GENERICITY OF PERIODIC MAXIMIZATION: PROOF OF CONTRERAS' THEOREM FOLLOWING HUANG, LIAN, MA, XU, AND ZHANG

JAIRO BOCHI

Abstract. A theorem due to Gonzalo Contreras essentially says that given a expanding map $T\colon X \to X$ and a *generic* Lispchitz (or Hölder) function $f\colon X \to \mathbb{R}$, there is a unique T-invariant probability measures that maximizes the average of f, and this measure is supported on a periodic orbit. In these lecture notes, we present a proof of Contreras' theorem following the recent preprint of Wen Huang, Zeng Lian, Xiao Ma, Leiye Xu, and Yiwei Zhang.

1. Introduction

Let (X, d) be a compact metric space. Given a continuous self-map $T\colon X \to X$, let $\mathcal{M}_T$ denote the (nonempty) set of T-invariant measures. The *ergodic maximum* of a continuous function $f\colon X \to \mathbb{R}$ is defined as:

$$\beta(f) := \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu \,. \tag{1}$$

Any $\mu \in \mathcal{M}_T$ that attains the supremum is called a *maximizing measure* for f w.r.t. T. The existence of a maximizing measure follows from compactness of $\mathcal{M}_T$ with respect to the weak-* topology. The description of maximizing measures (for reasonable classes of dynamics T and functions f) is the main general problem of a field called *ergodic optimization*. For motivation and much more information, see [Je1, BLL, Ga, Bo, Je2].

Let us assume henceforth that $T\colon X \to X$ is a expanding map (see Section 2 for the precise hypotheses). In particular, there is a dense subset of $\mathcal{M}_T$ formed by probability measures supported on periodic orbits: see e.g. [VO, Theorem 11.3.4].

Fix $\theta > 0$. Let $C^\theta(X)$ be the set of θ-Hölder functions.[1] We use the following norm:

$$\|f\|_\theta := \sup_{x \in X} |f(x)| + \sup_{x \neq y \in X} \frac{|f(x) - f(y)|}{[d(x,y)]^\theta} \,, \tag{2}$$

which makes $C^\theta(X)$ a Banach space.[2]

Let $\mathcal{L}$ be the set of $f \in C^\theta(X)$ with the following properties:

(a) f has a unique maximizing measure μ;
(b) μ is supported on a periodic orbit;
(c) for every $g \in C^\theta(X)$ sufficiently close to f, the measure μ is the unique maximizing measure of g.

Date: May 30, 2019.

[1] We do not assume $\theta \leqslant 1$. If X is a Cantor set then it may be useful to work with $\theta > 1$.

[2] This space is nonseparable, unless X is countable, since the subset $\{[d(x,\cdot)]^\theta \ ; \ x \in X\}$ is discrete.

If $f \in \mathcal{L}$ then we say that f has the *locking property.* (Actually, condition Item b follows from the other two, as shown by Yuan and Hunt [YH] and also as a consequence of Theorem 1.1.)

The following is a main result of ergodic optimization:

Theorem 1.1 (Contreras [Co])**.** *The set $\mathcal{L}$ is (open and) dense in $C^\theta(X)$.*

In particular, maximizing measures of generic Hölder functions are supported on periodic orbits. This provides a positive answer to the topological version of a conjecture of Hunt and Ott [HO]. In its original probabilistic formulation, the conjecture is still open. (See [BZ2] for a partial result). We refer the reader to [Je2, § 7] for more information on the history of the problem.

The aim of this note is to present a proof of Theorem 1.1. This proof is due to Huang, Lian, Ma, Xu, and Zhang [H$^+$]. These authors actually prove a more general result. Therefore the proof presented here is slightly simplified.

Before entering into details, let us mention that a major ingredient of Contreras' proof was an intermediate result of Morris [Mo1] stating that maximizing measures of generic Hölder functions have zero entropy. Morris' theorem, on the other hand, was based on a closing lemma due to Bressaud and Quas [BQ]. The proof of Huang et al. is more direct, using Bressaud–Quas but not Morris.

In Sections 2 and 3 we present preliminary results which are then used in Section 4 to prove Theorem 1.1.

2. Closing Lemmas

Our hypotheses on the dynamics $T\colon X \to X$ are:

(a) T is Lipschitz;
(b) T is distance-expanding: this means that there exist constants $r > 0$ and $\lambda > 1$ such that:
$$x, y \in X,\ d(x,y) \leqslant 2r \quad \Rightarrow \quad d(Tx,Ty) \geqslant \lambda d(x,y)\,; \tag{3}$$
(c) T is an open map.

The following result (related to the Shadowing Lemma in hyperbolic dynamics) says that a piece of orbit that almost closes can be well-approximated by a periodic orbit:

Theorem 2.1 (Basic closing property)**.** *There exist constants $L > 1$ and $s > 0$ such that whenever $x \in X$ and $n \geqslant 1$ satisfy $d(T^n x, x) \leqslant s$, it is possible to find a point $y \in X$ with $T^n y = y$ such that:*
$$\forall j \in \{0,1,\dots,n\}, \qquad d(T^j y, T^j x) \leqslant L\lambda^{-(n-j)}\, d(T^n x, x)\,, \tag{4}$$
where $\lambda > 1$ is as in (3).

Proof. The result follows from [PU, Corollary 4.2.5] and its proof. □

In particular, if $K \subseteq X$ is a nonempty compact forward-invariant set (in the sense that $T(K) \subseteq K$), then every neighborhood of K contains a periodic orbit (since K must contain recurrent points). The following theorem relates the period if this orbit with the size of the neighborhood:

Theorem 2.2 (Bressaud–Quas [BQ])**.** *Let $K \subseteq X$ be a nonempty compact forward-invariant set. Then, for every $\kappa > 0$ and every sufficiently small $\varepsilon > 0$, we can find a periodic orbit of period $p < (1/\varepsilon)^\kappa$ entirely contained in the ε-neighborhood of K.*

A variation of the proof that avoids the use of Markov partitions is presented in [BG, Appendix A.6]; that proof assumes T is a hyperbolic homeomorphism, but it can be adapted to expanding maps with very minor modifications.

If $\mathcal{O}$ is a periodic orbit, then its *half-gap* is defined as:

$$\Delta(\mathcal{O}) := \min\left\{ r,\ \frac{1}{2} \min_{\substack{x,y\in\mathcal{O} \\ x\neq y}} d(x,y) \right\}, \tag{5}$$

where r is the constant from (3). It is understood that $\min\varnothing = +\infty$, so:

$$\Delta(\mathcal{O}) = r \quad \text{if } \mathcal{O} \text{ consists on a single point.} \tag{6}$$

The following lemma is a cornerstone of the proof. The original formulation [H$^+$, Proposition 2.1] is stronger.

Lemma 2.3 (Huang, Lian, Ma, Xu, Zhang [H$^+$])**.** *Let $K \subseteq X$ be a nonempty compact forward-invariant set. Then for every $\theta > 0$ and $\tau > 0$ there exists a periodic orbit $\mathcal{O}$ such that:*

$$\sum_{x\in\mathcal{O}} [d(x,K)]^\theta \leqslant \tau[\Delta(\mathcal{O})]^\theta\,. \tag{7}$$

Proof. The proof will consist on the recursive construction of a chain periodic orbits $\mathcal{O}_0$, ..., $\mathcal{O}_m$, the last of which will have the desired property (7). We will use the notations:

$$\Sigma_k := \sum_{x\in\mathcal{O}} [d(x,K)]^\theta\,, \qquad \Delta_k := \Delta(\mathcal{O}_k)\,. \tag{8}$$

The idea of the construction is as follows. We start applying Theorem 2.2 in order to find an orbit $\mathcal{O}_0$ very close to the set K and whose period p_0 is not very large. If the half-gap $\Delta_0 = \Delta(\mathcal{O}_0)$ is not exceedingly small then we are done. Otherwise we look at a pair of distinct points in the orbit that are as close as possible. This pair of points divides the orbit $\mathcal{O}_0$ in two pieces. We choose the piece with smallest length, and then close it using Theorem 2.1. So we obtain a new periodic orbit of period $p_1 \leqslant p_0/2$ which is still close to K; furthermore Σ_1 is comparable to Σ_0. If the new orbit has a sufficiently large half-gap then we are done. Otherwise we repeat the procedure. This process necessarily ends quickly: after at most $\lfloor \log_2(p_0) \rfloor$ steps we would get a periodic orbit consisting of a single point, which by definition (6) has a large half-gap. We proceed with the precise proof.

Fix the set K and the positive numbers θ and τ. Let $0 < c < 1/2$ be a small constant, whose precise value will be exhibited lated. Let $0 < \varepsilon < 1$ be another small number such that:

$$\tau^{-1/\theta}\varepsilon^{1/2} \leqslant \min\left\{ r, \frac{s}{2} \right\}, \tag{9}$$

where the constants r and s come from property (3) and Theorem 2.1, respectively. We apply Bressaud–Quas Theorem 2.2 with $\kappa := c\theta$: so, reducing ε if necessary, we can find a periodic orbit $\mathcal{O}_0$ of period

$$p_0 < \varepsilon^{-c\theta} \tag{10}$$

contained in the ε-neighborhood of K. Consider the quantities Σ_0 and Δ_0 defined by (8); if they satisfy $\Sigma_0 \leqslant \tau\Delta_0^\theta$ then we are done: $\mathcal{O}_0$ is the desired periodic orbit. Therefore we may assume that:

$$\Sigma_0 > \tau\Delta_0^\theta\,. \tag{11}$$

On the other hand, we also have:

$$\begin{aligned}\Sigma_0 &\leqslant p_0\varepsilon^{\theta} && \text{(since } \mathcal{O}_0 \text{ is } \varepsilon\text{-close to } K) && (12)\\ &< \varepsilon^{-c\theta+\theta} && \text{(by (10))} && (13)\\ &\leqslant \varepsilon^{\theta/2} && \text{(since } c < 1/2 \text{ and } \varepsilon < 1). && (14)\end{aligned}$$

Let us describe the recursive step. Assume that we are given a periodic orbit $\mathcal{O}_{k-1}$ such that:

$$\tau\Delta_{k-1}^{\theta} < \Sigma_{k-1} < \varepsilon^{\theta/2}\,, \tag{15}$$

(These conditions hold when $k = 0$, by (11) and (14).) In particular, recalling (9),

$$\Delta_{k-1} < \tau^{-1/\theta}\varepsilon^{1/2} \leqslant \min\left\{r, \frac{s}{2}\right\}. \tag{16}$$

It follows that the period p_{k-1} of the orbit $\mathcal{O}_{k-1}$ is at least 2, since otherwise we would have $\Delta_{k-1} = r$ by (6), Let x, x' be distinct points in $\mathcal{O}_{k-1}$ with $d(x, x')$ as small as possible, that is, $d(x, x') = 2\Delta_{k-1}$. Interchanging x and x' if necessary, we have $x' = T^n x$ for some n in the range $1 \leqslant n \leqslant p_{k-1}/2$. Note that $d(x, x') = 2\Delta_{k-1}$ is less than s, by (16). This permits us to apply Theorem 2.1, and therefore find a point y such that $T^n y = y$ and (4) holds. Let $\mathcal{O}_k$ be the orbit of y; its period p_k is at most n, and so

$$p_k \leqslant \frac{p_{k-1}}{2}\,. \tag{17}$$

We estimate:

$$\begin{aligned}\Sigma_k &\leqslant \sum_{j=0}^{n-1} d(T^j y, K)^{\theta} && (18)\\ &\leqslant \sum_{j=0}^{n-1} \big[d(T^j x, K) + d(T^j x, T^j y)\big]^{\theta} && (19)\\ &\leqslant \sum_{j=0}^{n-1} 2\big[d(T^j x, K)^{\theta} + d(T^j x, T^j y)^{\theta}\big] && (20)\\ &\leqslant 2\Sigma_{k-1} + 2\sum_{j=0}^{n-1} \big[L\lambda^{-(n-j)}\, d(x, x')\big]^{\theta} && (21)\\ &\leqslant 2\Sigma_{k-1} + \frac{4L^{\theta}}{1-\lambda^{-\theta}}\,\Delta_{k-1}^{\theta} && (22)\\ &\leqslant \underbrace{\left(2 + \frac{4L^{\theta}}{1-\lambda^{-\theta}}\cdot\frac{1}{\tau}\right)}_{C}\Sigma_{k-1}\,. && (23)\end{aligned}$$

By iterating the conditions (17) and (23), we obtain the bounds:

$$p_k \leqslant \frac{p_0}{2^k} \quad \text{and} \quad \Sigma_k \leqslant C^k\Sigma_0\,. \tag{24}$$

Since $p_k \geqslant 1$, the first inequality yields $k \leqslant \log_2(p_0)$; substituting in the second inequality we obtain:

$$\begin{aligned}\Sigma_k &\leqslant p_0^{\log_2(C)}\Sigma_0 && (25)\\ &\leqslant p_0^{\log_2(C)+1}\varepsilon^{\theta} && \text{(by (12))} && (26)\end{aligned}$$

$$< \varepsilon^{[-c(\log_2(C)+1)+1]\theta} \quad \text{(by (10))}. \tag{27}$$

So if the constant c is initially defined as:

$$c := \frac{1}{2(\log_2(C)+1)} \tag{28}$$

then estimate (27) becomes $\Sigma_k < \varepsilon^{\theta/2}$. So the second inequality in the inductive assumption (15) is satisfied when we pass from $\mathcal{O}_{k-1}$ to $\mathcal{O}_k$. Of course, if $\Sigma_k \leqslant \tau\Delta_k^\theta$ then the construction terminates and $\mathcal{O}_k$ is the periodic orbit sought for. Otherwise we keep iterating the procedure. As seen above, this can go on for at most $\lfloor \log_2(p_0) \rfloor$ steps, so we ultimately find the desired period orbit. □

3. Tools from ergodic optimization

We will use the following notation for the Birkhoff sums of a function $f \colon X \to \mathbb{R}$:

$$f^{(n)} := \sum_{j=0}^{n-1} f \circ T^n . \tag{29}$$

The ergodic maximum (1) of a continuous function f has several alternative characterizations, among which:

$$\beta(f) = \sup_{x \in X} \inf_{n \geqslant 1} \frac{f^{(n)}(x)}{n} , \tag{30}$$

This formula was obtained in a more general (subadditive) setting by Morris (see [Mo2, Theorem A.3]); it is also a straightforward consequence of Peres Lemma [Pe].

The following fact will be convenient[3]:

Lemma 3.1 ([YH, Remark 4.5])**.** *Suppose that $f \in C^\theta(X)$ admits a maximizing measure μ supported on a periodic orbit. Then $f \in \overline{\mathcal{L}}$, i.e., f can be approximated in θ-Hölder topology by functions with the locking property.*

See [BZ1] for a proof (assuming $\theta = 1$).

The following theorem is a fundamental fact in ergodic optimization, which was discovered independently by several authors [CG, Sa, CLT]. See e.g. [Je2, § 6] for more information and a proof.

Theorem 3.2 (Mañé Lemma)**.** *For any $f \in C^\theta(X)$, there exists $h \in C^\theta(X)$ such that:*

$$f + h \circ T - h \leqslant \beta(f) . \tag{31}$$

A function of the form $h \circ T - h$ is called a *coboundary*; it is θ-Hölder if h is (since T is Lipschitz). Note that adding a coboundary to f does not affect its ergodic maximum $\beta(f)$. So when proving general properties of maximizing measures of θ-Hölder functions f, Theorem 3.2 allows us to reduce to the situation where $f \leqslant \beta(f)$. In this case, the following proposition provides a simple characterization of maximizing measures:

---

[3]But we could manage without it: see Footnote 4.

Proposition 3.3. *If $f \leqslant \beta(f)$ then*

$$K := \left\{x \in X \,;\, \forall n \geqslant 0,\ f(T^n x) = \beta(f)\right\} \tag{32}$$

is nonempty compact forward-invariant set, and a measure $\mu \in \mathcal{M}_T$ is maximizing for f if and only if its support is contained in K.

Proof. The set K is clearly compact and forward-invariant. If μ is any maximizing measure then $\int f \, \mathrm{d}\mu = \beta(f)$, and since we are assuming $f \leqslant \beta(f)$ we conclude that $f = \beta(f)$ μ-almost everywhere. The support of μ is forward-invariant, and therefore contained in K. Since maximizing measures always exist, we conclude that K is nonempty. Furthermore, any invariant measure whose support is contained in K is obviously maximizing. □

4. Proof of Theorem 1.1

In view of Lemma 3.1, it is sufficient to prove that any given $f \in C^\theta(X)$ can be perturbed in θ-Hölder topology to a function g that has a maximizing measure supported on a periodic orbit.[4] So we fix the function $f \in C^\theta(X)$ and a small positive number ε (less than 1, say).

Adding a constant to f if necessary, we may assume that $\beta(f) = 0$. By Theorem 3.2, adding a coboundary to f if necessary, we may also assume that $f \leqslant 0$. Let K be the set of points $x \in X$ whose entire forward orbits lie in the level set $f^{-1}(0)$. We can assume that K contains no periodic orbits, otherwise there is nothing to prove. By Proposition 3.3, K is a nonempty compact forward-invariant set. Take a positive $\tau \ll \varepsilon$ (to be defined precisely later). Applying Lemma 2.3 to K, we find a periodic orbit $\mathcal{O}$ of period p and half-gap $\Delta := \Delta(\mathcal{O})$ such that:

$$\sum_{x \in \mathcal{O}} \left[d(x, K)\right]^\theta < \tau \, \Delta^\theta \, . \tag{33}$$

Define a function $g \in C^\theta(X)$ by:

$$g := f - \varepsilon \left[d(\cdot, \mathcal{O})\right]^\theta \, . \tag{34}$$

Then $\|g - f\|_\theta = O(\varepsilon)$, so g is a perturbation of f. We will check that the unique probability measure $\mu_{\mathcal{O}}$ supported on $\mathcal{O}$ is g-maximizing, and this will conclude the proof of the theorem. That is, we need to prove that $\beta(g) = \gamma$, where:

$$\gamma := \int g \, \mathrm{d}\mu_{\mathcal{O}} = \frac{1}{p} \sum_{x \in \mathcal{O}} g(x) = \frac{1}{p} \sum_{x \in \mathcal{O}} f(x). \tag{35}$$

Note that $\gamma < 0$, since K contains no periodic orbits.

The very basic idea of the proof is as follows: The function g has no reason to be $\leqslant \gamma$ (in which case the conclusion would be trivial), but since $|\gamma|$ is small, we can guarantee that $g < \gamma$ outside some small neighborhood V of $\mathcal{O}$. Every orbit different from $\mathcal{O}$ that enters that dangerous region V will be repelled by the periodic orbit $\mathcal{O}$ and therefore will not only leave V but be sent sufficiently far away from V so that g becomes reasonably negative: indeed, the relatively large gap of $\mathcal{O}$ ensures that the orbit we are following cannot “accidentally” get close to $\mathcal{O}$ before it gets sufficiently away. So the Birkhoff average of the orbit is brought down. It is far

[4] Alternatively, one can show directly that the perturbation constructed in the following proof already satisfies the locking property: this is how it is done in [H+].

from obvious that this strategy actually works, but once the estimates are done carefully, it will turn out that (33) is exactly the necessary condition.

Let us begin the actual proof. We estimate:

$$p|\gamma| = \sum_{x\in\mathcal{O}} |f(x)| \qquad \text{(since } \gamma < 0,\ f \leqslant 0\text{)} \tag{36}$$

$$\leqslant \sum_{x\in\mathcal{O}} \|f\|_\theta \big[d(x,K)\big]^\theta \quad \text{(since } f = 0 \text{ on } K\text{)} \tag{37}$$

$$\leqslant \|f\|_\theta\, \tau\, \Delta^\theta\,. \qquad \text{(by (33))} \tag{38}$$

Let V be the closed ρ-neighborhood of $\mathcal{O}$, where:

$$\rho \coloneqq \varepsilon^{-1/\theta}|\gamma|^{1/\theta} \tag{39}$$

$$\leqslant (\tau/\varepsilon)^{1/\theta}\,\Delta\,. \tag{40}$$

Then:

$$g < \gamma \text{ outside of } V; \tag{41}$$

indeed:

$$y \notin V \quad \Rightarrow \quad d(y,\mathcal{O}) > \rho \tag{42}$$

$$\Rightarrow \qquad g(y) = \underbrace{f(y)}_{\leqslant 0} - \varepsilon[d(y,\mathcal{O})]^\theta < -\varepsilon\rho^\theta = \gamma \tag{43}$$

Recall that our aim is to show that $\beta(g) \leqslant \gamma$. By formula (30), it is sufficient to prove the following statement:

$$\forall y \in X\ \exists N = N(y) \geqslant 1 \text{ such that } g^{(N)}(y) \leqslant N\gamma\,. \tag{44}$$

There are two situations where it is easy to find the time N:

- If $y \in \mathcal{O}$ then we can take $N(y) = p$ (the period of $\mathcal{O}$), due to the very definition (35) of γ.
- If $y \notin V$ then we take $N(y) = 1$, due to property (41).

Let us consider the remaining case, so fix a point y in $V \smallsetminus \mathcal{O}$, that is, such that $0 < d(y,\mathcal{O}) < \rho$.

Let z be the point on the periodic orbit $\mathcal{O}$ which is closest to y. As seen before in (40), ρ is less than the half-gap of $\mathcal{O}$, and so the point z is uniquely determined.

By expansivity (3), the separation $d(T^j y, T^j z)$ between the orbits of y and z increase until it becomes bigger than r (which is $\geqslant \Delta$, by definition of the half-gap (5)). Let N be the least positive integer such that:

$$d(T^{N-1}y, T^{N-1}z) > \Delta\,. \tag{45}$$

Then:

$$j \in \{0,1,\dots,N-2\}, \quad d(T^j y,\mathcal{O}) = d(T^j y, T^j z) \quad \text{(by def. of half-gap } \Delta\text{)} \tag{46}$$

$$\leqslant \lambda^j d(y,z) \qquad \text{(by (3))}. \tag{47}$$

We will show that N has the desired property (44), that is, the Birkhoff sum $(\gamma - g)^{(N)}(y)$ is $\geqslant 0$.

Let $n \geqslant 1$ be the time it takes for y to leave the neighborhood V, that is, the least positive integer such that:

$$d(T^n y, T^n z) > \rho\,. \tag{48}$$

Since $\rho < \Delta$ (as a consequence of (40) and $\tau < \varepsilon$), we have $n \leqslant N-1$.

We break the Birkhoff sum we are interested in as follows:

$$(\gamma - g)^{(N)}(y) = \underbrace{(\gamma - g)^{(n)}(y)}_{①} + \underbrace{(\gamma - g)^{(N-n)}(T^n y)}_{②} \,. \tag{49}$$

While the Birkhoff sum ① may be negative, the other sum ② (which is nonempty, since $n < N$) consists entirely of nonnegative terms (due to property (41)). Let us bound the sum ② by its last term:

$$② \geqslant \gamma - g(T^{N-1}y) \tag{50}$$

$$= \gamma - \underbrace{f(T^{N-1}y)}_{\leqslant 0} + \varepsilon\left[d(T^{N-1}y, \mathcal{O})\right]^\theta \tag{51}$$

$$> \gamma + \varepsilon\Delta^\theta \,. \tag{52}$$

We need to show that ① cannot be too negative; we estimate it as follows:

$$① = n\gamma - g^{(n)}(y) = \underbrace{n\gamma - g^{(n)}(z)}_{③} + \underbrace{g^{(n)}(z) - g^{(n)}(y)}_{④} \,. \tag{53}$$

By euclidian division, $n = pq+r$ for appropriate integers $q \geqslant 0$ and $0 \leqslant r \leqslant p-1$. Since z belongs to the periodic orbit $\mathcal{O}$, which has period p, and the nonpositive function g has average γ on $\mathcal{O}$, we have:

$$g^{(n)}(z) = g^{(pq)}(z) + g^{(r)}(T^{pq}z) \leqslant pq\gamma \tag{54}$$

and therefore:

$$③ \geqslant r\gamma \geqslant (p-1)\gamma \tag{55}$$

(since $\gamma \leqslant 0$).

Next, we estimate:

$$\left|④\right| \leqslant \sum_{j=0}^{n-1} \left|g(T^j z) - g(T^j y)\right| \tag{56}$$

$$\leqslant \sum_{j=0}^{n-1} \left[\left|f(T^j z) - f(T^j y)\right| + \varepsilon \big[\underbrace{d(T^j y, \mathcal{O})}_{d(T^j y, T^j z)} \big]^\theta \right] \tag{57}$$

$$\leqslant \sum_{j=0}^{n-1} \big(\|f\|_\theta + \varepsilon\big) \left[d(T^j y, T^j z)\right]^\theta \tag{58}$$

$$\leqslant \sum_{j=0}^{n-1} \big(\|f\|_\theta + 1\big) \lambda^{-(n-1-j)\theta} \big[\underbrace{d(T^{n-1}y, T^{n-1}z)}_{<\rho} \big]^\theta \tag{59}$$

$$\leqslant \underbrace{\frac{\|f\|_\theta + 1}{1 - \lambda^{-\theta}}}_{C} \rho^\theta \tag{60}$$

$$= C\varepsilon^{-1}|\gamma| \qquad \text{(by (39))} \tag{61}$$

Collecting the estimates above, we obtain:

$$(\gamma - g)^{(N)}(y) = ② + ③ + ④ \tag{62}$$

$$> \varepsilon\Delta^\theta - p|\gamma| - C\varepsilon^{-1}|\gamma| \tag{63}$$

$$\geqslant \varepsilon\Delta^\theta - (1 + C\varepsilon^{-1})p|\gamma| \qquad \text{(since } p \geqslant 1\text{)} \tag{64}$$

$$\geqslant \big[\varepsilon - \|f\|_\theta\,(1 + C\varepsilon^{-1})\tau\big]\Delta^\theta \quad \text{(by (38))}\,. \tag{65}$$

Now we see explicitly how small τ needs to be (in addition to being less than ε): if

$$\tau < \frac{\varepsilon^2}{\|f\|_\theta\,(C+\varepsilon)} \tag{66}$$

then the estimates above ensure that $(\gamma - g)^{(N)}(y) > 0$. This completes the proof of property (44) and Theorem 1.1.

Acknowledgements. I thank Yiwei Zhang for kindly explaining to me the main ideas of the proof.

References

[BLL] Baraviera, A.; Leplaideur, R.; Lopes, A. – *Ergodic optimization, zero temperature limits and the max-plus algebra.* Publicações Matemáticas do IMPA. 29º Colóquio Brasileiro de Matemática. IMPA, Rio de Janeiro, 2013. (Cited on page 1.)

[Bo] Bochi, J. – Ergodic optimization of Birkhoff averages and Lyapunov exponents. *Proc. Int. Cong. of Math. – 2018 Rio de Janeiro*, Vol. 2, 1821–1842. (Cited on page 1.)

[BG] Bochi, J.; Garibaldi, E. – *Extremal norms for fiber-bunched cocycles.* arXiv:1808.02804 (Cited on page 3.)

[BZ1] Bochi, J.; Zhang, Y. – Note on robustness of periodic measures in ergodic optimization. `http://www.mat.uc.cl/~jairo.bochi/docs/lock.pdf` (Cited on page 5.)

[BZ2] Bochi, J.; Zhang, Y. – Ergodic optimization of prevalent super-continuous functions. *IMRN* 2016, no. 19, 5988–6017. (Cited on page 2.)

[BQ] Bressaud, X.; Quas, A. – Rate of approximation of minimizing measures. *Nonlinearity* 20 (2007), no. 4, 845–853. (Cited on page 2.)

[Co] Contreras, G. – Ground states are generically a periodic orbit. *Invent. Math.* 205 (2016), no. 2, 383–412. (Cited on page 2.)

[CLT] Contreras, G.; Lopes, A.O.; Thieullen, P. – Lyapunov minimizing measures for expanding maps of the circle. *Ergodic Theory Dynam. Systems* 21 (2001), no. 5, 1379–1409. (Cited on page 5.)

[CG] Conze, J.P.; Guivarc'h, Y. – Croissance des sommes ergodiques et principe variationnel. Unpublished manuscript, circa 1993. (Cited on page 5.)

[Ga] Garibaldi, E. – *Ergodic optimization in the expanding case: concepts, tools ad applications.* SpringerBriefs in Mathematics. Springer, Cham, 2017. (Cited on page 1.)

[H$^+$] Huang, W.; Lian, Z.; Ma, X.; Xu, L.; Zhang, Y. – Ergodic optimization theory for a class of typical maps. arXiv:1904.01915 (Cited on pages 2, 3, and 6.)

[HO] Hunt, B.R.; Ott, E. – Optimal periodic orbits of chaotic systems occur at low period. *Phys. Rev. E* 54 (1996), no. 1, 328–337. (Cited on page 2.)

[Je1] Jenkinson, O. – Ergodic optimization. *Discrete Contin. Dyn. Syst.* 15 (2006), no. 1, 197–224. (Cited on page 1.)

[Je2] Jenkinson, O. – Ergodic optimization in dynamical systems. *Ergodic Theory Dynam. Systems* (to appear). DOI:10.1017/etds.2017.142 (Cited on pages 1, 2, and 5.)

[Mo1] Morris, I.D. – Maximizing measures of generic Hölder functions have zero entropy. *Nonlinearity* 21 (2008), no. 5, 993–1000. (Cited on page 2.)

[Mo2] Morris, I.D. – Mather sets for sequences of matrices and applications to the study of joint spectral radii. *Proc. LMS* 107 (2013), no. 1, 121–150. (Cited on page 5.)

[Pe] Peres, Y. – A combinatorial application of the maximal ergodic theorem. *Bull. London Math. Soc.* 20 (1988), no. 3, 248–252. (Cited on page 5.)

[PU] Przytycki, F.; Urbański, M. – *Conformal fractals: ergodic theory methods.* London Mathematical Society Lecture Note Series, 371. Cambridge University Press, Cambridge, 2010. (Cited on page 2.)

[Sa] Savchenko, S.V. – Homological inequalities for finite topological Markov chains. *Funct. Anal. Appl.* 33 (1999), no. 3, 236–238. (Cited on page 5.)

[VO] Viana, M.; Oliveira, K. – *Foundations of ergodic theory.* Cambridge Studies in Advanced Mathematics, 151. Cambridge University Press, Cambridge, 2016. (Cited on page 1.)

[YH] Yuan, G.; Hunt, B.R. – Optimal orbits of hyperbolic systems. *Nonlinearity* 12 (1999), no. 4, 1207–1224. (Cited on pages 2 and 5.)

Facultad de Matemáticas, Pontificia Universidad Católica de Chile
URL: http://www.mat.uc.cl/~jairo.bochi
Email address: jairo.bochi@mat.uc.cl