

Majorization of invariant measures for orientation-reversing maps

OLIVER JENKINSON and JACOB STEEL

School of Mathematical Sciences, Queen Mary, University of London, Mile End Road,
London E1 4NS, UK

(e-mail: omj@maths.qmul.ac.uk, jacobfdsteel@gmail.com)

(Received 28 September 2008 and accepted in revised form 23 June 2009)

Abstract. Let the invariant probability measures for an orientation-reversing weakly expanding map of the interval $[0, 1]$ be partially ordered by majorization. The minimal elements of the resulting poset are shown to be convex combinations of Dirac measures supported on two adjacent fixed points. A consequence is that if $f : [0, 1] \rightarrow \mathbb{R}$ is strictly convex, then either its minimizing measure is unique and is a Dirac measure on a fixed point, or f has precisely two ergodic minimizing measures, namely Dirac measures on two adjacent fixed points. In the case where $\{0, 1\}$ is a period-two orbit, with corresponding invariant measure μ_{01} , the maximal elements of the poset are shown to be convex combinations of μ_{01} with the Dirac measure on either the leftmost, or the rightmost, fixed point. This facilitates the identification of f -maximizing measures when $f : [0, 1] \rightarrow \mathbb{R}$ is convex.

1. Introduction

Majorization is a way of making precise the notion that one measure is more *spread out* than another (see e.g. [3, 5, 6, 9, 19–22]). If μ and ν are Borel probability measures on the unit interval $[0, 1]$, we say that ν *majorizes* μ , and write $\mu < \nu$, if $\mu(f) \leq \nu(f)$ for all convex functions $f : [0, 1] \rightarrow \mathbb{R}$. Equivalently, $\mu < \nu$ if and only if

$$\int_0^t \mu[0, x] dx \leq \int_0^t \nu[0, x] dx \quad \text{for all } t \in [0, 1] \quad (1)$$

and

$$\int_0^1 \mu[0, x] dx = \int_0^1 \nu[0, x] dx.$$

If $T : [0, 1] \rightarrow [0, 1]$ is Borel, its set \mathcal{M}_T of invariant Borel probability measures becomes a partially ordered set when equipped with $<$. For the doubling map $T(x) = 2x \pmod{1}$, the poset $(\mathcal{M}_T, <)$ was investigated in [15] (see also [14, 16, 24]), where its minimal and maximal elements were identified.

THEOREM 1.1. *Let $T(x) = 2x \pmod{1}$ for $x \in [0, 1)$, and $T(1) = 1$. The minimal elements of (\mathcal{M}_T, \prec) are precisely the Sturmian[†] measures. The maximal elements of (\mathcal{M}_T, \prec) are precisely the convex combinations of the Dirac measures at the two fixed points 0 and 1.*

In this article we consider maps such as the reverse doubling map, defined by $T(x) = -2x \pmod{1}$ for $x \in (0, 1]$, and $T(0) = 1$. As for the doubling map, the invariant measures for the reverse doubling map are naturally identified with those for the full shift on two symbols; in particular, both maps have precisely two fixed points. The role of the fixed points for the reverse doubling map turns out to be the reverse of their role for the doubling map.

THEOREM 1.2. *Let $T(x) = -2x \pmod{1}$ for $x \in (0, 1]$, and $T(0) = 1$. The minimal elements in (\mathcal{M}_T, \prec) are precisely the convex combinations of the Dirac measures at the two fixed points $1/3$ and $2/3$.*

Perhaps surprisingly, and in contrast to Theorem 1.1, the statement of Theorem 1.2 is robust under nonlinear perturbation. It is a particular case of the following result.

THEOREM 1.3. *If $T : [0, 1] \rightarrow [0, 1]$ is the lift of a continuous orientation-reversing expanding circle map, with fixed points $x_1 < \dots < x_k$, then the minimal elements of (\mathcal{M}_T, \prec) are precisely those measures of the form $\lambda\delta_{x_i} + (1 - \lambda)\delta_{x_{i+1}}$ for some $\lambda \in [0, 1]$, $1 \leq i \leq k - 1$.*

For T the reverse doubling map, the maximal elements of (\mathcal{M}_T, \prec) are identified as follows.

THEOREM 1.4. *Let $T(x) = -2x \pmod{1}$ for $x \in (0, 1]$, and $T(0) = 1$. Let $\mu_{01} = (\delta_0 + \delta_1)/2$, the invariant measure supported by the period-two orbit $\{0, 1\}$.*

The maximal elements in (\mathcal{M}_T, \prec) are precisely the convex combinations of μ_{01} with $\delta_{1/3}$, and the convex combinations of μ_{01} with $\delta_{2/3}$.

Theorem 1.4 can be generalized as follows.

THEOREM 1.5. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is the lift of a continuous orientation-reversing expanding circle map, with $T(0) = 1$ and $T(1) = 0$, and with fixed points $x_1 < \dots < x_k$. The maximal elements of (\mathcal{M}_T, \prec) are precisely those measures of the form either $\lambda\mu_{01} + (1 - \lambda)\delta_{x_1}$ or $\lambda\mu_{01} + (1 - \lambda)\delta_{x_k}$, for some $\lambda \in [0, 1]$, where $\mu_{01} := (\delta_0 + \delta_1)/2$.*

A primary motivation for the above results is *ergodic optimization* (see, e.g., [4, 7, 8, 13]), that is, the study of the smallest possible ergodic average $\alpha(f) = \inf_{\mu \in \mathcal{M}_T} \mu(f)$ and largest possible ergodic average $\beta(f) = \sup_{\mu \in \mathcal{M}_T} \mu(f)$ of a given function $f : [0, 1] \rightarrow \mathbb{R}$, and of the invariant measures μ (so-called *minimizing* and *maximizing* measures,

[†] Sturmian measures are symbolic versions of rotations (see, e.g., [1, 2, 4, 12, 14, 15] for further details); in particular they are ergodic, supported by either a periodic orbit (in the case of a rational rotation) or a uniquely ergodic Cantor set.

cf. Definition 2.2) for which $\mu(f)$ attains these extrema. When $f : [0, 1] \rightarrow \mathbb{R}$ is convex[†], we solve the ergodic optimization problem as follows.

THEOREM 1.6. *Suppose that $T : [0, 1] \rightarrow [0, 1]$ is the lift of a continuous orientation-reversing expanding circle map, with $T(0) = 1$ and $T(1) = 0$, and with fixed points $x_1 < \cdots < x_k$. If $f : [0, 1] \rightarrow \mathbb{R}$ is convex, then its minimum ergodic average is*

$$\alpha(f) = \min_{1 \leq i \leq k} f(x_i),$$

and its maximum ergodic average is

$$\beta(f) = \max\left(f(x_1), f(x_k), \frac{f(0) + f(1)}{2}\right).$$

There exists $1 \leq i \leq k$ such that δ_{x_i} is an f -minimizing measure, and at least one of the measures $\delta_{x_1}, \mu_{01}, \delta_{x_k}$ is f -maximizing.

This article is organized as follows. In §2 we introduce the class of *orientation-reversing weakly expanding maps*, a suitable generalization of the maps T considered above. For orientation-reversing weakly expanding maps T we identify (see Theorem 3.1 in §3) the minimal elements of $(\mathcal{M}_T, <)$, a result which implies Theorem 1.3. Under the additional hypothesis that $\{0, 1\}$ is a period-two orbit, we then identify (see Theorem 4.1 in §4) the maximal elements of $(\mathcal{M}_T, <)$, a result which implies Theorem 1.5. Theorem 1.6 is a consequence of Theorems 3.4 and 4.4, which also give more precise information in the case where $f : [0, 1] \rightarrow \mathbb{R}$ is *strictly* convex. In §5, some of the fine structure of $(\mathcal{M}_T, <)$ is computed explicitly in the case where T is the reverse doubling map. In §6 we consider the extension of our results to interval maps with infinitely many branches, with Gauss's continued fraction map serving as an illustrative example.

2. Preliminaries

Notation 2.1. For a Borel probability measure μ on $[0, 1]$, let $b(\mu) := \int x d\mu(x)$ denote its *barycentre*. If $T : [0, 1] \rightarrow [0, 1]$ is Borel, let \mathcal{M}_T denote the set of T -invariant Borel probability measures. For $\varrho \in [0, 1]$, define the corresponding *barycentre class* $\mathcal{M}_\varrho := \{\mu \in \mathcal{M}_T \mid b(\mu) = \varrho\}$.

Definition 2.2. Suppose that $f : [0, 1] \rightarrow \mathbb{R}$ is bounded and Borel measurable, and that $T : [0, 1] \rightarrow [0, 1]$ is Borel. For $\varrho \in [0, 1]$, a measure $\mu \in \mathcal{M}_\varrho$ is called *f -minimizing in \mathcal{M}_ϱ* (respectively *f -maximizing in \mathcal{M}_ϱ*) if $\mu(f) = \inf_{m \in \mathcal{M}_\varrho} m(f)$ (respectively $\mu(f) = \sup_{m \in \mathcal{M}_\varrho} m(f)$).

Define the *minimum ergodic average*

$$\alpha(f) := \inf_{m \in \mathcal{M}_T} m(f),$$

and the *maximum ergodic average*

$$\beta(f) := \sup_{m \in \mathcal{M}_T} m(f).$$

[†] Clearly, the solution of the ergodic optimization problem for *concave* f follows immediately, since in this case $-f$ is convex.

A measure $\mu \in \mathcal{M}_T$ is called (globally) f -minimizing if $\mu(f) = \alpha(f)$, and (globally) f -maximizing if $\mu(f) = \beta(f)$.

Definition 2.3. Suppose $0 = a_0 < a_1 < \dots < a_k = 1$, where $k \geq 2$. Let J_1, \dots, J_k be disjoint sub-intervals of $[0, 1]$, with $\bigcup_{i=1}^k J_i = [0, 1]$, such that the left (respectively right) endpoint of each J_i is a_{i-1} (respectively a_i). Suppose that, for each $1 \leq i \leq k$, the restriction $T|_{J_i}$ is continuous. Suppose that $0 \in \overline{T(J_1)}$, that $\overline{T(J_i)} = [0, 1]$ for $1 < i < k$, and that $1 \in \overline{T(J_k)}$.

We say that T is *orientation reversing* if $T|_{J_i}$ is decreasing for each $1 \leq i \leq k$, and that T is *weakly expanding* if $|T(x) - T(y)| \geq |x - y|$ for all $x, y \in J_i$, $1 \leq i \leq k$.

For orientation-reversing weakly expanding maps, not every point in $[0, 1]$ is the barycentre of an invariant measure.

LEMMA 2.4. *If $T : [0, 1] \rightarrow [0, 1]$ is an orientation-reversing weakly expanding map, with fixed points $x_1 < \dots < x_k$, then its barycentre set $b(\mathcal{M}_T)$ is the interval $[x_1, x_k]$.*

Proof. The set $b(\mathcal{M}_T)$ is an interval, since it is the affine image of the convex space \mathcal{M}_T . Since $x_1 = b(\delta_{x_1})$ and $x_k = b(\delta_{x_k})$, it remains to show that $x_1 \leq b(\mu) \leq x_k$ for every $\mu \in \mathcal{M}_T$.

Define $\sigma : [0, 1] \rightarrow \mathbb{R}$ by $\sigma(x) = \min\{x, x_k\}$. If $x > x_k$, then $T(x) < x_k$, and $x - x_k \leq x_k - T(x)$ because T is weakly expanding, so $x - \sigma(x) + \sigma(Tx) = x - x_k + T(x) \leq x_k$. If $x \leq x_k$, then $x - \sigma(x) + \sigma(Tx) = \sigma(Tx) \leq \max \sigma = x_k$. Thus, if $\mu \in \mathcal{M}_T$ then $b(\mu) = \int x d\mu(x) = \int (x - \sigma(x) + \sigma(Tx)) d\mu(x) \leq x_k$.

Defining $\tau(x) = \max\{x_1, x\}$, a similar argument shows that $x - \tau(x) + \tau(Tx) \geq x_1$ for all $x \in [0, 1]$, hence that $b(\mu) = \int x d\mu(x) = \int (x - \tau(x) + \tau(Tx)) d\mu(x) \geq x_1$ for all $\mu \in \mathcal{M}_T$. \square

A necessary condition for two measures to be related by majorization is that they share the same barycentre (since $f(x) := x$ and $g(x) := -x$ are both convex). Consequently the poset (\mathcal{M}_T, \prec) is the disjoint union of the barycentre classes $(\mathcal{M}_\varrho, \prec)$, $\varrho \in [x_1, x_k]$. For $\varrho \in [x_1, x_k]$, a natural first question about the barycentre class $(\mathcal{M}_\varrho, \prec)$ is whether it has a smallest and a largest element. In §§3 and 4 this question is answered affirmatively[†], with the following measures m_ϱ and ν_ϱ identified as, respectively, the smallest and largest elements of $(\mathcal{M}_\varrho, \prec)$.

Notation 2.5. (The measures m_ϱ and ν_ϱ) Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \dots < x_k$. For each $\varrho \in [x_1, x_k]$, define

$$m_\varrho := \lambda \delta_{x_i} + (1 - \lambda) \delta_{x_{i+1}},$$

where $\lambda \in [0, 1]$ and $1 \leq i \leq k - 1$ are such that $\varrho = \lambda x_i + (1 - \lambda) x_{i+1}$.

[†] We emphasize that there are (well-known) interval maps T for which the barycentre classes $(\mathcal{M}_\varrho, \prec)$ do not have smallest or largest elements, see [24].

If furthermore $T(0) = 1$ and $T(1) = 0$, with $\mu_{01} := (\delta_0 + \delta_1)/2$, then define ν_ϱ by

$$\nu_\varrho := \begin{cases} \frac{1/2 - \varrho}{1/2 - x_1} \delta_{x_1} + \frac{\varrho - x_1}{1/2 - x_1} \mu_{01} & \text{for } \varrho \in [x_1, 1/2), \\ \mu_{01} & \text{for } \varrho = 1/2, \\ \frac{\varrho - 1/2}{x_k - 1/2} \delta_{x_k} + \frac{x_k - \varrho}{x_k - 1/2} \mu_{01} & \text{for } \varrho \in (1/2, x_k]. \end{cases}$$

3. Minimal elements of $(\mathcal{M}_T, <)$, and minimizing measures for convex f

Each barycentre class has a smallest element, which can be identified explicitly.

THEOREM 3.1. *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \dots < x_k$. For every $\varrho \in [x_1, x_k]$, the ordered set $(\mathcal{M}_\varrho, <)$ has a smallest element, namely m_ϱ .*

Proof. For every convex function $g : [0, 1] \rightarrow \mathbb{R}$, we must show that $m_\varrho(g) \leq \mu(g)$ for all $\mu \in \mathcal{M}_\varrho$. Note that this inequality holds if and only if

$$m_\varrho(f) \leq \mu(f) \quad \text{for all } \mu \in \mathcal{M}_\varrho, \quad (2)$$

where

$$f(x) := g(x) + \frac{g(x_{i+1}) - g(x_i)}{x_i - x_{i+1}} x,$$

and $1 \leq i \leq k - 1$ is chosen such that $x_i \leq \varrho \leq x_{i+1}$.

We now claim that all measures in $\mathcal{N} = \{\varepsilon \delta_{x_i} + (1 - \varepsilon) \delta_{x_{i+1}} \mid \varepsilon \in [0, 1]\}$ are f -minimizing. This implies that $m_\varrho(f) \leq \mu(f)$ for all $\mu \in \mathcal{M}_T$, which in particular implies (2).

Note that $f(x_i) = f(x_{i+1})$, and denote this common value by α . Define $\varphi : [0, 1] \rightarrow \mathbb{R}$ to be equal to $-\alpha$ on (x_i, x_{i+1}) , and by $\varphi(x) = -f(x)$ otherwise. We claim that, for all $x \in [0, 1]$,

$$(f + \varphi - \varphi \circ T)(x) \geq \alpha. \quad (3)$$

If $x \notin (x_i, x_{i+1})$, then $(f + \varphi - \varphi \circ T)(x) = -\varphi(Tx) \geq -\max_{y \in [0, 1]} \varphi(y) = \alpha$, so (3) holds.

If $x \in (x_i, x_{i+1})$, then $T(x) \notin (x_i, x_{i+1})$, so

$$(f + \varphi - \varphi \circ T)(x) = f(x) - \alpha + f(Tx). \quad (4)$$

Note that $x \in J_i \cup J_{i+1}$. If $x \in J_i$, then $T(x) < x_i < x$, and weak expansion gives

$$x_i - T(x) \geq x - x_i. \quad (5)$$

Since $f(x_i) = f(x_{i+1})$, the convexity of f implies it is non-increasing on $[0, x_i]$, so $f(x_i) - f(Tx) \leq 0$. Combining with (5) gives

$$f(x_i) - f(Tx) = \frac{f(x_i) - f(Tx)}{x_i - T(x)}(x_i - T(x)) \leq \frac{f(x_i) - f(Tx)}{x_i - T(x)}(x - x_i). \quad (6)$$

However, f is convex, so (see, e.g., [23, p. 113])

$$\frac{f(x_i) - f(Tx)}{x_i - T(x)} \leq \frac{f(x) - f(x_i)}{x - x_i}. \quad (7)$$

Combining (6) with (7) gives $f(x_i) - f(Tx) \leq f(x) - f(x_i)$, and substituting into (4) gives (3).

The argument when $x \in J_{i+1}$ is similar. In this case $x < x_{i+1} < T(x)$, and weak expansion gives

$$T(x) - x_{i+1} \geq x_{i+1} - x. \quad (8)$$

Since $f(x_i) = f(x_{i+1})$, the convexity of f implies that it is non-decreasing on $[x_{i+1}, 1]$, so $f(Tx) - f(x_{i+1}) \geq 0$. Combining this with (8), and the fact that

$$\frac{f(x_{i+1}) - f(x)}{x_{i+1} - x} \leq \frac{f(Tx) - f(x_{i+1})}{T(x) - x_{i+1}},$$

we deduce that $f(Tx) - f(x_{i+1}) \geq f(x_{i+1}) - f(x)$, and again (3) follows from (4).

Having established (3) for all $x \in [0, 1]$, integration of this inequality with respect to an arbitrary $\mu \in \mathcal{M}_T$ gives $\mu(f) \geq \alpha = m(f)$ for every $m \in \mathcal{N}$, so indeed every measure in \mathcal{N} is f -minimizing. \square

Theorem 3.1 asserts that if $f : [0, 1] \rightarrow \mathbb{R}$ is convex, then the measure m_ϱ is minimizing in \mathcal{M}_ϱ (cf. Definition 2.2). If f is *strictly* convex, then this conclusion can be strengthened as follows.

COROLLARY 3.2. *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \dots < x_k$. If $f : [0, 1] \rightarrow \mathbb{R}$ is strictly convex, and $\varrho \in [x_1, x_k]$, then m_ϱ is the unique f -minimizing measure in \mathcal{M}_ϱ .*

Proof. An alternative definition of majorization (see [3, 5]) is that $\mu < \nu$ if and only if ν is a *dilation* of μ , that is, there exists a family of probability measures $(D_x)_{x \in [0, 1]}$, with each $b(D_x) = x$, such that if $f : [0, 1] \rightarrow \mathbb{R}$ is bounded and Borel, then so is $x \mapsto D_x(f)$, and $\nu(f) = \int D_x(f) d\mu(x)$.

If $\mu \in \mathcal{M}_\varrho \setminus \{m_\varrho\}$, then $m_\varrho < \mu$ by Theorem 3.1, so there exists $(D_x)_{x \in [0, 1]}$ as above, such that $\mu(f) = \int D_x(f) dm_\varrho(x)$ for any bounded Borel f . Now $\mu \neq m_\varrho$, so there is a Borel set A , with $m_\varrho(A) > 0$, such that $D_x \neq \delta_x$ for all $x \in A$. Since Jensen's inequality is strict, that is $D_x(f) > f(x)$, whenever f is strictly convex and D_x is not the Dirac measure at x , we deduce that $\mu(f) > m_\varrho(f)$. \square

In particular, the *variance* $\text{var}(\mu) = \int (x - b(\mu))^2 d\mu(x)$ around the mean ϱ is minimized precisely when $\mu = m_\varrho$.

COROLLARY 3.3. *For every $\varrho \in [x_1, x_k]$, the measure m_ϱ is the unique measure with smallest variance in \mathcal{M}_ϱ .*

For convex $f : [0, 1] \rightarrow \mathbb{R}$ we are now able to identify the minimum ergodic average $\alpha(f)$, and deduce information about (globally) minimizing measures.

THEOREM 3.4. *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \dots < x_k$. If $f : [0, 1] \rightarrow \mathbb{R}$ is convex, then its minimum ergodic average is*

$$\alpha(f) = \min_{1 \leq i \leq k} f(x_i),$$

and there exists $1 \leq i \leq k$ such that δ_{x_i} is f -minimizing.

01 If $f : [0, 1] \rightarrow \mathbb{R}$ is strictly convex, then either there is a unique $1 \leq i \leq k$ such that
 02 $f(x_i) = \alpha(f)$, in which case δ_{x_i} is the unique f -minimizing measure, or there exists
 03 $1 \leq i \leq k - 1$ such that $f(x_i) = f(x_{i+1}) = \alpha(f)$, in which case the set of f -minimizing
 04 measures is $\{\lambda\delta_{x_i} + (1 - \lambda)\delta_{x_{i+1}} \mid \lambda \in [0, 1]\}$.

05 *Proof.* The map $F : \varrho \mapsto \inf_{\mu \in \mathcal{M}_\varrho} \mu(f)$ is clearly convex on $[x_1, x_k]$. Since $m_\varrho(f) =$
 06 $\inf_{\mu \in \mathcal{M}_\varrho} \mu(f)$ by Theorem 3.1, the map F is affine on each interval $[x_i, x_{i+1}]$, $1 \leq i \leq$
 07 $k - 1$. Consequently, there exists $1 \leq i \leq k$ such that $f(x_i) = \min_{\varrho \in [x_1, x_k]} F(\varrho) = \alpha(f)$,
 08 and therefore δ_{x_i} is f -minimizing.

09 Now suppose that f is strictly convex. If there is a unique $1 \leq i \leq k$ such that
 10 $f(x_i) = \min_{\varrho \in [x_1, x_k]} F(\varrho) = \alpha(f)$, then the f -minimizing measures are precisely those
 11 which are minimizing in \mathcal{M}_{x_i} ; Corollary 3.2 then implies that δ_{x_i} is the unique invariant
 12 measure which is minimizing in \mathcal{M}_{x_i} , hence it is the unique f -minimizing measure. If not,
 13 then the strict convexity of f implies that $f(x_i) = f(x_{i+1}) = \min_{\varrho \in [x_1, x_k]} F(\varrho) = \alpha(f)$
 14 for some $1 \leq i \leq k - 1$, with $f(x_j) > \alpha(f)$ for $j \in \{1, \dots, k\} \setminus \{i, i + 1\}$. So F attains
 15 its minimum value $\alpha(f)$ precisely on the interval $[x_i, x_{i+1}]$, hence an invariant measure is
 16 f -minimizing if and only if it is minimizing in \mathcal{M}_ϱ for some $\varrho \in [x_i, x_{i+1}]$. Corollary 3.2
 17 then implies that the set of f -minimizing measures is precisely $\{\lambda\delta_{x_i} + (1 - \lambda)\delta_{x_{i+1}} \mid \lambda \in$
 18 $[0, 1]\}$. \square

19 When T is the reverse doubling map, the result is particularly explicit.

20
 21 COROLLARY 3.5. Let $T(x) = -2x \pmod{1}$ for $x \in (0, 1)$, and $T(0) = 1$. If $f : [0, 1] \rightarrow$
 22 \mathbb{R} is strictly convex, then its minimum ergodic average is

$$23 \quad \alpha(f) = \min(f(1/3), f(2/3)).$$

24
 25 If $f(1/3) < f(2/3)$, then the unique f -minimizing measure is $\delta_{1/3}$, while if $f(1/3) >$
 26 $f(2/3)$, then the unique f -minimizing measure is $\delta_{2/3}$. If $f(1/3) = f(2/3)$, then the set
 27 of f -minimizing measures is precisely the convex hull of $\{\delta_{1/3}, \delta_{2/3}\}$.

28 29 4. Maximal elements of (\mathcal{M}_T, \prec) , and maximizing measures for convex f

30 We prove the following stronger version of Theorem 1.5.

31 THEOREM 4.1. Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding
 32 map, with fixed points $x_1 < \dots < x_k$, where $T(0) = 1$ and $T(1) = 0$. For every $\varrho \in$
 33 $[x_1, x_k]$, the ordered set $(\mathcal{M}_\varrho, \prec)$ has a largest element, namely v_ϱ .

34
 35 *Proof.* First suppose that $\varrho = 1/2$. In this case $v_\varrho = \mu_{01} = (\delta_0 + \delta_1)/2$ is the largest
 36 element not only in $(\mathcal{M}_{1/2}, \prec)$, but also in the larger set of all (not necessarily invariant)
 37 Borel probability measures with barycentre equal to $1/2$. To see this, note that if
 38 $b(\mu) = 1/2$, then $\mu_{01} = \int (x\delta_1 + (1 - x)\delta_0) d\mu(x)$. Therefore, μ_{01} is a dilation of μ
 39 (cf. [3, 5] and the proof of Corollary 3.2), hence $\mu \prec \mu_{01}$.

40 Now suppose that $\varrho \in [x_1, x_k] \setminus \{1/2\}$. We first prove the result in the case where
 41 $x_1 < 1/2 < x_k$.

42 It suffices to consider the case $\varrho \in [x_1, 1/2)$; the proof for $\varrho \in (1/2, x_k]$ is almost
 43 identical, and is omitted. For all convex functions $f : [0, 1] \rightarrow \mathbb{R}$, we wish to show that

$$44 \quad \mu(f) \leq v_\varrho(f) \quad \text{for all } \mu \in \mathcal{M}_\varrho. \quad (9)$$

01 We may assume that $f(0) = 0 = f(1)$ (since $\tilde{f}(x) := f(x) + (f(0) - f(1))x - f(0)$
 02 satisfies $\tilde{f}(0) = 0 = \tilde{f}(1)$, and (9) holds if and only if $\mu(\tilde{f}) \leq \nu_\varrho(\tilde{f})$ for all $\mu \in \mathcal{M}_\varrho$).
 03 So, since

$$04 \quad \nu_\varrho = \frac{1/2 - \varrho}{1/2 - x_1} \delta_{x_1} + \frac{\varrho - x_1}{1/2 - x_1} \mu_{01},$$

06 we wish to show that

$$07 \quad \mu(f) \leq \frac{1/2 - \varrho}{1/2 - x_1} f(x_1) \quad \text{for all } \mu \in \mathcal{M}_\varrho. \quad (10)$$

10 Now (10) is implied by the existence of a continuous function $\varphi : [0, 1] \rightarrow \mathbb{R}$ satisfying

$$11 \quad (f + \varphi - \varphi \circ T)(x) \leq \frac{1/2 - x}{1/2 - x_1} f(x_1) \quad \text{for all } x \in [0, 1], \quad (11)$$

14 and if we define $\hat{f} : [0, 1] \rightarrow \mathbb{R}$ by

$$15 \quad \hat{f}(x) = f(x) + \frac{f(x_1)}{1/2 - x_1} x,$$

18 then (11) becomes

$$19 \quad (\hat{f} + \varphi - \varphi \circ T)(x) \leq \frac{1/2}{1/2 - x_1} f(x_1) = \hat{f}(x_1) \quad \text{for all } x \in [0, 1]. \quad (12)$$

22 Defining φ by

$$23 \quad \varphi(x) = \begin{cases} -\hat{f}(x_1) & \text{for } x \in [0, x_1], \\ -\hat{f}(x) & \text{for } x \in [x_1, 1], \end{cases} \quad (13)$$

25 we verify that (12) holds[†]. If $x \in [x_1, 1]$, then $(\hat{f} + \varphi - \varphi \circ T)(x) = -\varphi(Tx) = \hat{f}(x_1)$,
 26 while if $x \in [0, x_1]$, then $(\hat{f} + \varphi - \varphi \circ T)(x) = \hat{f}(x) - \hat{f}(x_1) + \hat{f}(Tx)$, so (12) holds if
 27 and only if

$$28 \quad \hat{f}(x) + \hat{f}(Tx) \leq 2\hat{f}(x_1) \quad \text{for all } x \in [0, x_1]. \quad (14)$$

30 To prove (14), first note that the convexity of \hat{f} implies

$$31 \quad \hat{f}(x) \leq \frac{x}{x_1} \hat{f}(x_1) + \left(1 - \frac{x}{x_1}\right) \hat{f}(0) = \frac{x}{x_1} \hat{f}(x_1) \quad \text{for all } x \in [0, x_1], \quad (15)$$

34 and

$$35 \quad \hat{f}(y) \leq \frac{1 - y}{1 - x_1} \hat{f}(x_1) + \frac{y - x_1}{1 - x_1} \hat{f}(1) \quad \text{for all } y \in [x_1, 1]. \quad (16)$$

37 Now T is weakly expanding, so if $x \in [0, x_1]$, then $T(x) - x_1 \geq x_1 - x$, or in other
 38 words

$$39 \quad T(x) \geq 2x_1 - x. \quad (17)$$

40 However, $\hat{f} : [0, 1] \rightarrow \mathbb{R}$ is decreasing, since it is a convex function with

$$42 \quad \hat{f}(1) = 2\hat{f}(x_1) = \frac{f(x_1)}{1/2 - x_1} < 0 = \hat{f}(0), \quad (18)$$

44 [†] Note in particular that (12) implies that δ_{x_1} is (globally) \hat{f} -maximizing.

01 so (17) implies that

$$02 \quad \hat{f}(Tx) \leq \hat{f}(2x_1 - x). \quad (19)$$

03
04 Setting $y = 2x_1 - x$ in (16), and combining with (19), gives

$$05 \quad \hat{f}(Tx) \leq \frac{1+x-2x_1}{1-x_1} \hat{f}(x_1) + \frac{x_1-x}{1-x_1} \hat{f}(1). \quad (20)$$

06
07
08 Using (15) with (20), and the fact that $\hat{f}(1) = 2\hat{f}(x_1)$, we obtain

$$09 \quad \hat{f}(x) + \hat{f}(Tx) \leq \hat{f}(x_1) \left(\frac{x}{x_1} + \frac{1-x}{1-x_1} \right) \quad \text{for all } x \in [0, x_1]. \quad (21)$$

10
11
12 Now $x_1 < 1/2$, so the function $x \mapsto x/x_1 + (1-x)/(1-x_1)$ is increasing on $[0, x_1]$,
13 attaining its maximum value of two when $x = x_1$. Therefore, (21) implies the desired
14 inequality (14), and the proof in the case $x_1 < 1/2 < x_k$ is complete.

15 Now suppose that $x_1 < 1/2 < x_k$ does not hold. In this case, because T is weakly
16 expanding, either $x_1 < 1/2 = x_k$ or $x_1 = 1/2 < x_k$. In the former case $b(\mathcal{M}_T) = [x_1, 1/2]$,
17 so we need only consider $\varrho \in [x_1, 1/2)$, and the proof is as above. In the latter case
18 $b(\mathcal{M}_T) = [1/2, x_k]$, so we need only consider $\varrho \in (1/2, x_k]$, and, as noted above, the proof
19 is almost identical to the argument for $\varrho \in [x_1, 1/2)$. \square

20
21 The proofs of the following two corollaries are similar to those of Corollaries 3.2
22 and 3.3, so are omitted.

23 **COROLLARY 4.2.** *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding*
24 *map, with fixed points $x_1 < \dots < x_k$, where $T(0) = 1$ and $T(1) = 0$. If $f : [0, 1] \rightarrow \mathbb{R}$*
25 *is strictly convex, and $\varrho \in [x_1, x_k]$, then ν_ϱ is the unique maximizing measure in \mathcal{M}_ϱ .*

26
27 **COROLLARY 4.3.** *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding*
28 *map, with fixed points $x_1 < \dots < x_k$, where $T(0) = 1$ and $T(1) = 0$. For every $\varrho \in$*
29 *$[x_1, x_k]$, the measure ν_ϱ is the unique measure with largest variance in \mathcal{M}_ϱ .*

30
31 For convex functions $f : [0, 1] \rightarrow \mathbb{R}$, the maximum ergodic average $\beta(f)$ is determined
32 by simply evaluating f at the points 0, 1, x_1 , and x_k .

33 **THEOREM 4.4.** *Let $T : [0, 1] \rightarrow [0, 1]$ be an orientation-reversing weakly expanding*
34 *map, with fixed points $x_1 < \dots < x_k$, where $T(0) = 1$ and $T(1) = 0$. If $f : [0, 1] \rightarrow \mathbb{R}$*
35 *is convex, then*

$$36 \quad \beta(f) = \max(f(x_1), f(x_k), (f(0) + f(1))/2),$$

37
38 and at least one of the measures δ_{x_1} , μ_{01} , and δ_{x_k} is f -maximizing.

39 If $f : [0, 1] \rightarrow \mathbb{R}$ is strictly convex, then one of the following five possibilities holds:

- 40 (i) δ_{x_1} is the unique f -maximizing measure;
41 (ii) μ_{01} is the unique f -maximizing measure;
42 (iii) δ_{x_k} is the unique f -maximizing measure;
43 (iv) the set of f -maximizing measures is $\{\lambda\delta_{x_1} + (1-\lambda)\mu_{01} \mid \lambda \in [0, 1]\}$;
44 (v) the set of f -maximizing measures is $\{\lambda\mu_{01} + (1-\lambda)\delta_{x_k} \mid \lambda \in [0, 1]\}$.

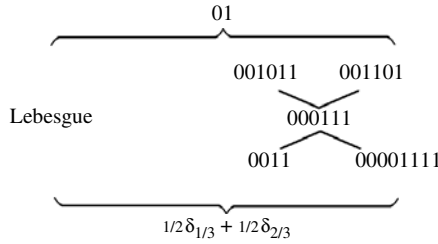


FIGURE 1. Hasse diagram for a portion of $(\mathcal{M}_{1/2}, <) \subset (\mathcal{M}_T, <)$, where T is the reverse doubling map. Symbolic codes denote the corresponding periodic orbit measures. The least element in $(\mathcal{M}_{1/2}, <)$ is the non-ergodic measure $(\delta_{1/3} + \delta_{2/3})/2$, while the largest element is the invariant measure supported by the period-two orbit $\{0, 1\}$.

Proof. The map $G : \varrho \mapsto \sup_{\mu \in \mathcal{M}_\varrho} \mu(f)$ is easily seen to be concave on $[x_1, x_k]$, and Theorem 4.1 implies that G is affine on both $[x_1, 1/2]$ and $[1/2, x_k]$. Consequently, G attains its maximum value at x_1 , $1/2$, or x_k , so at least one of the measures δ_{x_1} , μ_{01} , and δ_{x_k} is f -maximizing.

Now suppose that $f : [0, 1] \rightarrow \mathbb{R}$ is strictly convex. If G has a unique maximum then this must be attained at x_1 , $1/2$, or x_k , in which case Corollary 4.2 implies that the unique f -maximizing measure is, respectively, δ_{x_1} , μ_{01} , or δ_{x_k} .

We next show that $G : [x_1, x_k] \rightarrow \mathbb{R}$ cannot be a constant function. If $G \equiv c$, say, then $f(x_1) = f(x_k) = c$; the strict convexity of f then implies that $\min(f(0), f(1)) > c$, hence $G(1/2) = \mu_{01}(f) = (f(0) + f(1))/2 > c$, a contradiction.

There remains the case that G is constant on either $[x_1, 1/2]$ or on $[1/2, x_k]$: here Corollary 4.2 implies that the set of f -maximizing measures is, respectively, either $\{\lambda\delta_{x_1} + (1 - \lambda)\mu_{01} \mid \lambda \in [0, 1]\}$ or $\{\lambda\mu_{01} + (1 - \lambda)\delta_{x_k} \mid \lambda \in [0, 1]\}$. \square

5. Computations

The majorization criterion (1) can be used to compute some of the structure of $(\mathcal{M}_T, <)$. These computations are particularly tractable when μ and ν are purely atomic with finitely many atoms (which are necessarily periodic points for T). If the mass of each atom is rational, then (1) can be re-formulated in terms of a well-known criterion of Hardy *et al.* (see [9, 10, 14, 15]); for example, if $\mu := Q^{-1} \sum_{i=1}^Q \delta_{\mu_i}$ and $\nu := Q^{-1} \sum_{i=1}^Q \delta_{\nu_i}$, with $\mu_1 \leq \dots \leq \mu_Q$ and $\nu_1 \leq \dots \leq \nu_Q$, and $b(\mu) = b(\nu)$, then $\mu < \nu$ if and only if $\sum_{i=1}^n \mu_i \geq \sum_{i=1}^n \nu_i$ for all $1 \leq n \leq Q - 1$.

Now let $T : [0, 1] \rightarrow [0, 1]$ be the reverse doubling map, given by $T(x) = -2x \pmod{1}$ for $x \in (0, 1]$, and $T(0) = 1$. Any invariant measure supported by a single periodic orbit lies in \mathcal{M}_ϱ for some rational $\varrho \in [1/3, 2/3]$; conversely, if $\varrho \in (1/3, 2/3)$ is rational, then infinitely many such periodic orbit measures belong to \mathcal{M}_ϱ . For barycentre $\varrho = 1/2$, the majorization relations between some of the periodic orbit measures in $\mathcal{M}_{1/2}$ are depicted in Figure 1. Here symbolic codes denote the corresponding periodic orbit measures, where as usual the left half-interval is coded by 0, and the right half-interval by 1. For example 0011 denotes the measure $\frac{1}{4} \sum_{i=1}^4 \delta_{i/5}$, supported by the period-four orbit $\{1/5, 3/5, 4/5, 2/5\}$.

6. *Infinitely many branches: the Gauss map*

The results of this paper can be extended, with some modification, to maps $T : [0, 1] \rightarrow [0, 1]$ which are orientation reversing and weakly expanding, but with infinitely many branches. More precisely, the techniques of §3, and to a lesser extent those of §4, can be used to study the majorization structure of \mathcal{M}_T for maps T satisfying the analogue of Definition 2.3 where the finite set $(a_i)_{i \in \{1, \dots, k\}}$ is replaced by a countably infinite set $(a_i)_{i \in \mathcal{I}}$, with each $a_i < a_{i+1}$.

A treatment of general countable branch orientation-reversing weakly expanding maps would involve individual analyses of various sub-cases, according to which of 0 and 1 is an accumulation point of $(a_i)_{i \in \mathcal{I}}$, and the value of T at the accumulation points. Instead of performing this analysis, we fix ideas by concentrating on the most well-known such map, the *Gauss map*, which illustrates the way in which the infinite branch case differs from §§3 and 4.

The Gauss map is most naturally defined as a self-map of the set of irrationals in $(0, 1)$. However, it is usually extended to $(0, 1]$ by the formula

$$T(x) = 1/x \pmod{1},$$

and can be extended to a self-map of $[0, 1]$ by choosing $T(0)$ to be some element of $[0, 1]$. Since 0 cannot be a point of continuity of T , this choice of $T(0)$ is rather arbitrary; for definiteness, and in view of the assumptions in §4, we set $T(0) := 1$.

The fixed points of T are

$$z_i := (\sqrt{i^2 + 4} - i)/2 \quad \text{for } i \geq 1.$$

In particular, $z_1 = (\sqrt{5} - 1)/2 = \max_i z_i$, $z_2 = \sqrt{2} - 1$, and $\inf_i z_i = 0$.

By analogy with Notation 2.5, let $m_\varrho := \lambda \delta_{z_i} + (1 - \lambda) \delta_{z_{i+1}}$, where $\lambda \in [0, 1]$ and $i \geq 1$ are such that $\varrho = \lambda z_i + (1 - \lambda) z_{i+1}$. Proofs analogous to those of Lemma 2.4 and Theorem 3.1 give the following.

THEOREM 6.1. *For $T : [0, 1] \rightarrow [0, 1]$ the Gauss map, $b(\mathcal{M}_T) = (0, (\sqrt{5} - 1)/2]$. For every $\varrho \in (0, (\sqrt{5} - 1)/2]$, the ordered set $(\mathcal{M}_\varrho, <)$ has a smallest element, namely m_ϱ .*

Using the same arguments as in §3, it can also be shown that m_ϱ is the unique f -minimizing measure in \mathcal{M}_ϱ for strictly convex $f : [0, 1] \rightarrow \mathbb{R}$, and in particular that m_ϱ is the unique measure in \mathcal{M}_ϱ with smallest variance.

The following result on global minimization for convex functions is an analogue of Theorem 3.4, the significant difference being that f -minimizing measures need not exist (e.g. this occurs if the convex function f is strictly increasing).

THEOREM 6.2. *Let $T : [0, 1] \rightarrow [0, 1]$ be the Gauss map. If $f : [0, 1] \rightarrow \mathbb{R}$ is convex, then its minimum ergodic average is*

$$\alpha(f) = \inf_{i \geq 1} f(z_i). \tag{22}$$

If the infimum (22) is not attained by any $i \geq 1$, then there are no f -minimizing measures. Otherwise, at least one Dirac measure δ_{z_i} is f -minimizing.

If f is strictly convex, and (22) is attained by a unique $i \geq 1$, then the corresponding Dirac measure δ_{z_i} is the unique f -minimizing measure. If (22) is attained by at least two distinct values $i \geq 1$, then there exists $j \geq 1$ such that the set of f -minimizing measures is $\{\lambda\delta_{z_j} + (1 - \lambda)\delta_{z_{j+1}} \mid \lambda \in [0, 1]\}$.

Example 6.3. Let $T : [0, 1] \rightarrow [0, 1]$ be the Gauss map, and define $f : [0, 1] \rightarrow \mathbb{R}$ by $f(x) = (x - 1/2)^2$. Since f is strictly convex, Theorem 6.2 implies that its minimum ergodic average is

$$\alpha(f) = \inf_i f(z_i) = f(z_2) = (3/2 - \sqrt{2})^2 = \frac{17}{4} - 3\sqrt{2},$$

and that its unique f -minimizing measure is $\delta_{z_2} = \delta_{\sqrt{2}-1}$.

Remark 6.4. We make the following final remarks.

- (a) A completely different approach to ergodic optimization for infinite branch maps, based on symbolic dynamics, has been considered in [11, 17, 18].
- (b) The choice $T(0) = 1$ renders $\{0, 1\}$ a period-two orbit of T , so that $\mu_{01} = (\delta_0 + \delta_1)/2 \in \mathcal{M}_T$. Arguments similar to those of Theorem 4.1 can then be used to show that for $\varrho \in [1/2, (\sqrt{5} - 1)/2]$, the appropriate convex combination of μ_{01} with δ_{z_1} is the largest element in $(\mathcal{M}_\varrho, <)$. For $\varrho \in (0, 1/2)$, however, the absence of a smallest fixed point of T can be used to show that $(\mathcal{M}_\varrho, <)$ has no largest element.

REFERENCES

- [1] V. Anagnostopoulou, K. Díaz-Ordaz, O. Jenkinson and C. Richard. Sturmian maximizing measures for the piecewise-linear cosine family. *Preprint*.
- [2] V. Anagnostopoulou and O. Jenkinson. Which beta-shifts have a largest invariant measure? *J. London Math. Soc.* **79** (2009), 445–464.
- [3] D. Blackwell. Equivalent comparisons of experiments. *Ann. Math. Statist.* **24** (1953), 265–272.
- [4] T. Bousch. Le poisson n’a pas d’arêtes. *Ann. Inst. H. Poincaré Probab. Statist.* **36** (2000), 489–508.
- [5] P. Cartier, J. M. G. Fell and P.-A. Meyer. Comparaison des mesures portées par un ensemble convexe compact. *Bull. Soc. Math. France* **92** (1964), 435–445.
- [6] G. Choquet. Le théorème de représentation intégrale dans les ensembles convexes compacts. *Ann. Inst. Fourier (Grenoble)* **10** (1960), 333–344.
- [7] G. Contreras, A. Lopes and P. Thieullen. Lyapunov minimizing measures for expanding maps of the circle. *Ergod. Th. & Dynam. Sys.* **21** (2001), 1379–1409.
- [8] J.-P. Conze and Y. Guivarc’h. Croissance des sommes ergodiques et principe variationnel. *Manuscript*, 1993.
- [9] G. H. Hardy, J. E. Littlewood and G. Pólya. Some simple inequalities satisfied by convex functions. *Messenger Math.* **58** (1929), 145–152.
- [10] G. H. Hardy, J. E. Littlewood and G. Pólya. *Inequalities*, Cambridge University Press, Cambridge, (1st edn, 1934; 2nd edn, 1952).
- [11] G. Iommi. Ergodic optimization for renewal type shifts. *Monatsh. Math.* **150** (2007), 91–95.
- [12] O. Jenkinson. Frequency locking on the boundary of the barycentre set. *Experiment. Math.* **9** (2000), 309–317.
- [13] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst.* **15** (2006), 197–224.
- [14] O. Jenkinson. Optimization and majorization of invariant measures. *Electron. Res. Announc. Amer. Math. Soc.* **13** (2007), 1–12.
- [15] O. Jenkinson. A partial order on $\times 2$ -invariant measures. *Math. Res. Lett.* **15** (2008), 893–900.

- 01 [16] O. Jenkinson. Balanced words and majorization. *Preprint*.
- 02 [17] O. Jenkinson, R. D. Mauldin and M. Urbański. Ergodic optimization for countable alphabet subshifts of
finite type. *Ergod. Th. & Dynam. Sys.* **26** (2006), 1791–1803.
- 03 [18] O. Jenkinson, R. D. Mauldin and M. Urbański. Ergodic optimization for non-compact dynamical systems.
04 *Dyn. Syst.* **22** (2007), 379–388.
- 05 [19] J. Karamata. Sur une inégalité relative aux fonctions convexes. *Publ. Math. Univ. Belgrade* **1** (1932),
06 145–148.
- 07 [20] S. Karlin and A. Novikoff. Generalized convex inequalities. *Pacific J. Math.* **13** (1963), 1251–1279.
- 08 [21] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications (Mathematics in
Science and Engineering, 143)*. Academic Press, New York, 1979.
- 09 [22] M. Rothschild and J. Stiglitz. Increasing risk I: a definition. *J. Econom. Theory* **2** (1970), 225–243.
- 10 [23] H. L. Royden. *Real Analysis*, 3rd edn. MacMillan, New York, 1988.
- 11 [24] J. Steel. *PhD Thesis*, Queen Mary, University of London, in preparation.
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44

AUTHOR QUERIES

Please reply to these questions on the relevant page of the proof;
please do not write on this page.

Q1 (page 12)

Please update Refs. [1, 16, 24] if possible.

PROOF

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44