N}$.

We say that the orbit of y *ε-shadows* a pseudo-orbit $(x_n)_{n\in\mathbb{N}}$ if $\forall n \in \mathbb{N}$, $d(T^n(y), x_n) < \varepsilon$.

2.3. **Proposition** (Shadowing Lemma)**.**

If $(x_k)_{k\in\mathbb{N}}$ is a δ-pseudo-orbit then there is $y \in X$ whose orbit ε-shadows (x_k) with $\varepsilon = \frac{\delta}{1-\lambda}$. If (x_k) is a periodic pseudo-orbit then y is a periodic orbit with the same period.

Proof: Write $B(x, r) := \{\, z \in X \mid d(z, x) \le r \,\}$ and $a := \frac{\lambda\,\delta}{1-\lambda}$. Let S_k be the branch of the inverse of T such that $S_k(T(x_k)) = x_k$. We have that

$$S_k\big(B(x_{k+1}, a)\big) \subseteq S_k\big(B(T(x_k), a+\delta)\big) \subseteq B\big(x_k, \lambda(a+\delta)\big) = B(x_k, a).$$

Let $y \in X$ be given by

$$y \in \bigcap_{k=0}^{\infty} S_0 \circ \cdots \circ S_k\big(B(x_{k+1}, a)\big).$$

The point y exists and is unique because it is the intersection of a nested family of non-empty compact sets with diameter smaller than $2a\lambda^k$. We have that $T^k(y) \in B(x_k, a)$. Thus y a-shadows (x_k). Now suppose x_k is p-periodic. Then also $T^p(y)$ a-shadows (x_k). The uniqueness of y implies that $T^p(y) = y$.

$\square$

We show now a condition which permits to obtain a perturbation with maximizing measure supported on a periodic orbit. The argument appeared first in Yuan and Hunt [19]. The proof below is a modification that we shall need of the arguments by Quas and Siefken [16] which we adapt to pseudo-orbits.

Let $y \in \mathrm{Per}(T) = \cup_{p\in\mathbb{N}^+} \mathrm{Fix}(T^p)$ be a periodic point for T. Let P_y be the set of Lipschitz functions $F \in \mathrm{Lip}(X, \mathbb{R})$ such that there is a unique F-maximizing measure and it is supported on the positive orbit of y. Let $\overset{\circ}{P}_y$ be the interior of P_y in $\mathrm{Lip}(X, \mathbb{R})$.

2.4. **Proposition.** *Let $F, u \in \mathrm{Lip}(X, \mathbb{R})$ with $\mathcal{L}_F(u) = u$ and let $\overline{F}$ be defined by (2).*

Let $M \in \mathbb{N}^+$. Suppose that for any $\delta > 0$ there is a $p(\delta)$-periodic δ-pseudo orbit $(x_k^\delta)_k$ in $[\overline{F} = 0]$ with at most M jumps such that if $\gamma_\delta := \min_{0\le i<j<p(\delta)} d(x_i^\delta, x_j^\delta)$ then $\lim_{\delta\to 0} \frac{\gamma_\delta}{\delta} = \infty$.

Then F is in the closure of $\cup_{y\ \text{periodic}} \overset{\circ}{P}_y$.

Proof: Observe that fixing u, for any $H \in \mathrm{Lip}(X,\mathbb{R})$ the functions H and $H + \alpha(F) + u - u\circ T$ have the same maximizing measures. Therefore it is enough to prove that the function $\overline{F}$ is in the closure of $\cup_{y \text{ periodic}} \overset{\circ}{P}_y$.

Let $\varepsilon > 0$. We will show a perturbation of F with Lipschitz norm smaller than ε such that it has a unique maximizing measure supported on a periodic orbit. Moreover, we will exhibit a neighborhood of the perturbed function in which the same periodic orbit is the unique maximizing measure for all functions in the neighborhood. The neighborhood will depend on the periodic orbit.

Let

$$K := \max\left\{ \frac{M\,\mathrm{Lip}(\overline{F})}{(1-\lambda)^2}, \frac{\mathrm{Lip}(\overline{F})+2}{1-\lambda} \right\},$$
$$\gamma_3 := \tfrac{1}{\mathrm{Lip}(T)} \left(\gamma - \tfrac{2\delta}{1-\lambda} - \delta\right).$$

Assume that $\frac{\gamma}{\delta}$ is so large that

$$3K\delta - \varepsilon\gamma_3 =: -2a < 0. \tag{4}$$

Let y be the p-periodic point which $\left(\frac{\delta}{1-\lambda}\right)$-shadows (x_k). Write $y_k := T^k(y)$ and

$$\mathcal{O}(y) = \{T^i(y) \mid i = 0,\dots,p-1\} = \{y_0,\dots,y_{p-1}\}.$$

For a function $G : X \to \mathbb{R}$ write

$$\langle G\rangle(y) = \frac{1}{p}\sum_{i=0}^{p-1} G(T^i(y)).$$

Let n_i, $i = 1,\dots,\ell$, $\ell \le M$, be the jumps of (x_k); i.e. $d\big(T(x_k), x_{k+1}\big) = 0$ if $k \notin \{n_1,\dots,n_\ell\}$. We have that

$$\left| \sum_{k=n_{i-1}}^{n_i} \overline{F}(y_k) - \sum_{k=n_{i-1}}^{n_i} \overline{F}(x_k) \right| \le \sum_{k=n_{i-1}}^{n_i} \mathrm{Lip}(\overline{F})\; d(y_k,x_k) \le \sum_{k=0}^{n_i-n_{i-1}} \lambda^k \,\frac{\delta}{1-\lambda}\; \mathrm{Lip}(\overline{F})$$
$$\le \frac{\mathrm{Lip}(\overline{F})}{(1-\lambda)^2}\,\delta.$$

Thus

$$\left| \sum_{k=0}^{p-1} \overline{F}(y_k) - \sum_{k=0}^{p-1} \overline{F}(x_k) \right| \le \frac{M\;\mathrm{Lip}(\overline{F})}{(1-\lambda)^2}\,\delta.$$

By hypothesis $\forall k$, $\overline{F}(x_k)=0$, thus $\sum_0^{p-1}\overline{F}(x_k)=0$. Therefore

$$\sum_{k=0}^{p-1}\overline{F}(y_k)\ge -\frac{M\ \mathrm{Lip}(\overline{F})}{(1-\lambda)^2}\ \delta\ge -K\delta,$$

$$\langle\overline{F}\rangle(y)\ge -\frac{K\delta}{p}. \tag{5}$$

Observe that if $0\le i<j<p$,

$$d(y_i,y_j)\ge -d(y_i,x_i)+d(x_i,x_j)-d(x_j,y_j)>\gamma-\frac{2\,\delta}{1-\lambda}=:\gamma_2.$$

Claim: Assume that $d(z,y_k)\le\delta<e_0$. *Take* $w_1\in T^{-1}\{z\}$ *such that* $d(w_1,y_{k-1})<\lambda\delta$. *If* $w_2\in T^{-1}\{z\}\setminus\{w_1\}$ *then*

$$d(w_2,\mathcal{O}(y))\ge\gamma_3:=\frac{\gamma_2-\delta}{\mathrm{Lip}(T)}\gg\delta.$$

Proof: Let $y_j\in\mathcal{O}(y)$ be such that $d(w_2,\mathcal{O}(y))=d(w_2,y_j)$. Observe that $j\ne k-1$ because T is injective in the ball $d(\,\cdot\,,y_{k-1})\le\lambda\delta$. Then

$$\gamma_2<d(y_k,y_{j+1})\le d(y_k,z)+d(z,y_{j+1})\le\delta+\mathrm{Lip}(T)\,d(w_2,y_j).$$

This proves the claim.

Now we make two perturbations to $\overline{F}$. The first perturbation is the addition of $-\varepsilon g(x)$, where

$$g(x):=d(x,\mathcal{O}(y)),$$

and ε depends on δ and γ. This is a perturbation with $\mathrm{Lip}(\varepsilon g)=\varepsilon$. The second is a perturbation by any function with norm

$$\|h\|_0<\frac{K\delta}{2p}. \tag{6}$$

This perturbation depends on $\mathcal{O}(y)$, and in particular on its period p. We shall prove that the function $G_1:=\overline{F}-\varepsilon g+h$ has a unique maximizing measure supported on the periodic orbit $\mathcal{O}(y)$. Since the set of such functions G_1 contains an open ball centered at $\overline{F}-\varepsilon g$, this proves the proposition.

Let

$$G=\overline{F}-\varepsilon g+h+\beta=G_1+\beta, \tag{7}$$

where

$$\beta = -\sup_{\mu \in \mathcal{M}(T)} \int (\overline{F} - \varepsilon g + h)\, d\mu.$$

It is enough to prove the claim for G because G and G_1 have the same maximizing measures.

Using (5), we have that

$$\begin{aligned} \beta &\le -\langle \overline{F} - \varepsilon g + h\rangle(y) = -\langle \overline{F} + h\rangle(y) \\ &\le -\langle \overline{F}\rangle(y) + \|h\|_0 \\ &\le \frac{K\delta}{p} + \|h\|_0 \end{aligned} \tag{8}$$

Let v be a calibrated sub-action for G, $\mathcal{L}_G(v) = v$. Given any $z \in X$ let $(z_k)_{k\le 0}$ be a pre-orbit of z which calibrates v. Let $0 > t_1 > t_2 > \cdots$ be the times on which $d(z_k, \mathcal{O}(y)) > \delta$. Then for each $n \in \mathbb{N}^+$ there is $s_n \in \mathbb{Z}$ such that the orbit segment $(z_k)_{k=t_{n+1}+1}^{t_n-1}$ δ-shadows $(y_{-i+s_n})_{i=t_n-t_{n+1}-1}^{1}$, thus

$$d(z_{-i+t_{n+1}}, y_{-i+s_{n+1}}) \le \lambda^{i-1}\,\delta, \quad \forall n \in \mathbb{N}, \quad \forall i = 1, \dots, t_n - t_{n+1} - 1.$$

By the Claim, we have that

$$d(z_{t_n}, \mathcal{O}(y)) \ge \gamma_3.$$

Since both terms in $\overline{F} - \varepsilon g$ are non-positive, from (7) and (8) we obtain

$$G \le h + \beta \le \frac{K\delta}{p} + 2\,\|h\|_0\,. \tag{9}$$

On a shadowing segment we have

$$\left| \sum_{t_{n+1}+1}^{t_n-1} G(z_k) - \sum_{s_n-t_n+t_{n+1}-1}^{s_n-1} G(y_k) \right| \le \mathrm{Lip}(G) \sum_{i=0}^{+\infty} \lambda^i\,\delta \le \mathrm{Lip}(G)\, \frac{\delta}{1-\lambda} \le K\delta.$$

Write

$$t_n - t_{n+1} - 1 = mp + r$$

with $0 \le r < p$ and separate the shadowing segment in m loops around the orbit $\mathcal{O}(y)$ and a residue with at most $p-1$ iterates. Using (9) for $(p-1)$ times, we have that

$$\sum_{t_{n+1}+1}^{t_n-1} G(z_k) \le mp\,\langle G\rangle(y) + (p-1)\frac{K\delta}{p} + 2(p-1)\,\|h\|_0 + \mathrm{Lip}(G)\,\frac{\delta}{1-\lambda}.$$

By the definition of β we have that $\langle G\rangle(y)\le 0$. Therefore

$$\sum_{t_{n+1}+1}^{t_n-1} G(z_k) \le (p-1)\frac{K\delta}{p} + 2(p-1)\left\|h\right\|_0 + K\delta. \tag{10}$$

On the points z_{t_n} far from $\mathcal{O}(y)$, using (9) we have that

$$G(z_{t_n}) \le 0 - \varepsilon\,\gamma_3 + \left\|h+\beta\right\|_0 \le -\varepsilon\,\gamma_3 + \frac{K\delta}{p} + 2\left\|h\right\|_0 . \tag{11}$$

Thus, adding (10) and (11), and using (6) and (4)

$$\sum_{t_{n+1}}^{t_n-1} G(z_k) \le 2p\,\left\|h\right\|_0 + 2K\delta - \varepsilon\gamma_3 < -a < 0. \tag{12}$$

Since v is finite and (z_k) is a calibrating pre-orbit for v, we have that for every $k\le 0$,

$$v(z) = v(z_k) + \sum_{i=k+1}^{-1} G(z_i)$$

Thus

$$\sum_{-\infty}^{-1} G(z_k) \ge -2\left\|v\right\|_0 > -\infty.$$

From (12) we obtain that the sequence t_n is finite. Thus every calibrating pre-orbit has α-limit $\mathcal{O}(y)$. By remark 2.2, this implies that every maximizing measure for G has support on $\mathcal{O}(y)$.

□

3. Proof of Theorem A

Proof of theorem A:

We prove that $\cup_{y\in\mathrm{Per}(T)}\overset{\circ}{P}_y$ is open and dense. It is clearly open.

Suppose that there is a non-empty open set

$$\mathcal{U}\subset \mathrm{Lip}(X,\mathbb{R}) \tag{13}$$

which is disjoint from $\cup_{y\in\mathrm{Per}(T)}\overset{\circ}{P}_y$. Let $F\in\mathcal{U}$. Let μ be an ergodic maximizing measure for F. By Theorem 1.2 we can assume that the entropy of μ is

$$h_\mu(T) = 0. \tag{14}$$

By (iii) for any calibrating subaction u, $\operatorname{supp}(\mu) \subset [\overline{F} = 0]$. Let $q \in X$ be a generic point for μ, i.e. for any continuous function $f : X \to \mathbb{R}$,

$$\int f \, d\mu = \langle f \rangle(q) = \lim_N \frac{1}{N} \sum_{i=0}^{N-1} f(T^i(q)).$$

By Proposition 2.4 with $M = 2$ there is $Q > 0$ and $\delta_0 > 0$ such that if $(x_k)_{k\geq 0} \subset \mathcal{O}(q)$ is a p-periodic δ-pseudo-orbit with at most 2 jumps made with elements of the positive orbit of q and $0 < \delta < \delta_0$ then $\gamma = \min_{1\leq i<j<p} d(x_i, x_j) < Q\delta$.

Let N_0 be such that

$$2\, Q^{-N_0} < \delta_0. \tag{15}$$

Fix a point $w \in \operatorname{supp}(\mu)$ for which Brin-Katok Theorem holds [6], i.e.

$$h_\mu(T) = -\lim_{L\to+\infty} \frac{1}{L} \log \mu\big(V(w, L, \varepsilon)\big), \tag{16}$$

where $V(w, L, \varepsilon)$ is the *dynamic ball*:

$$V(w, L, \varepsilon) := \big\{\, x \in X \;\big|\; d(T^k x, T^k w) < \varepsilon\,,\ \forall k = 0, \dots, L \,\big\}. \tag{17}$$

Given $N > N_0$ let $0 \leq t_1^N < t_2^N < \cdots$ be all the Q^{-N} returns to w, i.e.

$$\{t_1^N, t_2^N, \dots\} = \{n \in \mathbb{N} \,|\, d(T^n q, w) \leq Q^{-N}\}.$$

We need the following

3.1. **Proposition.** *For any $\ell \geq 0$,* $\quad t_{\ell+1}^N - t_\ell^N \geq \sqrt{2}^{N-N_0-1}$.

Using Proposition 3.1 we continue the proof of Theorem A.

Write

$$B(w, r) := \{\, x \in X \mid d(x, w) \leq r \,\}.$$

Given $N \gg N_0$, let $f_N : X \to \mathbb{R}$ be a continuous function such that $0 \leq f \leq 1$, $f|_{B(w,Q^{-N-1})} \equiv 1$ and $\operatorname{supp} f \subseteq B(w, Q^{-N})$. Using that q is a generic point for μ and

Proposition 3.1, we have that

$$\begin{aligned}\mu\big(B(w,Q^{-N-1})\big) &\le \int f_N\,d\mu = \lim_{L\to+\infty}\frac{1}{L}\sum_{i=0}^{L-1} f_N(T^i q)\\ &\le \lim_{L\to+\infty}\frac{1}{L}\,\#\Big\{\,0\le i<L\;\Big|\;d(q_i,w)\le Q^{-N}\,\Big\}\\ &\le \lim_{L\to+\infty}\frac{1}{L}\,\#\Big\{\,\ell\;\Big|\;t_\ell^N\le L\,\Big\}\\ &\le \sqrt{2}^{\,-N+N_0+1}. \qquad (18)\end{aligned}$$

Recall that the dynamic ball about w is

$$V(w,L,\varepsilon):=\big\{\,x\in X\;\big|\;d(T^kx,T^kw)<\varepsilon\,,\ \forall k=0,\dots,L\,\big\}.$$

We have that

$$V(w,L,\varepsilon)=S_1\circ\cdots\circ S_L\big(B(T^Lw,\varepsilon)\big),$$

where S_k is the branch of the inverse of T such that $S_k(T^kw)=T^{k-1}w$. Therefore

$$V(w,L,\varepsilon)\subseteq B(w,\lambda^L\varepsilon).$$

Let N be such that

$$Q^{-N-2}\le\lambda^L\varepsilon\le Q^{-N-1}.$$

Then

$$-N\le L\,\frac{\log\lambda}{\log Q}+\frac{\log\varepsilon}{\log Q}+2.$$

Using (18), we have that

$$\mu\big(V(w,L,\varepsilon)\big)\le\mu\big(B(w,\lambda^L\varepsilon)\big)\le\mu\big(B(w,Q^{-N-1})\big)\le\sqrt{2}^{\,-N+N_0+1}.$$

$$\begin{aligned}\frac{1}{L}\,\log\mu\big(V(w,L,\varepsilon)\big)&\le\frac{1}{L}\,\big(\log\sqrt{2}\big)\big(-N+N_0+1\big)\\ &\le\frac{\log\lambda}{\log Q}\,\log\sqrt{2}+\frac{1}{L}\,\big(\log\sqrt{2}\big)\left(2+\frac{\log\varepsilon}{\log Q}+N_0+1\right).\end{aligned}$$

By Brin-Katok Theorem [6] we have that

$$h_\mu(T)=-\lim_{L\to+\infty}\frac{1}{L}\,\log\mu\big(V(w,L,\varepsilon)\big)\ge\frac{\log\lambda^{-1}}{\log Q}\,\log\sqrt{2}>0.$$

This contradicts the choice of F in (14). Therefore such non-empty open set $\mathcal{U}$ in (13) does not exist. This implies that the (open) set $\cup_{y\in\mathrm{Per}(T)}\overset{\circ}{P}_y$ is dense. $\square$

Now we prove

3.1. **Proposition.** *For any* $\ell \geq 0$, $\quad t^N_{\ell+1} - t^N_\ell \geq \sqrt{2}^{N-N_0-1}$.

Proof: For $N \in \mathbb{N}$, let

$$\mathbb{A}_N := \{(x,y) \in X \times X \mid d(x,y) \leq Q^{-N}\}.$$

From (15) if $N > N_0$ and $(x_k)_{k=0}^{p-1}$ is a p-periodic $2Q^{-N}$ pseudo-orbit in $\mathcal{O}(q)$ with at most 2 jumps, then there is a Q^{-N+1}-return $(x_i, x_j) \in \mathbb{A}_{N-1}$ with $0 \leq i < j \leq p-1$.

Write $q_i := T^i(q)$. The sequence $(q_k)_{k=t^N_\ell}^{t^N_{\ell+1}-1}$ is a periodic Q^{-N} pseudo-orbit in $\mathcal{O}(q)$ with 1 jump. Therefore there is a Q^{-N+1}-return $d(q_i, q_j) \leq Q^{-N+1}$ with $t^N_\ell \leq i < j \leq t^N_{\ell+1} - 1$. This gives rise to two Q^{-N+1} periodic pseudo-orbits in $\mathcal{O}(q)$ with at most 2 jumps. Namely, $(q_i, \ldots, q_{j-1})$ and $(q_j, \ldots, q_{t^N_{\ell+1}-1}, q_{t^N_\ell}, \ldots, q_{i-1})$. Each of them give rise to a Q^{-N+2} return.

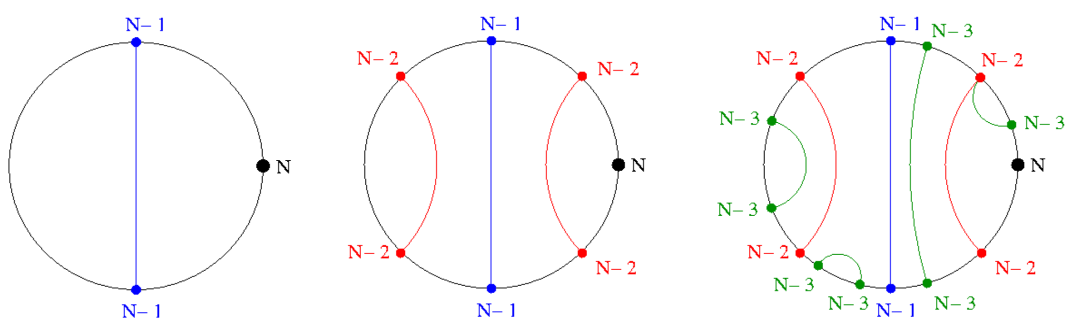

FIGURE 1. Example of a cascade of returns implied by the inductive process.

It is simpler to show the inductive process in a picture. Draw a circle $\mathbb{S}$ with the elements of the pseudo-orbit $(q_k)_{k=t^N_\ell}^{t^N_{\ell+1}-1}$. Inside the disk $\mathbb{D}$, draw a line from q_i to q_j. It may be that $q_i = q_{t^N_\ell}$ but in that case $q_j \neq q_{t^N_{\ell+1}}$. The line $\ell_1 = \overline{q_i q_j}$ separates the disk in two components. Each component is a Q^{-N+1} pseudo-orbit with at most two jumps (one jump of size $\leq Q^{-N+1}$ and possibly another with size $\leq Q^{-N} < Q^{-N+1}$). Thus, each component has at least one Q^{-N+2} return… The interior of the lines in this construction do not intersect.

We will also draw a tree with the returns, in order to see that their number grows exponentially. An example appears in figure 2. The nodes of the tree are the returns implied by Proposition 2.4. The height[1] of the node corresponds to the size of the return. The numbers near a node are the quantity of returns in upper levels of the tree which are adjacent to the return of the node, either at its left or at its right. These numbers are also equal to -1 +the quantity of jumps of the two new periodic pseudo-orbits determined by the node.

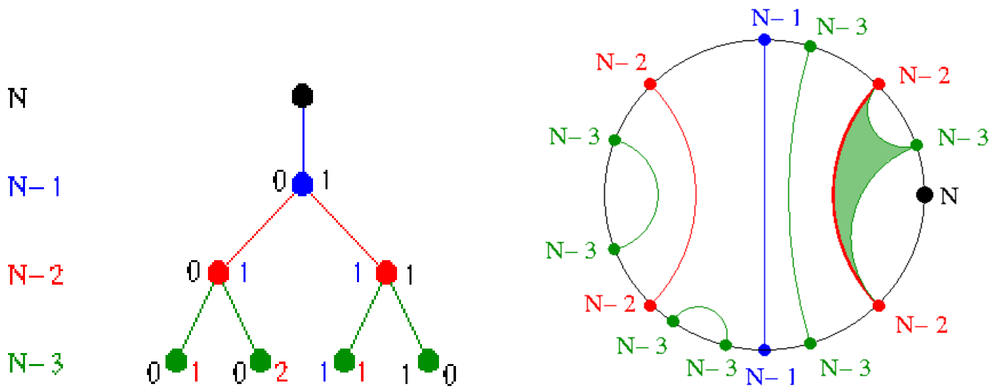

FIGURE 2. An example of a distribution of returns implied by Proposition 2.4 and the tree representing it. The shadow is explained after inequality (19).

We show how the tree is constructed in the example of figure 2. We begin with a return in $\mathbb{A}_N$. This gives a periodic Q^{-N} pseudo-orbit with no other jump. It implies the existence of a return in $\mathbb{A}_{N-1}$. In the tree we draw a vertical line from level N to level $N-1$. At this stage, the line in the circle corresponding to the $\mathbb{A}_{N-1}$ return divides the disk in two components. One side has 1 return in $\mathbb{A}_N$ that appears in a previous level in the tree and the other side has 0 returns appearing above in the tree. We write the numbers 0 and 1 at the sides of the node of the tree corresponding to the $\mathbb{A}_{N-1}$ return. The $\mathbb{A}_{N-1}$ return divides the circle in two components. The component at the left is a periodic Q^{N-1} pseudo-orbit with only one Q^{N-1} jump, corresponding to the number 0 in the tree. The component at the right is a Q^{N-1} pseudo-orbit with a Q^{N-1} jump and also a Q^N jump, and corresponds to the number 1 in the tree in the node at level $N-1$. Proposition 2.4 implies the existence of other returns in $\mathbb{A}_{N-2}$ for both pseudo-orbits. In the right hand side of figure 2 we draw the case in which the pseudo-orbit segment between the $\mathbb{A}_{N-2}$ return contains a Q^{-N} jump. Cutting the Q^{N-1} pseudo-orbit of the right hand side of the circle at the $\mathbb{A}_{N-2}$ return we obtain two Q^{N-2} periodic pseudo-orbits. The one

[1]A node in level $N-2$ corresponds to a return $(q_i, q_j) \in \mathbb{A}_{N-2}$ with $d(q_i, q_j) \le Q^{-N+2}$. The distance could actually be smaller. The height in the tree is just the distance implied by the previous steps in the construction of the tree. It if was smaller it would imply a larger subtree from that point, this helps the argument.

at the right has a Q^N jump which appears previously in the tree and the one at the left has a Q^{N-1} jump appearing previously in the tree. We write the numbers 1 and 1 in the corresponding node of the tree.

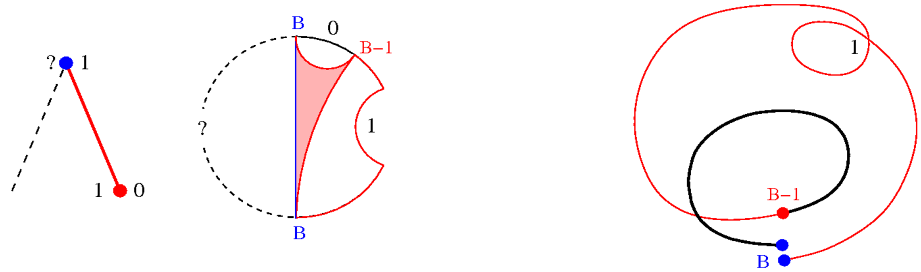

FIGURE 3. If the return implied by Proposition 2.4 contains one of the endpoints of the mother periodic pseudo-orbit we observe that it divides the mother pseudo-orbit in two child pseudo-orbits. We draw lines connecting the ends of these pseudo-orbits and shadow the internal part of the disk $\mathbb{D}$ which does not contain an interval in the circle $\mathbb{S}$.

We want to assure that each node of the tree corresponds to a return which has at least one new point of the pseudo-orbit $(q_{t_\ell^N}, \dots, q_{t_{\ell+1}^N-1})$ which did not appear in the returns corresponding to the other nodes of the tree. So that we have

$$t_{\ell+1}^N - t_\ell^N \geq \#\{\text{nodes}\}. \tag{19}$$

Given a Q^{-B} periodic pseudo-orbit $(q_a, \dots, q_{b-1})$ with at most $M = 2$ jumps we get a return (i,j) with $d(q_i, q_j) \leq Q^{-B+1}$, $a \leq i < j \leq b$. It may happen that $i = a$ or $j = b-1$ but not both, because it would be the same return we started with. Say $i = a$. The next return in the construction of the tree could be (j, b) and in that case the new node in the tree does not correspond to a new element in the pseudo-orbit, the points q_a, q_i, q_b were already counted. To avoid this situation we observe that $(q_j, \dots, q_{b-1})$ is indeed a $2Q^{-B}$ periodic pseudo-orbit because $d(q_j, q_b) \leq d(q_j, q_a) + d(q_a, q_b) = d(q_j, q_i) + d(q_a, q_b)$. In the disk we draw two new lines $\overline{q_a q_j}$ and $\overline{q_j q_b}$ and shadow the region bounded by these two lines and $\overline{q_a q_b}$. We treat the region as a new line. This is, the two new components in the (disk $\mathbb{D}$) $\setminus$ {shadow} are treated as the two sides of a new line in $\mathbb{D}$ corresponding to a node in the tree. The numbers in the node are -1 + the quantity of jumps of each of the pseudo-orbits determined by the components.

We describe in figures 4 – 7 all the possibilities for a vertical step of size one of the tree's construction. The "bitten" parts are returns which appeared previously in the tree's construction. In the case of a node at level B with numbers $(0, 2)$, for $M = 2$ Proposition 2.4 does not give a new return for the side which already has 2 returns. In this case Proposition 2.4 only implies that there is one return in $\mathbb{A}_{B-1}$ in the side which had 0 previous returns. Then for a node at level B with numbers $(0,2)$ we draw only one vertical line to level $B-1$.

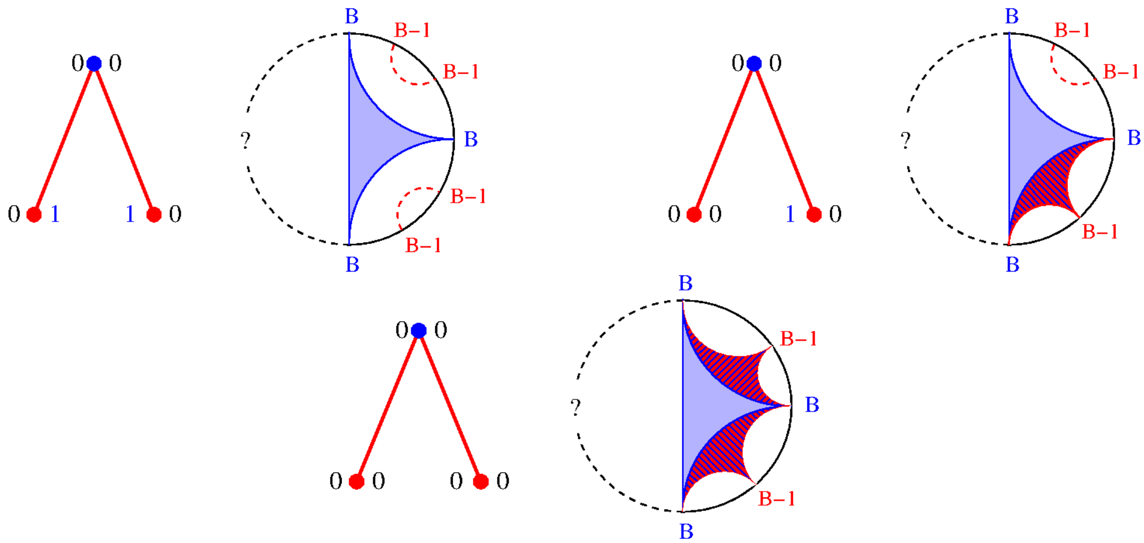

FIGURE 4. These are all the cases for a node in the tree with numbers $(0,0)$. Since the beginning numbers are $(0,1)$, the numbers $(0,0)$ appear only with a shadow i.e. returns to the beginning (and end) of the pseudo-orbit.

The tree is constructed by joining the pieces shown in these figures. Except for the nodes with numbers $(0,2)$ all the other nodes have two child nodes. All the nodes have at least one child. The nodes with numbers $(0,2)$ in figure 7 have one child with numbers different from $(0,2)$. This implies that they have at least two grandchildren nodes. This is graphically shown in figure 8. Therefore the quantity of nodes at least duplicates in every sequence of two vertical steps. The tree continues growing while the height number is larger than N_0. Thus it has at least $N-N_0$ vertical steps and then at least $(N-N_0-1)/2$ sequences of two vertical steps (and hence duplication of nodes). We obtain

$$t^N_{\ell+1}-t^N_\ell \geq \#\{\text{nodes}\} \geq 2^{\frac{N-N_0-1}{2}}.$$

$\square$

References

[1] A. T. Baraviera, R Leplaideur, and A. O. Lopes, *Ergodic optimization, zero temperature limits and the max-plus algebra*, IMPA, Rio de Janeiro, 2013, 29^0 Coloquio Brasileiro de Matematica.

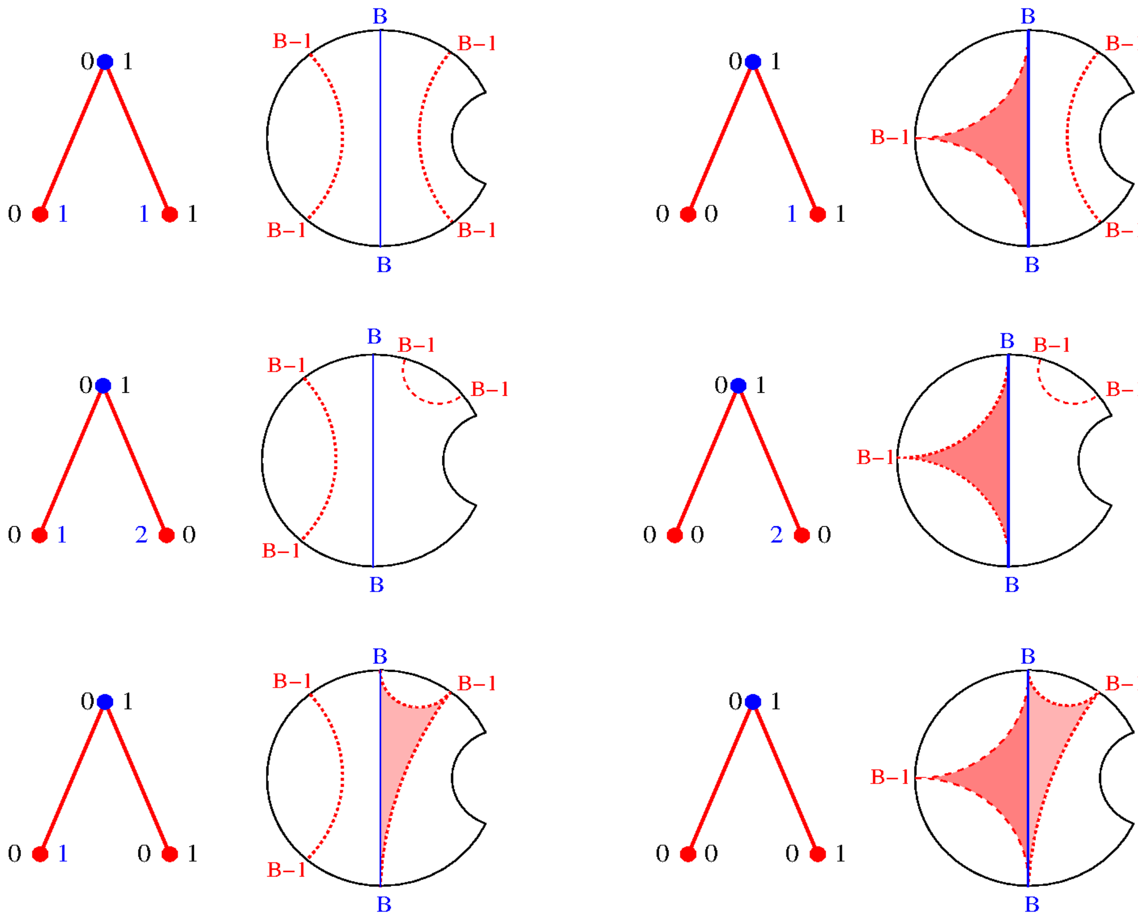

FIGURE 5. All the cases for a node in the tree with numbers $(0, 1)$.

[2] Thierry Bousch, *Le poisson n'a pas d'arêtes*, Ann. Inst. H. Poincaré Probab. Statist. **36** (2000), no. 4, 489–508.

[3] ———, *La condition de Walters*, Ann. Sci. École Norm. Sup. (4) **34** (2001), no. 2, 287–311.

[4] Julien Brémont, *Gibbs measures at temperature zero*, Nonlinearity **16** (2003), no. 2, 419–426.

[5] Xavier Bressaud and Anthony Quas, *Rate of approximation of minimizing measures*, Nonlinearity **20** (2007), no. 4, 845–853.

[6] Michael Brin and Anatole Katok, *On local entropy.*, Geometric dynamics, Proc. int. Symp., Rio de Janeiro/Brasil 1981, Lect. Notes Math. 1007, 1983, pp. 30–38.

[7] J.-R. Chazottes, J.-M. Gambaudo, and E. Ugalde, *Zero-temperature limit of one-dimensional Gibbs states via renormalization: the case of locally constant potentials*, Ergodic Theory Dynam. Systems **31** (2011), no. 4, 1109–1161.

[8] G. Contreras, A. O. Lopes, and Ph. Thieullen, *Lyapunov minimizing measures for expanding maps of the circle*, Ergodic Theory Dynam. Systems **21** (2001), no. 5, 1379–1409.

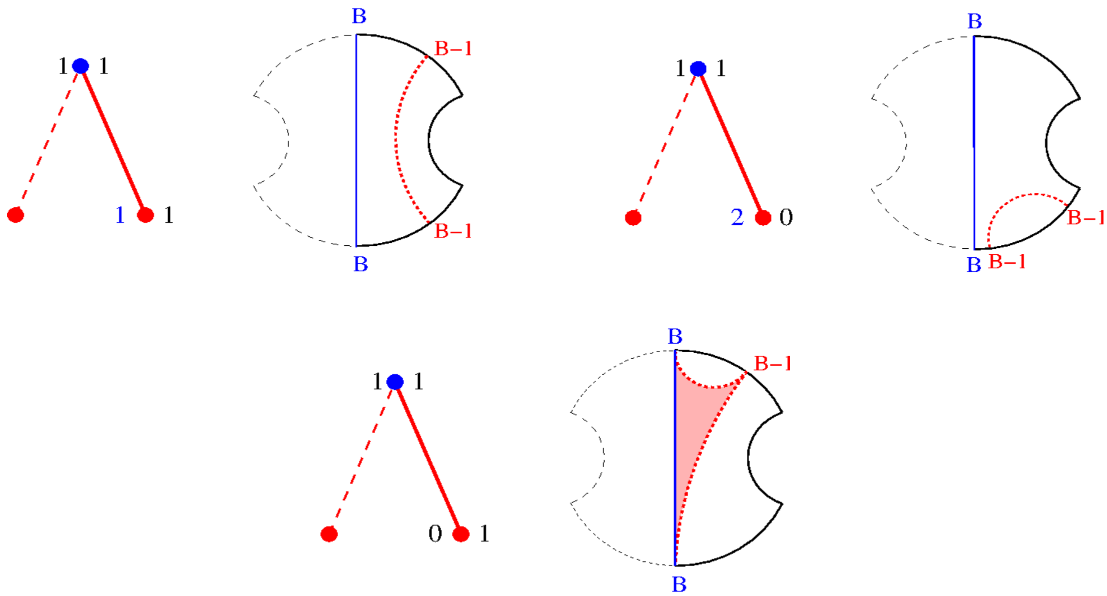

Figure 6. The case of a node with numbers $(1,1)$ is symmetric with respect to a reflection with axis $\overline{BB}$ = the line joining the returning points. Here we show the three possibilities at the right hand side of the axis $\overline{BB}$. There are 9 possibilities for a node with numbers $(1,1)$ given by all the combinations of the three cases shown in the figure at the right hand side and the three symmetric cases at the left hand side. All the possibilities have two child nodes.

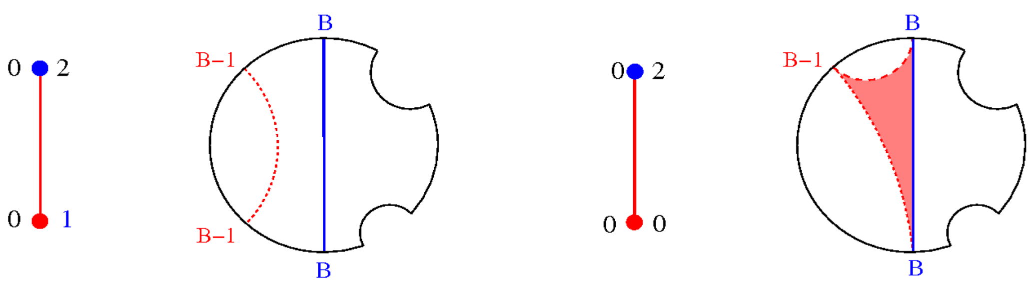

Figure 7. In a node with numbers $(0,2)$ one side is a pseudo-orbit with $M+1=3$ nodes. Since we have chosen $M=2$, Proposition 2.4 does not apply and we leave that side without a child node. The other side has a child node with numbers $(0,0)$ or $(0,1)$. Both cases have been analyzed before: the child node has in turn two children nodes.

[9] Albert Fathi, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Sér. I Math. **324** (1997), no. 9, 1043–1046.

[10] E. Garibaldi, A. O. Lopes, and Ph. Thieullen, *On calibrated and separating sub-actions*, Bull. Braz. Math. Soc. (N.S.) **40** (2009), no. 4, 577–602.

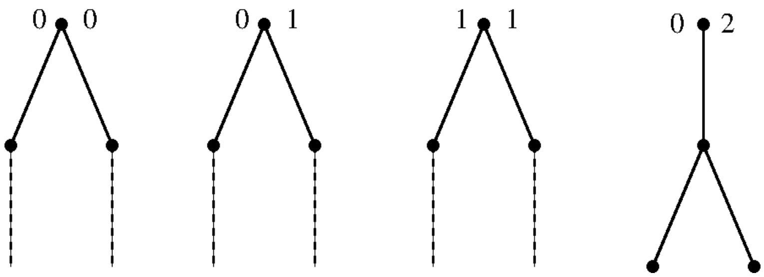

Figure 8. As seen in the previous figures, every node has at least one child node. Nodes with numbers $(0,0)$, $(0,1)$ and $(1,1)$ have two children and thus in two generation they have at least two grandchildren nodes. As seen in figure 7, a node with numbers $(0,2)$ has two grandchildren. Therefore every node has at least two grandchildren nodes in two generations.

[11] Kazimierz Goebel and W. A. Kirk, *Topics in metric fixed point theory*, Cambridge University Press, Berlin, 1990.

[12] Oliver Jenkinson, *Ergodic optimization*, Discrete Contin. Dyn. Syst. **15** (2006), no. 1, 197–224.

[13] Renaud Leplaideur, *A dynamical proof for the convergence of Gibbs measures at temperature zero*, Nonlinearity **18** (2005), no. 6, 2847–2880.

[14] Ricardo Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), no. 2, 273–310.

[15] Ian D Morris, *Maximizing measures of generic hölder functions have zero entropy*, Nonlinearity **21** (2008), 993–1000.

[16] Anthony Quas and Jason Siefken, *Ergodic optimization of super-continuous functions on shift spaces*, Ergodic Theory and Dynamical Systems **32** (2012), no. 6, 2071–2082.

[17] David Ruelle, *Thermodynamic formalism. The mathematical structures of equilibrium statistical mechanics. 2nd edition.*, Cambridge Mathematical Library. Cambridge: Cambridge University Press, 2004.

[18] Peter Walters, *An introduction to ergodic theory*, Graduate Texts in Math. 79, Springer, 1982.

[19] G.C. Yuan and B.R. Hunt, *Optimal orbits of hyperbolic systems*, Nonlinearity **12** (1999), no. 4, 1207–1224.

Appendix A. Zero Entropy.

In this appendix we describe the modifications to the proof of Theorem 2 in [15] to obtain Theorem 1.2.

We need two lemmas.

A.1. **Lemma.** *Let $a_1, \dots, a_n$ be non-negative real numbers, and let $A = \sum_{i=1}^n a_i \geq 0$. Then*

$$\sum_{i=1}^{n} -a_i \, \log a_i \leq 1 + A \, \log n,$$

where we use the convention $0 \, \log 0 = 0$.

Proof: Applying Jensen's inequality to the concave function $x \mapsto -x \log x$ yields

$$\frac{1}{n} \sum_{i=1}^{n} -a_i \, \log a_i \leq - \left(\frac{1}{n} \sum_{i=1}^{n} a_i \right) \log \left(\frac{1}{n} \sum_{i=1}^{n} a_i \right) = -\frac{A}{n} \, \log A + \frac{A}{n} \, \log n$$

from which the result follows. □

A.2. **Lemma.** *Let $f \in \mathrm{Lip}(X, \mathbb{R})$ and suppose that $\mathcal{M}_{\max}(f) = \{\mu\}$ for some $\mu \in \mathcal{M}(T)$. Then there is $C > 0$ such that for every $\nu \in \mathcal{M}(T)$,*

$$-\alpha(f) - C \int d(x, K) \, d\nu \leq \int f \, d\nu,$$

where $K = \operatorname{supp} \mu$.

Proof: By Proposition 2.1 there exists $g \in \mathrm{Lip}(X, \mathbb{R})$ such that $f + g - g \circ T \leq -\alpha(f)$. Define $\tilde{f} = f + g - g \circ T$. Since $\tilde{f}$ is continuous, $\int \tilde{f} \, d\mu = \int f \, d\mu = -\alpha(f)$ and $\tilde{f} \leq -\alpha(f)$, it follows that $\tilde{f}(x) = -\alpha(f)$ for every $x \in K = \operatorname{supp} \mu$. Let $C = \mathrm{Lip}(\tilde{f})$. Given $x \in X$, let $z \in K$ be such that $d(x, z) = d(x, K)$; we then have

$$\tilde{f}(x) \geq \tilde{f}(z) - C \, d(x, z) = -\alpha(f) - C \, d(x, K)$$

from which the result follows. □

Proof of Theorem 1.2:

For $p \geq 1$ let $\mathcal{M}^p(T)$ be the set of invariant probabilities supported on periodic of period smaller or equal to p. In this appendix we will identify a periodic orbit $\{z, Tz, \dots, T^{p-1}z\}$ with the corresponding invariant measure $\mu = \frac{1}{p} \sum_{i=0}^{p-1} \delta_{T^i z}$.

Let

$$\varepsilon_0 > 0, \qquad 0 < \lambda < 1 \tag{20}$$

be such that for every $x \in X$ the branches of the inverses of T at x are well defined, injective, and are λ-contractions on the ball $B(x, \varepsilon_0)$ of radius ε_0 centered at x.

Let

$$\mathcal{E}_\gamma := \{\, f \in \mathrm{Lip}(X, \mathbb{R}) \mid h(\mu) < 2\,\gamma\, h_{\mathrm{top}}(T) \quad \forall \mu \in \mathcal{M}_{\max}(f)\,\}.$$

It is enough to prove that $\mathcal{E}_\gamma$ is open and dense for every $\gamma > 0$.

Step 1 and Step 2 of the proof are the same as in [15]. The Claim of step 3 is

Step 3.

Claim: Given any $0 < \theta < 1$, *there is a sequence of integers* $(m_n)_n$ *and a sequence of periodic orbits* $\mu_n \in \mathcal{M}^n(T)$ *such that*

$$\int d(x, K)\, d\mu_n = o(\theta^{m_n}) \qquad \text{and} \qquad \lim_{n\to\infty} \frac{\log n}{m_n} = 0.$$

which has the same proof as in [15].

Step 4. For each $n \geq 1$ define $L_n := \operatorname{supp} \mu_n$. Using (20), fix

$$0 < \theta < \min\{\varepsilon_0, \lambda\}. \tag{21}$$

Claim: There is $N_\gamma > 0$ *such that when* $n \geq N_\gamma$

$$\nu(\{\, x \in X \mid d(x, L_n) \geq \theta^{m_n}\,\}) > \gamma$$

for every invariant measure $\nu \in \mathcal{M}(T)$ *such that* $h(\nu) \geq 2\gamma\, h_{\mathrm{top}}(T)$.

Proof of the Claim.

Recall that a Markov partition for T is a finite collection of sets S_i which cover X such that

(a) $S_i = \overline{\mathrm{int}\, S_i}$.
(b) If $i \neq j$ then $\mathrm{int}\, S_i \cap \mathrm{int}\, S_j = \emptyset$.
(c) $f(S_i)$ is a union of sets S_j.

Ruelle [17, §7.29] proves that for expanding maps there are Markov partitions of arbitrarily small diameter. Let $\mathbb{P}$ be a Markov partition with $\operatorname{diam} \mathbb{P} < \varepsilon_0$. The elements of the partition

$$\mathbb{P}^{(n)} := \bigvee_{i=0}^{n-1} T^{-i}\mathbb{P} = \Big\{ \bigcap_{i=0}^{n-1} A_i \;\Big|\; A_i \in T^{-i}\mathbb{P} \Big\}$$

have diameter smaller than $\lambda^{n-1}\varepsilon_0$ and contain an open set. Then the partition $\mathbb{P}$ is generating because the σ-algebra

$$\mathbb{P}^\infty = \sigma\big(\cup_n \mathbb{P}^{(n)}\big) = \mathcal{B}orel(X).$$

contains all the open sets. Therefore [18, Theorem 4.18] for every invariant measure $\nu \in \mathcal{M}(T)$,

$$h(\nu) = \inf_k \frac{1}{k} \sum_{A\in\mathbb{P}^{(k)}} -\nu(A)\,\log\nu(A).$$

From the definition of topological entropy using covers [18, §7.1] we have that

$$\lim_{k\ge 1} \frac{1}{k}\,\log\#\,\mathbb{P}^{(k)} \le h_{\text{top}}(T).$$

Choose N_γ large enough that for all $n \ge N_\gamma$

$$\frac{2+\log\#\mathbb{P}}{m_n} + \frac{\log n}{m_n} + \frac{\gamma}{m_n}\log\#\mathbb{P}^{(m_n)} < 2\gamma\, h_{\text{top}}(T). \tag{22}$$

Let $\nu \in \mathcal{M}(T)$ and suppose that

$$\nu(\{x\in X \mid d(x,L_n) \ge \theta^{m_n}\}) \le \gamma \tag{23}$$

for some $n \ge N_\gamma$. We will show that necessarily $h(\nu) < 2\gamma\, h_{\text{top}}(\nu)$.

Let

$$W_n := \{\, A \in \mathbb{P}^{(m_n)} \mid d(x,L_n) < \theta^{m_n} \quad \text{for some } x\in A \,\}.$$

From (23),

$$\tilde{\gamma}_n := \sum_{A\in\mathbb{P}^{(m_n)}\setminus W_n} \nu(A) \le \gamma.$$

Using lemma A.1 we have that

$$\begin{aligned} h(\nu) &\le \frac{1}{m_n}\sum_{A\in W_n} -\nu(A)\,\log\nu(A) + \frac{1}{m_n}\sum_{A\in\mathbb{P}^{(m_n)}\setminus W_n} -\nu(A)\,\log\nu(A) \\ &\le \frac{1}{m_n}\big(1+(1-\tilde{\gamma}_n)\log\#W_n\big) + \frac{1}{m_n}\big(1+\gamma\,\log\#\mathbb{P}^{(m_n)}\big). \end{aligned} \tag{24}$$

Unsig (21), observe that since $\frac{\theta^{m_n}}{\lambda^{m_n}} < \varepsilon_0$ for any $y \in L_n$ there is a branch g of the inverse of T^{m_n} such that the ball

$$B(y, \theta^{m_n}) \subseteq g\big(B(T^{m_n}y, \varepsilon_0)\big).$$

Since $\mathbb{P}$ is a Markov partition with $\operatorname{diam}\mathbb{P} < \varepsilon_0$,

$$\mathbb{P}^{m_n} = \{\, g(A) \mid A \in \mathbb{P}, \quad g \text{ is branch of } T^{-m_n}\,\}.$$

Therefore the ball $B(y, \theta^{m_n})$ intersects at most $\#\mathbb{P}$ elements of $\mathbb{P}^{m_n}$ because by applying T^{m_n}

$$\#\{\, B \in \mathbb{P}^{m_n} \mid B \cap B(y, \theta^{m_n}) \neq \emptyset\} \leq \#\{\, A \in \mathbb{P} \mid A \cap B(T^{m_n}y, \varepsilon_0) \neq \emptyset\,\} \leq \#\mathbb{P}.$$

Since L_n has at most n elements, $\#W_n \leq n\,\#P$. Thus from (22) and (24) we have that

$$\begin{aligned} h(\nu) &\leq \frac{1}{m_n}\big(1 + (1 - \tilde{\gamma}_n)\log n\,\#\mathbb{P}\big) + \frac{1}{m_n}\big(1 + \gamma\, \log \#\mathbb{P}^{(m_n)}\big). \\ &\leq \frac{2 + \log\#\mathbb{P}}{m_n} + \frac{\log n}{m_n} + \frac{\gamma}{m_n}\log\#\mathbb{P}^{(m_n)} < 2\gamma\, h_{\text{top}}(T). \end{aligned}$$

This proves the claim.

Step 5 is the same as in [15]. This ends the proof.

□

CIMAT, A.P. 402, 36.000, Guanajuato. GTO, México.

E-mail address: gonzalo@cimat.mx