Regularity of calibrated sub-actions for circle expanding maps and Sturmian optimization

RUI GAO

College of Mathematics, Sichuan University, Chengdu 610064, China
(e-mail: gaoruimath@scu.edu.cn)

Abstract. In this short and elementary note, we study some ergodic optimization problems for circle expanding maps. We first make an observation that if a function is not far from being convex, then its calibrated sub-actions are closer to convex functions in a certain effective way. As an application of this simple observation, for a circle doubling map, we generalize a result of Bousch saying that translations of the cosine function are uniquely optimized by Sturmian measures. Our argument follows the mainline of Bousch's original proof, while some technical part is simplified by the observation mentioned above, and no numerical calculation is needed.

1. *Introduction*

Let X be a compact metric space and let $C(X)$ denote the collection of real-valued continuous functions on X. Let $T : X \to X$ be a continuous map. For the topological dynamical system (X, T), let $\mathcal{M}(X, T)$ denote the collection of T-invariant Borel probability measures on X. The *ergodic optimization* problem of (X, T) is concerned about, for given $f \in C(X)$, the extreme values (maximum or minimum) of $\mu \mapsto \int_X f \, d\mu$ for $\mu \in \mathcal{M}(X, T)$, and which measures attain the maximum or minimum. For an overview of the research topic of ergodic optimization, one may refer to Jenkinson's survey papers [10, 11] and Bochi's survey paper [2]. See also Contreras, Lopes, and Thieullen [6], and Garibaldi [9].

Since

$$\min_{\mu \in \mathcal{M}(X,T)} \int_X f \, d\mu = - \max_{\mu \in \mathcal{M}(X,T)} \int_X (-f) \, d\mu,$$

to be definite, we shall mainly focus on the maximum problem and denote

$$\beta(f) := \max_{\mu \in \mathcal{M}(X,T)} \int_X f \, d\mu,$$

$$\mathcal{M}_{\max}(f) := \left\{ \mu \in \mathcal{M}(X,T) : \int_X f \, d\mu = \beta(f) \right\}.$$

Measures contained in $\mathcal{M}_{\max}(f)$ are called *maximizing measures* of f for (X, T). Given $f \in C(X)$, a basic tool to study maximizing measures is usually referred to as *Mañé's lemma*, which aims at establishing the existence of $g \in C(X)$, called a *sub-action* of f, such that

$$\tilde{f} := f + g - g \circ T \leq \beta(f)$$

holds. One of its advantages is based on a simple observation as follows. Once such g exists, then $\beta(f) = \beta(\tilde{f})$ and $\mathcal{M}_{\max}(f) = \mathcal{M}_{\max}(\tilde{f})$; moreover, for any $\mu \in \mathcal{M}_{\max}(f)$, $\tilde{f} = \beta(f)$ holds on the support of μ. For a review of results on the existence or construction of sub-actions, one may refer to [11, §6]; see also [6, Remark 12].

Given $f \in C(X)$, a particular way to construct sub-action g of f is to solve the equation below:

$$g(x) + \beta(f) = \max_{Ty=x} (f(y) + g(y)) \quad \text{for all } x \in X. \tag{1}$$

A solution $g \in C(X)$ to equation (1) is automatically a sub-action of f, and it is called a *calibrated sub-action* of f. It is well known that when (X, T) is expanding and f is Hölder continuous, for the one-parameter family of potentials $(tf)_{t>0}$, the zero-temperature limit points (as $t \to +\infty$) of their equilibrium states are contained in $\mathcal{M}_{\max}(f)$; see, for example, [11, Theorem 4.1] and references therein. Calibrated sub-actions arise naturally in this limit process; see, for example, [6, Proposition 29]. For expanding (X, T) and Hölder continuous f, the first proofs on the existence of calibrated sub-actions were given by, among others, Savchenko [12] and Bousch [3] with different approaches. Savchenko's proof is based on a zero-temperature limit argument mentioned above; while Bousch's proof is to treat equation (1) as a fixed point problem and to apply the Schauder–Tychonoff fixed point theorem.

In the first half of this paper, for a smooth expanding system (X, T), given an observable in $C(X)$ with certain smoothness, we are interested in the regularity (or smoothness) of its calibrated sub-actions. For simplicity, let us focus on the simplest case that $T : X \to X$ is a linear circle expanding map; more precisely, $X = \mathbb{T} := \mathbb{R}/\mathbb{Z}$ is the unit circle, and $Tx = dx$ for $x \in \mathbb{T}$, where $d \geq 2$ is an integer. Our main results in this part are Theorem 2.3 and Corollary 2.4, and they roughly read as follows.

THEOREM A. *For $f \in C(\mathbb{T})$ identified as a function on $\mathbb{R}$ of period 1, if f can be written as a sum of a convex function and a quadratic function, then any calibrated sub-action g of f for the $x \mapsto dx$ system can be written in the same form. In particular, both one-sided derivatives of g exist everywhere and only allow at most countably many discontinuities, and the set of points where g is not differentiable is also at most countable.*

It might be interesting to compare the statement above with the following result of Bousch and Jenkinson [**4**, Theorem B]: for the $x \mapsto 2x$ system, there exists a real-analytic function such that any sub-action of it fails to be C^1.

In the second half of this paper, as an application of Theorem 2.3, we shall generalize a classic result of Bousch [**3**] roughly stated below with a simpler proof.

Theorem. (Bousch [**3**, Théorème A]) For each $\omega \in \mathbb{T}$, the function $x \mapsto \cos 2\pi(x - \omega)$ has a unique maximizing measure for the $x \mapsto 2x$ system; moreover, the maximizing measure is Sturmian, that is, supported on a semi-circle.

Our main technical result in the second half of this paper is Theorem 3.5. As a corollary of it, we can easily prove Theorem 3.8 as a generalization of [**3**, Théorème A]. A slightly weaker and easier to read version of Theorem 3.8 is as follows.

THEOREM B. *Let $f : \mathbb{T} \to \mathbb{R}$ be a C^2 function with the following properties, where f is identified as a function on $\mathbb{R}$ of period* 1*:*
- $f(x + \frac{1}{2}) = -f(x)$;
- $f(x) = f(-x)$;
- $f'' \le 0$ on $[-\frac{1}{4}, \frac{1}{4}]$;
- $f(0) > 1/40 \cdot \max_{x \in \mathbb{T}} f''(x)$.

Then for each $\omega \in \mathbb{T}$, the function $x \mapsto f(x - \omega)$ has a unique maximizing measure for the $x \mapsto 2x$ system; moreover, the maximizing measure is Sturmian.

To prove Theorem 3.5, we follow Bousch's proof of [**3**, Théorème A], while Theorem 2.3 is used to simplify some technical steps in Bousch's argument, especially [**3**, Lemme Technique 2]. No numerical calculation is needed in our argument.

The paper is organized as follows. In §2, we mainly prove Theorem 2.3. The statements are given in §2.1 and the proof is given in §2.2. In §3, we try to generalize [**3**, Théorème A]; the main results are Theorem 3.5 and its corollary Theorem 3.8. In §3.1, we recall Bousch's ideas in his proof of [**3**, Théorème A] that will be used for preparation. In §3.2, with the help of Theorem 2.3, we prove Theorem 3.5. Finally, Theorem 3.8 is proved in §3.3.

The following notation or conventions are used throughout the paper.
- For a topological space X, $C(X)$ denotes the collection of real-valued continuous functions on X.
- For a metric space X, $\mathrm{Lip}(X)$ denotes the collection of real-valued Lipschitz functions on X.
- For a (one-dimensional) smooth manifold (possibly with boundary) X, $C^k(X)$ denotes the collection of real-valued functions that are continuously differentiable up to order k, where $k \ge 1$ is an integer.
- $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ denotes the unit circle. The following notation will be used when there is no ambiguity.
  - For $x \in \mathbb{T}$ and $y \in \mathbb{R}$, $x + (y \bmod 1) \in \mathbb{T}$ will be denoted by $x + y$ for short.
  - For $a \in \mathbb{R}$ and $b \in (a, a + 1)$, the arc $\{x \bmod 1 : x \in [a, b]\} \subset \mathbb{T}$ will be denoted by $[a, b]$ for short.

- A function $f : \mathbb{T} \to \mathbb{R}$ will be frequently identified as a function on $\mathbb{R}$ of period 1.
- 'Lebesgue almost everywhere' is abbreviated as 'a.e.'.

2. *Regularity of calibrated sub-actions*

Throughout this section, the dynamical system is fixed to be

$$T_d : \mathbb{T} \to \mathbb{T}, \quad x \mapsto dx,$$

where $d \geq 2$ is an integer. We focus on the regularity of calibrated sub-actions for this system. Let us begin with introducing the following notation to express equation (1) of calibrated sub-action in a more convenient way. Define $\mathcal{M}_d : C(\mathbb{T}) \to C(\mathbb{T})$ as follows:

$$\mathcal{M}_d f(x) := \max_{T_d y = x} f(y) = \max_{0 \leq k < d} f\left(\frac{x + k}{d}\right) \quad \text{for all } f \in C(\mathbb{T}), \tag{2}$$

where, in the expression $f((x + k)/d)$, f is considered as a function on $\mathbb{R}$ of period 1. Note that by definition, for $f \in C(\mathbb{T})$,

$$g \text{ is a calibrated sub-action of } f \iff g \in C(\mathbb{T}) \quad \text{and} \quad g + \beta(f) = \mathcal{M}_d(f + g).$$

Also note that $\mathrm{Lip}(\mathbb{T})$ is $\mathcal{M}_d$-invariant. The following is well known.

LEMMA 2.1. *For every $f \in \mathrm{Lip}(\mathbb{T})$, there exists $g \in \mathrm{Lip}(\mathbb{T})$ such that g is a calibrated sub-action of f. Moreover, if f has a unique maximizing measure, then g is uniquely determined by f up to an additive constant.*

For the existence of g, see, for example [6, Proposition 29(iii)]. For the uniqueness part, see, for example [3, Lemme C], which is stated for $d = 2$, but the proof there is easily adapted to arbitrary $d \geq 2$ or more general systems; see also [9, Proposition 6.7].

2.1. *Statement of results.* To characterize the regularity of calibrated sub-actions and to state Theorem 2.3 later on, we need the following quantity $\eta(\cdot)$ to measure how far a continuous function is away from being convex.

Definition 2.1. Given $f \in C(\mathbb{R})$, let $\eta(f)$ be the infimum of $a \in \mathbb{R}$ (we adopt the convention that $\inf \varnothing = +\infty$) such that the following holds:

$$f(x + \delta) + f(x - \delta) - 2f(x) \geq -a\delta^2 \quad \text{for all } x, \delta \in \mathbb{R}.$$

By definition, the following is evident.

LEMMA 2.2. *Given $f \in C(\mathbb{R})$ and $a \in \mathbb{R}$, $\eta(f) \leq a$ if and only if $x \mapsto f(x) + (a/2)x^2$ is convex on $\mathbb{R}$.*

It will be useful to introduce the following notation closely related to $\eta(\cdot)$. Given $\delta > 0$, $x \in \mathbb{R}$, and $f \in C(\mathbb{R})$, denote

$$\xi_\delta^x(f) := 2f(x) - f(x + \delta) - f(x - \delta), \tag{3}$$

$$\xi_\delta^*(f) := \sup_{x \in \mathbb{R}} \xi_\delta^x(f). \tag{4}$$

Note that by definition,

$$\eta(f) = \sup_{\delta > 0} \delta^{-2} \xi_\delta^*(f) \quad \text{for all } f \in C(\mathbb{R}). \tag{5}$$

Also note that $\eta(f)$, $\xi_\delta^x(f)$, and $\xi_\delta^*(f)$ are well defined for $f \in C(\mathbb{T})$ considered as a function on $\mathbb{R}$ of period 1. It might be mentioned that when $f \in C(\mathbb{T})$, $\xi_\delta^*(f) \geq 0$ for any $\delta > 0$, and therefore $\eta(f) \geq 0$, although we shall not use these facts. To see $\xi_\delta^*(f) \geq 0$, let us consider f as a continuous function on $\mathbb{R}$ of period 1. Then f attains its maximum on $\mathbb{R}$ at some x, and hence $\xi_\delta^*(f) \geq \xi_\delta^x(f) \geq 0$.

Now we are ready to state our main result in this section.

THEOREM 2.3. *Given $d \geq 2$, let $f \in C(\mathbb{T})$ and suppose that there exists a calibrated sub-action $g \in C(\mathbb{T})$ of f for the $x \mapsto dx$ system. Then we have*

$$\xi_\delta^*(g) \leq \xi_{d^{-1}\delta}^*(f) + \xi_{d^{-1}\delta}^*(g) \quad \text{for all } \delta > 0, \tag{6}$$

$$\eta(g) \leq (d^2 - 1)^{-1} \cdot \eta(f). \tag{7}$$

The proof of Theorem 2.3 will be given in the next subsection. Let us state a corollary of it first, which indicates, under the regularity condition $\eta(f) < +\infty$, how regular a calibrated sub-action of f could be. It should be mentioned that, as the referee pointed out to the author, in [3, Lemme B], Bousch already showed that for any $f \in \text{Lip}(\mathbb{T})$ and any calibrated sub-action $g \in C(\mathbb{T})$ of f, g is automatically Lipschitz and $lip(g) \leq lip(f)/(d - 1)$ holds, where $lip(\cdot)$ denotes the Lipschitz constant of a function (although Bousch only focused on $d = 2$, his argument works equally well for general $d \geq 2$). The corollary below asserts that, under the additional assumption $\eta(f) < +\infty$, we can say something more about the regularity of g.

COROLLARY 2.4. *Given $d \geq 2$, let $f \in \text{Lip}(\mathbb{T})$ be with $\eta(f) < +\infty$ and let $g \in C(\mathbb{T})$ be a calibrated sub-action of f for the $x \mapsto dx$ system. Then the following hold.*
- *Both of the one-sided derivatives $g'_\pm(x) := \lim_{\Delta \searrow 0}((g(x \pm \Delta) - g(x))/\pm\Delta)$ of g exist at each $x \in \mathbb{T}$.*
- *$g'_\pm$ are of bounded variation on $\mathbb{T}$; that is to say, identifying $g'_\pm$ as functions on $\mathbb{R}$ of period 1, $g'_\pm$ are of bounded variation on finite subintervals of $\mathbb{R}$.*
- *Given $x \in \mathbb{T}$, $g'(x)$ does not exist if and only if $g'_+(x) > g'_-(x)$; the set $\{x \in \mathbb{T} : g'(x) \text{ does not exist}\}$ is at most countable.*

Proof. By equation (7) in Theorem 2.3, $\eta(g) < +\infty$. Then, according to Lemma 2.2, identifying g as a function on $\mathbb{R}$ of period 1, $x \mapsto g(x) + (\eta(g)/2)x^2$ is convex on $\mathbb{R}$. All the statements follow from basic properties of convex functions. $\square$

2.2. *Proof of Theorem 2.3.* This subsection is devoted to the proof of Theorem 2.3. Let us begin with some common elementary properties shared by $\eta(\cdot)$ and functionals defined in equations (4) and (4).

LEMMA 2.5. *Let X be a topological space and recall that $C(X)$ denotes the vector space of real-valued continuous functions on X. Let Γ denote the collection of γ :*

$C(X) \to (-\infty, +\infty]$ *satisfying the following assumptions (we adopt the convention* $0 \cdot (+\infty) = 0$ *below):*

- $\gamma(f + g) \le \gamma(f) + \gamma(g)$;
- $\gamma(\max\{f, g\}) \le \max\{\gamma(f), \gamma(g)\}$;
- $\gamma(af + b) = a\gamma(f)$ *for* $a \ge 0, b \in \mathbb{R}$.

Then Γ *has the following properties:*

- *if* $a, b \ge 0$ *and* $\alpha, \beta \in \Gamma$, *then* $a\alpha + b\beta \in \Gamma$;
- *if* $\Lambda \subset \Gamma$, *then* $\sup_{\gamma \in \Lambda} \gamma \in \Gamma$.

Moreover, for $X = \mathbb{R}$, $\xi_\delta^x \in \Gamma$ *for any* $\delta > 0$, $x \in \mathbb{R}$. *As a result,* $\xi_\delta^* \in \Gamma$ *for any* $\delta > 0$ *and hence* $\eta \in \Gamma$.

Proof. All the statements are more or less evident and can be easily proved by checking the definition directly. Let us only verify the relatively less obvious fact that

$$\xi_\delta^x(\max\{f, g\}) \le \max\{\xi_\delta^x(f), \xi_\delta^x(g)\},$$

and the rest is left to the reader. To prove the inequality above, denote $h := \max\{f, g\}$. If $h(x) = f(x)$, then

$$\xi_\delta^x(h) = 2f(x) - h(x - \delta) - h(x + \delta) \le \xi_\delta^x(f),$$

where the '$\le$' is due to $f \le h$; otherwise, $h(x) = g(x)$ and the same argument shows that $\xi_\delta^x(h) \le \xi_\delta^x(g)$. The proof is done. $\square$

The key ingredient to Theorem 2.3 is the simple fact below. Recall the operator $\mathcal{M}_d$ defined in equation (2).

PROPOSITION 2.6. *Given* $d \ge 2$ *and* $f \in C(\mathbb{T})$, *the following hold:*

$$\xi_\delta^*(\mathcal{M}_d f) \le \xi_{d^{-1}\delta}^*(f) \quad \text{for all } \delta > 0; \tag{8}$$

$$\eta(\mathcal{M}_d f) \le d^{-2} \cdot \eta(f). \tag{9}$$

Proof. Identify f as a function on $\mathbb{R}$ of period 1. For $0 \le k < d$, let $f_k \in C(\mathbb{R})$ be defined by $f_k(x) := f((x + k)/d)$. By definition, $\mathcal{M}_d f = \max_{0 \le k < d} f_k$, and

$$\xi_\delta^*(f_k) = \xi_{d^{-1}\delta}^*(f) \quad \text{for all } 0 \le k < d, \delta > 0.$$

It follows that

$$\xi_\delta^*(\mathcal{M}_d f) = \xi_\delta^*(\max_{0 \le k < d} f_k) \le \max_{0 \le k < d} \xi_\delta^*(f_k) = \xi_{d^{-1}\delta}^*(f),$$

where the '$\le$' is due to Lemma 2.5. This completes the proof of equation (8). Equation (9) follows from equations (8) and (5) immediately. $\square$

Now we are ready to prove Theorem 2.3.

Proof of Theorem 2.3. Since $g + \beta(f) = \mathcal{M}_d(f + g)$, equation (6) follows from equation (8) in Proposition 2.6 and Lemma 2.5 directly. To prove equation (7), first

note that by equations (6) and (5), the following holds for any $\delta > 0$:

$$\xi_\delta^*(g) \leq \xi_{d^{-1}\delta}^*(f) + \xi_{d^{-1}\delta}^*(g) \leq (d^{-1}\delta)^2 \cdot \eta(f) + \xi_{d^{-1}\delta}^*(g).$$

Iterating this inequality repeatedly and noting that $\xi_{d^{-n}\delta}^*(g) \to 0$ as $n \to \infty$, we obtain that

$$\xi_\delta^*(g) \leq \sum_{n=1}^{\infty} (d^{-n}\delta)^2 \cdot \eta(f) = \delta^2 \cdot (d^2 - 1)^{-1} \cdot \eta(f).$$

Since $\delta > 0$ is arbitrary, equation (7) follows. $\qquad\square$

3. *Sturmian optimization*

In this section, we fix $d = 2$ and denote $Tx = 2x$, $x \in \mathbb{T}$. For this system, we shall show that certain functions are uniquely maximized by Sturmian measures, which generalizes the result of Bousch [3, Théorème A]. The results are stated in Theorem 3.5 and its consequences in Corollary 3.6 and Theorem 3.8. Recall that for any closed semicircle $S \subset \mathbb{T}$ (that is, $S = [\gamma, \gamma + \frac{1}{2}]$ for some $\gamma \in \mathbb{R}$), there exists a unique T-invariant Borel probability measure supported on S, called the *Sturmian measure* on S. Bullett and Sentenac [5] studied Sturmian measures systematically. See also [8, §6] for connection of Sturmian measures with symbolic dynamics.

3.1. *Bousch's criterion of Sturmian condition.* This subsection is devoted to reviewing the basic ideas of Bousch for his proof of [3, Théorème A].

Definition 3.1. (Bousch [3, p. 497]) Let $f \in \text{Lip}(\mathbb{T})$ and let $g \in \text{Lip}(\mathbb{T})$ be a calibrated sub-action of f for the $x \mapsto 2x$ system. Denote

$$R(x) = R_{f,g}(x) := f(x) + g(x) - f\left(x + \tfrac{1}{2}\right) - g\left(x + \tfrac{1}{2}\right), \quad x \in \mathbb{T}. \tag{10}$$

If $\{x \in \mathbb{T} : R(x) = 0\}$ consists of a single pair of antipodal points, then we say that (f, g) satisfies the *Sturmian condition*. We also say that f satisfies the Sturmian condition if (f, g) satisfies the Sturmian condition for some g.

By definition, it is easy to see the following holds, which is a combination of the Proposition in of [3, p. 497] and [3, Lemme C].

LEMMA 3.1. (Bousch [3]) *If $f \in \text{Lip}(\mathbb{T})$ satisfies the Sturmian condition, then f admits a unique maximizing measure. Moreover, this measure is a Sturmian measure, and up to an additive constant, there exists a unique calibrated sub-action of f.*

Remark. As a corollary, if $f \in \text{Lip}(\mathbb{T})$ satisfies the Sturmian condition, then for the equilibrium state of tf $(t > 0)$, its zero-temperature limit (as $t \to +\infty$) exists. See, for example, [11, Theorem 4.1] for details.

The basic idea of Bousch [3] to verify the Sturmian condition is summarized below.

LEMMA 3.2. *Let $f \in \mathrm{Lip}(\mathbb{T})$ and let $g \in \mathrm{Lip}(\mathbb{T})$ be a calibrated sub-action of f. Assume that for R defined by equation (10), there exist $a \in \mathbb{R}$ and $b \in (a, a + \frac{1}{2})$ with the following properties:*

(S1) $R > 0$ *on* $[a, b]$;
(S2)   *R is strictly decreasing on* $[b, a + \frac{1}{2}]$.
Then (f, g) satisfies the Sturmian condition.

Let us provide a proof for completeness.

Proof. By definition, $R(x + \frac{1}{2}) = -R(x)$, so that $\{x \in \mathbb{T} : R(x) = 0\}$ is non-empty and consists of pairs of antipodal points. However, the assumptions (S1) and (S2) imply that $\{x \in \mathbb{T} : R(x) = 0\} \cap [a, a + \frac{1}{2}]$ contains at most one point. The conclusion follows. □

To get an effective test for assumption (S1) in Lemma 3.2, Bousch made the following observation in his proof of [3, Lemme Technique 1]. Recall the notation introduced in equations (3) and (4).

LEMMA 3.3. *Let $f \in \mathrm{Lip}(\mathbb{T})$ and let $g \in \mathrm{Lip}(\mathbb{T})$ be a calibrated sub-action of f. Then for any $n \geq 2$ and any $x \in \mathbb{T}$, we have*

$$-2\left(g(x) - g\left(x + \frac{1}{2}\right)\right) \leq \max_{y \in T^{-(n-1)}(x+1/2)} \sum_{k=2}^{n} \xi_{2^{-k}}^{T^{n-k}y}(f) + \xi_{2^{-n}}^{*}(g). \qquad (11)$$

Let us reproduce Bousch's proof for completeness.

Proof. We follow the beginning part in the proof of [3, Lemme Technique 1]. Without loss of generality, assume that $\beta(f) = 0$ for simplicity, so that equation (1) becomes

$$g(x) = \max_{Ty=x}(f(y) + g(y)) \quad \text{for all } x \in \mathbb{T}.$$

To prove equation (11), fix $x \in \mathbb{T}$, denote $x_1 = x + \frac{1}{2}$, and for each $n \geq 2$, choose $x_n \in T^{-1}x_{n-1}$ such that

$$g(x_{n-1}) = f(x_n) + g(x_n)$$

inductively on n. Noting that $T(x_n \pm 2^{-n}) = x_{n-1} \pm 2^{-n+1}$, we have

$$g(x_{n-1} + 2^{-n+1}) \geq f(x_n + 2^{-n}) + g(x_n + 2^{-n}),$$

$$g(x_{n-1} - 2^{-n+1}) \geq f(x_n - 2^{-n}) + g(x_n - 2^{-n}).$$

It follows that for each $n \geq 2$,

$$\xi_{2^{-(n-1)}}^{x_{n-1}}(g) \leq \xi_{2^{-n}}^{x_n}(f) + \xi_{2^{-n}}^{x_n}(g),$$

and therefore

$$-2\left(g(x) - g\left(x + \frac{1}{2}\right)\right) = \xi_{2^{-1}}^{x_1}(g) \leq \sum_{k=2}^{n} \xi_{2^{-k}}^{x_k}(f) + \xi_{2^{-n}}^{x_n}(g).$$

The proof of equation (11) is completed. □

3.2. *Sufficient conditions for the Sturmian condition.* In this subsection, based on Lemma 3.2 and Theorem 2.3, we shall prove Theorem 3.5 and Corollary 3.6, which provide sufficient conditions for the Sturmian condition that are easy to check for specific observable *f*. Let us begin with a simple fact that will be used to verify assumption (S2) in Lemma 3.2.

LEMMA 3.4. *The following holds for any $f \in \mathrm{Lip}(\mathbb{R})$ with $\eta(f) < +\infty$ and any $\delta > 0$:*

$$f'(x) - f'(x + \delta) \le \eta(f) \cdot \delta \quad a.e. \ x \in \mathbb{R}.$$

Proof. Let $\tilde{f}(x) := f(x) + (\eta(f)/2)x^2$. Then by Lemma 2.2, $\tilde{f}$ is convex on $\mathbb{R}$, so that for any $\delta > 0$,

$$f'(x) - f'(x + \delta) - \eta(f) \cdot \delta = \tilde{f}'(x) - \tilde{f}'(x + \delta) \le 0 \quad a.e. \ x \in \mathbb{R}. \qquad \square$$

Our main technical result in this section is the following. Its statement might look complicated, so we shall present a simplified version in Corollary 3.6 and further deduce another result in Theorem 3.8.

THEOREM 3.5. *Let $f \in \mathrm{Lip}(\mathbb{T})$ be with $\eta(f) < +\infty$. Suppose that there exist $a \in \mathbb{R}$ and $b \in (a, a + \frac{1}{2})$ such that equations (12) and (13) below hold:*

$$f(x) - f\left(x + \frac{1}{2}\right) - \frac{1}{2} \cdot \max_{y \in T^{-1}(x+1/2)} \xi_{1/4}^y(f) > \frac{1}{96}\eta(f) \quad \textit{for all } x \in [a, b]; \quad (12)$$

$$f'(x) - f'\left(x + \frac{1}{2}\right) < -\frac{1}{6}\eta(f) \quad a.e. \ x \in \left[b, a + \frac{1}{2}\right]. \quad (13)$$

Then f satisfies the Sturmian condition.

Proof. Let $g \in \mathrm{Lip}(\mathbb{T})$ be a calibrated sub-action of *f*. Note that by equation (7), $\eta(g) \le \frac{1}{3}\eta(f)$. It suffices to verify the two assumptions (S1) and (S2) in Lemma 3.2.

Verifying (S1). First, taking $n = 2$ in equation (11) yields that

$$g(x) - g\left(x + \frac{1}{2}\right) \ge -\frac{1}{2} \cdot \max_{y \in T^{-1}(x+1/2)} \xi_{1/4}^y(f) - \frac{1}{2} \cdot \xi_{1/4}^*(g).$$

Second, due to equation (5) and $\eta(g) \le \frac{1}{3}\eta(f)$,

$$\xi_{1/4}^*(g) \le \frac{1}{16}\eta(g) \le \frac{1}{48}\eta(f).$$

Combining the inequalities in the two displayed lines above with equation (12) and the definition of *R*, assumption (S1) is verified.

Verifying (S2). It suffices to show that $R' < 0$ a.e. on $[b, a + \frac{1}{2}]$. Identify *g* as a function on $\mathbb{R}$ of period 1. Applying Lemma 3.4 to *g* with $\delta = \frac{1}{2}$, we have

$$g'(x) - g'(x + \tfrac{1}{2}) \le \tfrac{1}{2}\eta(g) \quad a.e. \ x \in \mathbb{R}.$$

Combining this with $\eta(g) \le \frac{1}{3}\eta(f)$, equation (13), and the definition of *R*, the conclusion follows. $\qquad \square$

To simplify the expressions in equations (12) and (13), let us add certain symmetry ((A0) in Definition 3.2) to f and introduce the following families of functions. Then the statement of Theorem 3.5 can be reduced to Corollary 3.6 below.

Definition 3.2. Given $a \in \mathbb{R}$, $b \in (a, a + \frac{1}{2})$ and $v \in \mathbb{R}$, let $\mathcal{A}_{a,b}^{v}$ denote the collection of $f \in \text{Lip}(\mathbb{T})$ such that the assumptions (A0), (A1), (A2) below hold.

(A0) $f(x) + f(x + \frac{1}{2}) = 2v$ for $x \in \mathbb{T}$.

(A1) $2f(x) - v - \max_{y \in \mathbb{T}} f(y) > (1/96)\eta(f)$ for $x \in [a, b]$.

(A2) $f'(x) < -(1/12)\eta(f)$ for a.e. $x \in [b, a + \frac{1}{2}]$.

Moreover, denote $\mathcal{A}_{a,b} := \bigcup_{v \in \mathbb{R}} \mathcal{A}_{a,b}^{v}$.

Remark. By definition, if $f \in \mathcal{A}_{a,b}^{v}$, then for $f_{\omega}(x) := f(x - \omega)$, $f_{\omega} \in \mathcal{A}_{a+\omega, b+\omega}^{v}$ for any $\omega \in \mathbb{R}$. Moreover, $\mathcal{A}_{a,b}$ is a 'cone' in the sense that if $f, g \in \mathcal{A}_{a,b}$ and $t > 0$, then $tf, f + g \in \mathcal{A}_{a,b}$. This can be easily checked by definition and Lemma 2.5.

In the statement below, by saying that a T-invariant measure is a *minimizing measure* of $f \in C(\mathbb{T})$, we mean that it is a maximizing measure of $-f$.

COROLLARY 3.6. *For $\mathcal{A}_{a,b}$ defined in Definition 3.2, let $f \in \mathcal{A}_{a,b}$. Then for any $\omega \in \mathbb{R}$, $f_{\omega}(x) := f(x - \omega)$ satisfies the Sturmian condition. As a result, f_{ω} admits a unique maximizing measure and a unique minimizing measure, and both of them are Sturmian measures.*

Remark. Results in [1, Theorem 1] and [7, Theorem 4.1] are of the same style as Corollary 3.6 but cannot be recovered by it.

Proof. Let $f \in \mathcal{A}_{a,b}^{v}$ for some $v \in \mathbb{R}$. According to the remark ahead of the corollary, to show f_{ω} satisfies the Sturmian condition, it suffices to prove that f satisfies the Sturmian condition. First, since f satisfies (A0),

$$f(x) - f\left(x + \tfrac{1}{2}\right) - \tfrac{1}{2} \cdot \xi_{1/4}^{y}(f) = 2f(x) - v - f(y) \quad \text{for all } x, y \in \mathbb{T}.$$

Combining this with assumption (A1), equation (12) holds. Second, combining assumptions (A0) with (A2), equation (13) holds. Therefore, by Theorem 3.5, f satisfies the Sturmian condition.

Finally, since f_{ω} satisfies the Sturmian condition, by Lemma 3.1, it admits a unique maximizing measure which is Sturmian; since $f_{\omega} = 2v - f_{\omega+1/2}$, the minimizing measure of f_{ω} is also unique and Sturmian. □

3.3. *A family of functions looking like cosine.* As a concrete application of Corollary 3.6, we can recover the result of Bousch [3, Théorème A]: for $f(x) = \cos 2\pi x$, $\eta(f) = 4\pi^{2}$ and it is easy to check that $f \in \mathcal{A}_{-1/8, 1/8}$ (or $f \in \mathcal{A}_{-1/10, 1/10}$).

To give a sufficient condition for the Sturmian condition more explicit than Corollary 3.6, by mimicking behaviors of the cosine function, we add more properties to the observable f and introduce the following family of functions. A sufficient condition will be given in Theorem 3.8.

Definition 3.3. Let $\mathcal{B}$ denote the collection of $f \in \mathrm{Lip}(\mathbb{T})$ with the following properties.

- $x \mapsto f(x) + f(x + \frac{1}{2})$ is constant on $\mathbb{T}$.
- $f(-x) = f(x)$ for $x \in \mathbb{T}$.
- $\eta(f) < +\infty$.
- f is concave on $[-\frac{1}{4}, \frac{1}{4}]$.

By definition, the following is evident, where esssup (essinf) means the essential supremum (infimum) relative to Lebesgue measure.

LEMMA 3.7. $\mathcal{B}$ is a subset of $\{f \in C^1(\mathbb{T}) : f' \in \mathrm{Lip}(\mathbb{T}), f'(0) = 0\}$. *Moreover,*

$$\eta(f) = \mathrm{esssup}_{x \in \mathbb{T}} f''(x) = -\mathrm{essinf}_{x \in \mathbb{T}} f''(x) \quad \text{for all } f \in \mathcal{B}.$$

THEOREM 3.8. *Denote* $\kappa := 7/96 - \sqrt{3}/36 > 0$. *Let* $f \in \mathcal{B}$ *and suppose that*

$$f(0) - f(\tfrac{1}{4}) > \kappa \cdot \eta(f).$$

Then $x \mapsto f(x - \omega)$ *satisfies the Sturmian condition for any* $\omega \in \mathbb{R}$.

Remark.
- $7/96 - \sqrt{3}/36 < 0.02481 < 1/40$; however,

$$\sup_{f \in \mathcal{B}} \frac{f(0) - f(1/4)}{\eta(f)} = \frac{1}{32} = 0.03125,$$

  and the supremum is attained at $f \in \mathcal{B}$ defined by $f(x) = -x^2$, $x \in [0, \frac{1}{4}]$.
- For $f(x) = \cos 2\pi x \in \mathcal{B}$, $(f(0) - f(\frac{1}{4}))/\eta(f) = 1/4\pi^2 > 0.02533$.

Proof. Without loss of generality, assume that $f(\frac{1}{4}) = 0$ for simplicity. It suffices to find $c \in (0, \frac{1}{4})$ such that $f \in \mathcal{A}_{a,b}$ for $a = -c$, $b = c$. Assumption (A0) is automatically satisfied for $v = 0$. Since f is decreasing on $[0, \frac{1}{2}]$ and since $f(-x) = f(x)$, assumption (A1) is reduced to $2f(c) - f(0) > (1/96)\eta(f)$. Since f' is decreasing on $[0, \frac{1}{4}]$ and since $f'(\frac{1}{2} - x) = f'(x)$, assumption (A2) is reduced to $f'(c) < -(1/12)\eta(f)$. Moreover, $h := f(0) - f$ satisfies all the assumptions in Lemma 3.9 below, and assumptions (A1), (A2) are further reduced to assumptions (H1), (H2) in Lemma 3.9. Then, applying Lemma 3.9 to this h, the conclusion follows. □

LEMMA 3.9. *Let* $h : [0, \frac{1}{4}] \to \mathbb{R}$ *satisfy the following.*
- h *is* C^1, *and* h' *is Lipschitz and increasing.*
- $h(0) = h'(0) = 0$.
- $h(\frac{1}{4}) > \kappa \cdot \eta$ *for* $\kappa := 7/96 - \sqrt{3}/36$ *and* $\eta := \mathrm{esssup}_{x \in [0,1/4]} h''(x)$.

Then there exists $c \in (0, \frac{1}{4})$ *such that*
(H1) $h(\frac{1}{4}) - 2h(c) > \eta/96$,
(H2) $h'(c) > \eta/12$.

Proof. Without loss of generality, we may assume $\eta = 1$, because all the assumptions of h are positively homogeneous. Then $h'' \leq 1$ a.e. Let $b \in [0, \frac{1}{4}]$ be maximal such that

$h'(b) \leq 1/12$. Since $h'(0) = 0$ and $h'' \leq 1$ a.e., $b > 1/12$ and

$$h'(x) \leq \begin{cases} x, & 0 \leq x \leq \dfrac{1}{12}, \\ \dfrac{1}{12}, & \dfrac{1}{12} < x \leq b, \\ \dfrac{1}{12} + x - b, & b < x \leq \dfrac{1}{4}. \end{cases}$$

It follows that

$$h(b) = \int_0^b h'(x)\,dx \leq \frac{b}{12} - \frac{1}{288},$$

$$h\left(\frac{1}{4}\right) = h(b) + \int_b^{1/4} h'(x)\,dx \leq \frac{5}{288} + \frac{1}{2} \cdot \left(\frac{1}{4} - b\right)^2.$$

Then from

$$\frac{7}{96} - \frac{\sqrt{3}}{36} < h\left(\frac{1}{4}\right) \leq \frac{5}{288} + \frac{1}{2} \cdot \left(\frac{1}{4} - b\right)^2,$$

we deduce that $b < b_0 := (5 - 2\sqrt{3})/12 < \frac{1}{4}$, and hence $h(b) < b_0/12 - 1/288$. It follows that

$$h\left(\frac{1}{4}\right) - 2h(b) > \frac{7}{96} - \frac{\sqrt{3}}{36} - \frac{b_0}{6} + \frac{1}{144} = \frac{1}{96}.$$

Then the conclusion follows by choosing $c > b$ sufficiently close to b. $\qquad\square$

Motivated by Theorem 3.8, let us end this subsection with the following question that might be of interest.

Question. What is the optimal lower bound of $\kappa > 0$ such that if $f \in \mathcal{B}$ and $f(0) - f(\frac{1}{4}) > \kappa \cdot \eta(f)$ holds, then $x \mapsto f(x - \omega)$ satisfies the Sturmian condition for any $\omega \in \mathbb{R}$?

REFERENCES

[1] V. Anagnostopoulou, K. Díaz-Ordaz, O. Jenkinson and C. Richard. Sturmian maximizing measures for the piecewise-linear cosine family. *Bull. Braz. Math. Soc. (N.S.)* **43**(2) (2012), 285–302.

[2] J. Bochi. Ergodic optimization of Birkhoff averages and Lyapunov exponents. *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018 (Invited Lectures, III)*. Eds. B. Sirakov, P. Ney de Souza and M. Viana. World Scientific Publishing, Hackensack, NJ, 2018, pp. 1825–1846.

[3] T. Bousch. Le poisson n'a pas d'arêtes. *Ann. Inst. Henri Poincaré Probab. Stat.* **36**(4) (2000), 489–508.

[4] T. Bousch and O. Jenkinson. Cohomology classes of dynamically non-negative C^k functions. *Invent. Math.* **148**(1) (2002), 207–217.

[5]   S. Bullett and P. Sentenac. Ordered orbits of the shift, square roots, and the devil's staircase. *Math. Proc. Cambridge Philos. Soc.* **115**(3) (1994), 451–481.

[6]   G. Contreras, A. O. Lopes and P. Thieullen. Lyapunov minimizing measures for expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **21**(5) (2001), 1379–1409.

[7]   A. Fan, J. Schmeling and W. Shen. L^{∞}-estimation of generalized Thue–Morse trigonometric polynomials and ergodic maximization. *Discrete Contin. Dyn. Syst.* **41**(1) (2021), 297–327.

[8]   N. P. Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics (Lecture Notes in Mathematics, 1794)*. Eds. V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. Springer-Verlag, Berlin, 2002.

[9]   E. Garibaldi. *Ergodic Optimization in the Expanding Case: Concepts, Tools and Applications (Springer-Briefs in Mathematics)*. Springer, Cham, 2017.

[10] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst.* **15**(1) (2006), 197–224.

[11] O. Jenkinson. Ergodic optimization in dynamical systems. *Ergod. Th. & Dynam. Sys.* **39**(10) (2019), 2593–2618.

[12] S. V. Savchenko. Homological inequalities for finite topological Markov chains. *Funktsional. Anal. i Prilozhen.* **33**(3) (1999), 91–93.