Queen Mary University of London

Majorisation ordering of measures invariant under transformations of the interval

Steel, Jacob

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without the prior written consent of the author

For additional information about this publication click this link.
http://qmro.qmul.ac.uk/jspui/handle/123456789/1292

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

Majorisation Ordering of Measures Invariant Under Transformations of the Interval

by

Jacob Steel

PhD Thesis

Queen Mary, University of London

Submitted September 2010

Declaration

Chapters 1, 2, 3 and 4 are my own work. Chapter 5 is joint work with Oliver Jenkinson. Chapter 2 is closely based on a paper [32] accepted for publication in the Procedings of the American Mathematical Society. Chapter 2 is based on the paper [33], and Chapter 3 is based on the paper [34], both submitted for publication. Chapter 5 is based on a joint paper [23] with Oliver Jenkinson, published in Ergodic Theory & Dynamical Systems.

Abstract

Majorisation is a partial ordering that can be applied to the set of probability measures on the unit interval $I = [0, 1)$. Its defining property is that one measure μ majorises another measure ν, written $\mu \succ \nu$, if $\int_I f d\mu \geq \int_I f d\nu$ for every convex real-valued function $f : I \to \mathbb{R}$.

This means that studying the majorisation of $\mathcal{M}_T$, the set of measures invariant under a transformation $T : I \to I$, can give us insight into finding the maximising and minimising T-invariant measures for convex and concave f.

In this thesis I look at the majorisation ordering of $\mathcal{M}_T$ for four categories of transformations T: concave unimodal maps, the doubling map $T : x \mapsto 2x \pmod 1$, the family of shifted doubling maps $T_\beta : x \mapsto 2x + \beta \pmod 1$, and the family of orientation-reversing weakly-expanding maps.

Contents

Declaration **2**

Abstract **3**

List of Figures **5**

1 Introduction **6**

2 Concave Unimodal Maps **9**

3 Shifted Doubling Maps **14**

3.1 Range of possible barycentres for the family of shifted doubling maps 15

3.2 Majorisation of measures corresponding to Sturmian measures 22

3.3 Majorisation of measures corresponding to given T-invariant measures 25

3.4 Majorisation of $T_{\frac{1}{2}}$-invariant measures 27

4 Periodic orbits of the Doubling Map **30**

4.1 Possible orderings of elements of periodic orbits of T 31

4.2 Majorisation ordering of periodic orbits of weight 2 33

4.3 Words of Weight 3 36

5 Weakly Expanding Orientation-Reversing Maps **49**

5.1 Introduction 49

5.2 Preliminaries 51

5.3 Minimal elements of $(\mathcal{M}_T, \prec)$, and minimizing measures for convex functions f 53

5.4 Maximal elements of $(\mathcal{M}_T, \prec)$, and maximizing measures for convex f 56

5.5 Computations 59

5.6 Infinitely many branches: the Gauss map 60

Bibliography **63**

List of Figures

2.1 In unimodal maps in which A is not dense, majorisation can occur, but the function w_Δ must still be 0 on the closure of A. 13

3.1 Centre of mass of the Sturmian measure supported on the connected interval $[\beta, \beta \pm \frac{1}{2}]$ 17

3.2 Range of possible barycentres of T_β-invariant measures 18

4.1 Hasse diagram of majorisation ordering on gap sequences of words of weight 2, length 8 34

4.2 The Hasse diagram for majorisation ordering of words of weight 3 at a given length l, indexed by gap sequence as in Figure 4.3, can be constructed from this infinite diagram simply by restricting it to the set of gap sequences that correspond to points that are minimal in their own orbit at length l. For example, the Hasse diagrams for majorisation at weights 3,4,6,8 and 15 can be generated by taking the entries above the red lines indicated, and replacing "l" with the relevant value. 37

4.3 Hasse diagrams of majorisation ordering of periodic orbits of weight 3, lengths 6-10, denoted by the gap sequences of their minimal elements. Lines indicate majorisation; the entry above and/or to the left majorises the other. So, for example, the line linking 411 to 321 means that the orbit of 7/63 (binary expansion $\dot{0}0011\dot{1}$) majorises the orbit of 11/63 (binary expansion $\dot{0}0101\dot{1}$). 37

5.1 Hasse diagram for a portion of $(\mathcal{M}_{1/2}, \prec) \subset (\mathcal{M}_T, \prec)$, where T is the reverse doubling map. Symbolic codes denote the corresponding periodic orbit measures. The least element in $(\mathcal{M}_{1/2}, \prec)$ is the non-ergodic measure $(\delta_{1/3} + \delta_{2/3})/2$, while the largest element is the invariant measure supported by the period-2 orbit $\{0, 1\}$. 60

Chapter 1

Introduction

Majorisation (alternatively known as *second order stochastic dominance*, or as *dilation*) is a partial ordering. Equivalent partial orderings all denoted by the same name can be applied to various different sorts of sets, but this thesis is concerned with majorisation as it is applied to the set of Borel probability measures on the semi-open unit interval $I = [0, 1)$. Heuristically, one measure majorises another if they have the same centre of mass but one is more spread out than the other. The formal definition is most easily stated in terms of the cumulative density function, $w_\mu(x) = \int_0^x \mu[0, s)ds$: $\mu \succ \nu$ if $w_\mu(1) = w_\nu(1)$ and $w_\mu(x) \geq w_\nu(x)$ for all $0 \leq x \leq 1$. The connection between the formal and heuristic definitions is made clearer by the observation that the centre of mass $\varrho(\mu)$ satisfies $\varrho(\mu) = 1 - w_\mu(1)$.

Remark 1.0.1. Majorisation is also known as second-order stochastic dominance; analogous definitions of stochastic dominance of other orders also exist. Assuming for the sake of elegance that both μ and ν are absolutely continuous with respect to Lebesgue measure, with corresponding density functions r_μ, r_n, then zeroth order stochastic dominance is the property $r_\mu(x) \geq r_\nu(x)$ for all x and guarantees that $\int_I f d\mu \geq \int_I f d\nu$ for positive functions; first order stochastic dominance is the property that $\int_0^1 r_\mu(s)dx = \int_0^1 r_\nu(s)dx$ (i.e. $\mu[0, 1) = \nu[0, 1)$) and $\int_0^x r_\mu(s)ds \leq \int_0^x r_\nu(s)ds$ and guarantees that $\int_I f d\mu \geq \int_I f d\nu$ for increasing functions, and so on. Note that the absolute continuity assumption is purely for ease of stating the expressions; entirely analogous results exist when μ and ν are not absolutely continuous.

The motivation for studying this ordering is the following result, and its applications to ergodic optimisation:

Lemma 1.0.2. *Let μ, ν be Borel probability measures on the unit interval with $\mu \succ \nu$, and let $f : I \to \mathbb{R}$ be a convex real valued function. Then $\int_I f d\mu \geq \int_I f d\nu$.*

Corollary 1. Obviously, if in the above lemma we replace "convex" with "concave" then $\int_I f d\mu \leq \int_I f d\nu$.

As a result, understanding the majorisation ordering of measures invariant under a given transformation is a good way of understanding the ergodic optimisation of convex and concave functions with respect to that transformation.

The basic problem of ergodic optimisation is as follows: Denote the set of T-invariant Borel probability measures on X by $\mathcal{M}_T$. Given a transformation $T : X \to X$ on some space X and a function $f : X \to \mathbb{R}$, what is the value of $\max_{\mu \in \mathcal{M}_T} \int f d\mu$ and which measures in $\mathcal{M}_T$ attain this maximum?

The simplest space to study ergodic optimisation on is the interval, usually parameterised as either $I = [0, 1]$ or $[0, 1)$, and a good deal of study has been applied to the ergodic optimisation of specific combinations of transformations T and functions f. The transformations looked at are usually those which have been most studied previously in other branches of dynamics - For example the logistic maps $T_\mu : x \mapsto \mu x(1 - x)$, the tent map $T(x) = \begin{cases} 2x \text{ if } 0 \leq x \leq \frac{1}{2} \\ 2(1-x) \text{ if } \frac{1}{2} \leq x \leq 1 \end{cases}$, the sawtooth maps $T : x \mapsto nx \pmod 1$ and the Gauss map $T : x \mapsto 1/x \pmod 1$. Choices of function to optimise usually follow from the transformation, but common choices are classes of functions defined by some information about where they are largest or smallest - increasing functions, convex or concave functions, sine waves and characteristic functions are all natural choices, whereas e.g. the class of all polynomials or the class of all piecewise linear functions would generally be less likely to yield interesting results.

In this thesis, I will look at four classes of transformation and their corresponding sets of invariant measures: the family of concave unimodal maps (Chapter 2), the family of shifted doubling maps $T_\beta : x \mapsto 2x + \beta \pmod 1$ (Chapter 3), the doubling map $T : x \mapsto 2x \pmod 1$ (Chapter 4), and (jointly with Oliver Jenkinson) the family of orientation-reversing weakly expanding maps (Chapter 5).

For the family of concave unimodal maps I show that if the set $A = \{x : T^n(x) = 1 \text{ for some } n\}$ is dense then no T-invariant measure can majorise any other.

For the family of shifted doubling maps $T_\beta : x \mapsto 2x + \beta \pmod 1$, I show that every T_β-invariant measure is simply a T_0-invariant measure transposed by β, and that every measure majorises a measure corresponding in this way to a Sturmian measure. I categorise the set of possible centres of mass of a T_β-invariant measure. Given a pair of T_0-invariant measures μ, ν, I give a way of finding out for which, if any, β do the corresponding T_β-invariant measures μ_β, ν_β measures satisfy $\mu_\beta \succ \nu_\beta$, and I show that in the case $\beta = \frac{1}{2}$ all $T_{\frac{1}{2}}$-ergodic measures majorise the fixed point measure $\delta_{\frac{1}{2}}$, and no other majorisation occurs.

For the doubling map $T : x \mapsto 2x \pmod 1$, it is already known that for any centre of mass there exists a unique measure of a class known as Sturmian that is majorised by all other T-invariant measures of that centre of mass (see [19] for this result, and [6] for more on Sturmian measures in general), and that for any given measure there exists a T-invariant measure that majorises it, so the problem of maximising concave measures is solved and the maximisation of convex measures is impossible. I illustrate some of the structure between these two extremes, looking at periodic orbits of weight[1] 2 and weight 3, and I also give some associated results about possible cyclic orderings of periodic orbits under the doubling map.

For the family of orientation reversing weakly expanding maps, Jenkinson and I show that every measure majorises a measure of the form $\lambda\delta_{x_i}+(1-\lambda)\delta_{x_{i+1}}$, consisting of a linear combination of Dirac delta measures supported on consecutive fixed points. We also show that if $T(0) = 1$ and $T(1) = 0$ then every measure is majorised by a measure of the form $\frac{\lambda}{2}\delta_0 + (1-\lambda)\delta_{x_i} + \frac{\lambda}{2}\delta_1$.

[1]The *weight* of a periodic orbit is the number of 1's in its symbolic representation, cf. Definition 4.2.1.

Chapter 2

Concave Unimodal Maps

Theorem 2.0.3. *Let $T : I \to I$ be unimodal (i.e. there exists $a \in I$ such that $T|_{[0,a]}$ is increasing and $T|_{[a,1]}$ is decreasing), concave, and such that $\{x : T^n(x) = T(a)$ for some $n\}$ is dense in I.*

If μ, ν are T-invariant ergodic measures on I, with $\mu \prec \nu$, then necessarily $\mu = \nu$; in other words, no non-trivial majorisation can occur among such measures.

Remark 2.0.4. In terms of ergodic optimisation, a consequence of Theorem 2.0.3 is that, by contrast to the situation for the doubling map (cf. [18, 19, 20]), for concave unimodal T with $\{x : T^n(x) = T(a)$ for some $n\}$ dense, no class of invariant measures plays a privileged role in terms of maximising the integrals of concave functions.

Proof. Let $A := \{x : T^n(x) = T(a)$ for some $n\}$ denote the set of iterated preimages of the maximum. Let μ, ν be T-invariant ergodic Borel probability measures with $\mu \prec \nu$. We will show that $\mu \equiv \nu$ on the closure of A. Since T has a maximum at a, $\mu((T(a), 1]) = 0$. So, by rescaling, without loss of generality we may assume that $T(a) = 1$. Similarly, if $T(1) > 0$ then $\mu(0, T(1)) = 0$, and the Dirac measure at the leftmost (fixed) point 0 is clearly incomparable to all other measures (none of which share its barycentre), so without loss of generality we may assume that $T(1) = 0$.

Let $T_-^{-1}(x)$ and $T_+^{-1}(x)$ be the left-half and right-half preimages of x, respectively, or $T_-^{-1}(x) = 0$ if $x < T(0)$. By the above assumption, these are well-defined.

Next, we focus attention on two of the nth preimages of a point x, defined by

$$T_{--}^{-n}(x) := (T_-^{-1})^n(x) \qquad \text{and} \qquad T_{+-}^{-n}(x) := T_+^{-1}(T_-^{-(n-1)}(x)) .$$

In particular, note that if $y > T_{+-}^{-n}(1)$ then $T_{+-}^{-n}(T^n(y)) = T^n(T_{+-}^{-n}(y)) = y$. This is only possible if $T_{+-}^{-n}(1) < 1$.

We now consider constraints imposed on w_μ by the T-invariance of μ. By definition $w_\mu(x) = \int_0^x \mu[0,s)ds$, and

$$\begin{aligned}
\mu[0,s) &= \mu\left[0, T_-^{-1}(s)\right) + \mu\left(T_+^{-1}(s), 1\right] \\
&= \mu\left[0, T_-^{-2}(s)\right) + \mu\left(T_{+-}^{-2}(s), 1\right] + \mu\left(T_{+-}^{-1}(s), 1\right] \\
&= \lim_{n\to\infty} \mu\left[0, T_{--}^{-n}(s)\right) + \sum_{n=1}^{\infty} \mu\left(T_{+-}^{-n}(s), 1\right] .
\end{aligned}$$

Clearly the Dirac measure concentrated at the fixed point 0 does not majorise any other measure, and since we are only interested in majorisation among *ergodic* measures μ, we may assume that $0 = \mu(\{0\})$. Therefore $0 = \lim_{n\to\infty} \mu\left[0, T_{--}^{-n}(s)\right)$ for any s, so

$$\mu[0,s) = \sum_{n=1}^{\infty} \mu\left(T_{+-}^{-n}(s), 1\right] .$$

We are interested in the difference between w_μ and w_ν. Defining Δ to be the signed measure $\Delta = \mu - \nu$, and $w_\Delta = w_\mu - w_\nu$, we see that

$$w_\Delta(x) = \int_0^x \sum_{n=1}^{\infty} \Delta\left(T_{+-}^{-n}(s), 1\right] ds = \sum_{n=1}^{\infty} \int_0^x \Delta\left(T_{+-}^{-n}(s), 1\right] ds$$

since $\sum_{n=1}^{k} \nu\left[0, T_{+-}^{-n}(s)\right] - \mu\left[0, T_{+-}^{-n}(s)\right]$ is a monotone function of s for sufficiently large k. Note also that $\Delta[0,1] = 0$, as Δ is the difference of two probability measures.

Now T need not be smooth, but since it is concave, the first derivative is defined except at countably many points, the left and right derivatives are defined everywhere and the second derivative can be defined everywhere as a sum of a real-valued function and a countable number of Dirac Delta functions. The same is true of T^n.

Suppose that $x > T_{+-}^{-n}(1)$ - that is to say, x is to the right of the rightmost maxima of T^n. It follows that $x > a$, and $T^i(x) < a$ for $1 \le i \le n-1$. Unimodality of T then implies that $T'(x) < 0$, and $T'(T^i x) > 0$ for $1 \le i \le n-1$, and hence

$$(T^n)'(x) = \prod_{i=0}^{n-1} T'(T^{i-1}(x)) < 0 \quad \text{for } x \in [T_{+-}^{-n}(1), 1] . \tag{2.1}$$

In addition, concavity of T means that $T'' \le 0$, so

$$(T^n)''(x) = \sum_{i=0}^{n-1} T''(T^i x) \prod_{j=i+1}^{n-1} T'(T^j x) \left(\prod_{j=0}^{i-1} T'(T^j x)\right)^2 \le 0 \ \text{ for } x \in (T_{+-}^{-n}(1), 1] . \tag{2.2}$$

Now

$$w_\Delta(x) = \sum_{n=1}^{\infty} \int_0^x \Delta\left(T_{+-}^{-n}(s), 1\right] ds = \sum_{n=1}^{\infty} \int_0^x -\Delta\left[0, T_{+-}^{-n}(s)\right] ds\,.$$

Letting $y = T_{+-}^{-n}(s)$, and noting that $T_{+-}^{-n}(x) \geq T_{+-}^{-n}(1)$ implies that either $T_{+-}^{-n}(x) = 1$ or $T^n(y) = s$, we obtain

$$w_\Delta(x) = \sum_{n=1}^{\infty} \int_{T_{+-}^{-n}(x)}^1 \Delta\,[0, y]\, T^{n\prime}(y)\, dy = \sum_{n=1}^{\infty} \int_{T_{+-}^{-n}(x)}^1 w_\Delta'(y)\, T^{n\prime}(y)\, dy\,.$$

Integration by parts then gives

$$w_\Delta(x) = \sum_{n=1}^{\infty} \left(w_\Delta(1)\, T^{n\prime}(1) - w_\Delta(T_{+-}^{-n}(x))\, T^{n\prime}(T_{+-}^{-n}(x)) - \int_{T_{+-}^{-n}(x)}^1 w_\Delta(y) T^{n\prime\prime}(y)\, dy \right). \tag{2.3}$$

Now suppose that μ majorises ν. It follows that $w_\Delta(0) = w_\Delta(1) = 0$, and $w_\Delta \geq 0$. Setting $x = 1$ in (2.3) gives

$$0 = w_\Delta(1) = \sum_{n=1}^{\infty} \left(-w_\Delta(T_{+-}^{-n}(1))\, T^{n\prime}(T_{+-}^{-n}(1)) - \int_{T_{+-}^{-n}(1)}^1 w_\Delta(y) T^{n\prime\prime}(y)\, dy \right),$$

and hence

$$\sum_{n=1}^{\infty} w_\Delta(T_{+-}^{-n}(1))\, T^{n\prime}(T_{+-}^{-n}(1)) = -\sum_{n=1}^{\infty} \int_{T_{+-}^{-n}(1)}^1 w_\Delta(y)\, T^{n\prime\prime}(y)\, dy \;\geq 0 \tag{2.4}$$

because $T^{n\prime\prime} \leq 0$ on $[T_{+-}^{-n}(1), 1]$, by (2.2).

Now w_Δ is non-negative and not identically zero (since $\mu \neq \nu$), and (2.1) gives $T^{n\prime}(T_{+-}^{-n}(1)) \leq 0$ for each $n \geq 1$, so the only way that (2.4) can be satisfied is if

$$w_\Delta(T_{+-}^{-n}(1)) = 0 \quad \text{for all } n \geq 1\,. \tag{2.5}$$

We can now apply the above argument iteratively. For any $m \geq 1$, set $x = T_{+-}^{-m}(1)$ in (2.3), and using (2.5) (with $n = m$), we see that

$$0 = w_\Delta(T_{+-}^{-m}(1)) = \sum_{n=1}^{\infty} \left(-w_\Delta(T_{+-}^{-n}T_{+-}^{-m}(1))\, T^{n\prime}(T_{+-}^{-n}T_{+-}^{-m}(1)) - \int_{T_{+-}^{-n}T_{+-}^{-m}(1)}^1 w_\Delta(y) T^{n\prime\prime}(y)\, dy \right),$$

and hence

$$\sum_{n=1}^{\infty} w_\Delta(T_{+-}^{-n}T_{+-}^{-m}(1))\,T^{n\prime}(T_{+-}^{-n}T_{+-}^{-m}(1)) = -\sum_{n=1}^{\infty}\int_{T_{+-}^{-n}T_{+-}^{-m}(1)}^{1} w_\Delta(y)\,T^{n\prime\prime}(y)\,dy \;\geq 0\,, \tag{2.6}$$

and by arguing as above we deduce that necessarily

$$w_\Delta(T_{+-}^{-n}T_{+-}^{-m}(1)) = 0 \quad \text{for all } m, n \geq 1\,. \tag{2.7}$$

Now, every point satisfying $x \geq a, T^k(x) = 1$, for some k must be of form $T_{+-}^{n_1}(T_{+-}^{n_2}(...(x)...))$ for some $n_1 + n_2... = k$. So, by repeating the above steps process we see that $w_\Delta(x) = 0$ whenever $T^k(x) = 1$ for some k and $x > a$. Since the set of these points is dense by assumption, and w_Δ is continuous, this implies that w_Δ is zero on $[a, 1]$, and hence everywhere by T-invariance, and so the measures μ and ν must be identical, completing the proof.

□

From the above proof it is clear that even if the set $A = \{x : T^n(x) = T(a) \text{ for some } n\}$ is not dense, if one T-invariant ergodic measure majorises another then the function w_Δ defined above must be identically zero on the closure of A, as illustrated by the following examples.

Example 2.0.5. Consider the map $T_1 : I \to I$ defined by $T_1(x) = \frac{x}{2}+\frac{3}{4}$ for $0 \leq x \leq \frac{1}{2}$ and $T_1(x) = 2 - 2x$ for $\frac{1}{2} \leq x \leq 1$. Here points in the open intervals $(0, \frac{1}{2})$ and $(\frac{3}{4}, 1)$ are never iterated onto 1, so these intervals do not contain any preimages of 1; thus the density hypothesis of Theorem 2.0.3 is not satisfied. Note that the (ergodic) T_1-invariant measure supported on the period-4 orbit $\left\{0, \frac{1}{2}, \frac{3}{4}, 1\right\}$ does majorise the (ergodic) T_1-invariant measure supported on the period-2 orbit $\left\{\frac{1}{4}, \frac{7}{8}\right\}$. However, for these two measures the function w_{Δ_a} is zero except on the open intervals $(0, \frac{1}{2})$ and $(\frac{3}{4}, 1)$, which contain no preimages of 1.

Alternatively, consider any map $T_2 : I \to I$ such that $T_2(x) = 2k - x$ for some interval containing the point k, where $\frac{1}{2} < k < 1$. Such a map has infinitely many period-2 orbits (any pair of points of the form $(k + \epsilon, k - \epsilon)$ is a period-2 orbit), whose corresponding (ergodic) invariant measures majorise each other, as well as the Dirac measure concentrated on the fixed point k. Again, note that for such pairs of measures the function w_{Δ_2} will be 0 except on the interval of periodic points.

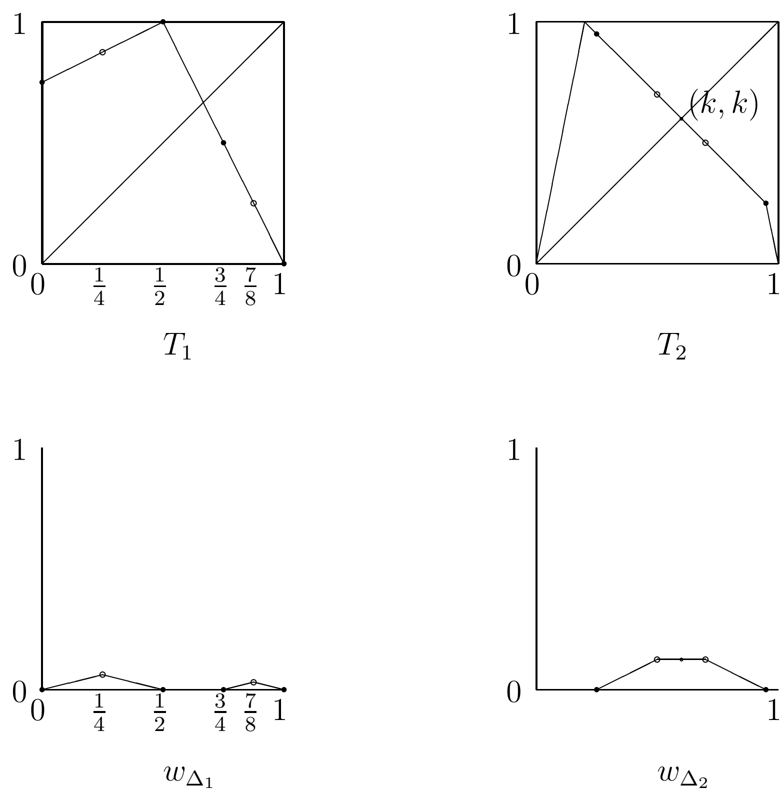

Figure 2.1: In unimodal maps in which A is not dense, majorisation can occur, but the function w_Δ must still be 0 on the closure of A.

Chapter 3

Shifted Doubling Maps

The dynamics of the doubling map on the unit interval, $T_0 : x \mapsto 2x \pmod 1$, and the properties of its set of invariant probability measures $\mathcal{M}_{T_0}$, have been extensively studied (see e.g. [6, 19]), but less is known about the family of maps $T_\beta : x \mapsto 2x + \beta \pmod 1$. In this chapter we look at a partial ordering on the elements of the associated sets of T_β-invariant measures $\mathcal{M}_{T_\beta}$, which has significance to the ergodic optimisation of concave and convex functions on the unit interval with respect to T_β (see e.g. [17, 19, 23]).

Recall that majorisation is a partial ordering on the set of Borel probability measures on the semi-open unit interval $I = [0, 1)$ (it also occurs in other contexts, see for example [26]). It is defined in terms of the cumulative density function $w_\mu(x) := \int_0^x \mu[0, s)ds$ as follows:

Definition 3.0.6. (Majorisation)

Let μ, ν be two Borel probability measures on $I = [0,1)$. We say that μ is *majorised* by ν, written $\mu \prec \nu$, if $w_\mu(x) \leq w_\nu(x)$ for all $0 \leq x \leq 1$, and $w_\mu(1) = w_\nu(1)$.

Remark 3.0.7. Note that the centre of mass or barycentre of a measure, defined by $\varrho(\mu) = \int_0^1 x d\mu(x)$, satisfies $\varrho(\mu) = 1 - w_\mu(1)$ and so, heuristically, $\mu \prec \nu$ means that μ and ν have the same barycentre but ν is more spread out.

Majorisation has the useful property that if $\mu \succ \nu$ then for all convex real-valued functions f we have $\int_I f d\mu \geq \int_I f d\nu$. For a proof of this, see [19]. This means that understanding the majorisation ordering of the elements of $\mathcal{M}_T$, the set of T-invariant probability measures, can often help us solve the problem of finding $\max_{\mu \in \mathcal{M}_T} \int_I f d\mu$ for concave and convex f. This approach has been exploited in [18, 19] for the standard unshifted doubling map $T(x) = 2x \pmod 1$. In this chapter we apply it to the family of shifted doubling maps, $T_\beta(x) = 2x + \beta \pmod 1$.

3.1 Range of possible barycentres for the family of shifted doubling maps

Let $\mathcal{M}_{T_\beta}$ denote the set of those Borel probability measures on the interval which are invariant under T_β. Our first result concerns the set of points $\varrho(\mathcal{M}_{T_\beta})$ which arise as a barycentre for some T_β-invariant probability measure. Clearly each $\varrho(\mathcal{M}_{T_\beta})$ is an interval (since $\mathcal{M}_{T_\beta}$ is convex and the map $\mu \mapsto \varrho(\mu)$ is affine). To determine the endpoints, we will need a class of measures known as Sturmian measures:

Definition 3.1.1 (Sturmian orbits and measures)**.** For any two numbers $\varrho \in [0,1)$, $s \in [0,1]$ let $x(\varrho, s, \pm)$ be the two points such that their binary expansions, given by $x(s,\varrho,\pm) = \sum_{i=1}^{\infty} x(s,\varrho,\pm)_i 2^{-i}$ have the following properties:

$$\begin{array}{lll} n\varrho + s - 1 & < \sum_{i=1}^{n} x(\varrho,s,+)_i & \leq n\varrho + s \\ n\varrho + s - 1 & \leq \sum_{i=1}^{n} x(\varrho,s,-)_i & < n\varrho + s \end{array} \tag{3.1}$$

$x(\varrho, 0, -)$ and $x(\varrho, 1, +)$ are undefined, but otherwise for any values of $\varrho, s, \pm$ these equations uniquely define $x(\varrho, s, \pm)_i$ to be either 0 or 1.

Example 3.1.2. For example, $x(\frac{1}{2}, 0, +) = \frac{1}{3}$, which has binary expansion 010101..., while $x(\frac{1}{2}, 0, +) = \frac{2}{3}$ which has binary expansion 101010..., and $x(\frac{2}{5}, \frac{1}{100}, -) = \frac{5}{31}$ which has binary expansion 0010100101... and $x(\frac{3}{4}, \frac{19}{20}, +) = \frac{14}{15}$ which has binary expansion 11101110....

The orbit of any point of form $x(\varrho, s, \pm)$ is called a *Sturmian orbit.*

Sturmian orbits have, among others, the following interesting properties (see [6, 29, 36]):

Proposition 3.1.3.

(1) For all s, $x(\varrho,0,+) \leq x(\varrho,s,\pm) \leq x(\varrho,1,-)$.

(2) If ϱ is rational then for all values of s and $\pm$, $x(\varrho,s,\pm)$ have the same T-orbit.

(3) If ϱ is irrational then for all values of s, $\pm$, $x(\varrho,s,\pm)$has the same orbit closure.

(4) If ϱ is rational then $x(\varrho,1,-) < x(\varrho,0,+) + \frac{1}{2}$

(5) If ϱ is irrational then $x(\varrho,1,-) = x(\varrho,0,+) + \frac{1}{2}$

(6) The orbit of a point $x \in I$ can be contained within a semi-circle (a set of form $[a, a+\frac{1}{2}]$ or $[0,a] \cup [a+\frac{1}{2}, 1]$) if and only if x is Sturmian.

(7) For every semi-circle $I_\alpha = [\alpha - \frac{1}{2} \text{ mod } 1, \alpha]$, there exists a unique $\varrho(\alpha)$ such that the orbit of $x(\varrho, s, \pm)$ is contained in I_α for all values of $s, \pm$.

(8) The support of any sturmian measure is a cantor set of zero Hausdorff measure. The transformation T is order-preserving upon this set, and so can be extended to a circle homeomorphism of rotation, which will have rotation number ϱ.

Definition 3.1.4. The only T_0-invariant Borel probability measure supported on a semi-circle is that given by the hitting frequency of the Sturmian orbits contained in it (all Sturmian orbits on a given barycentre have the same orbit closure, and the same hitting frequency on any Borel set). We call this a *Sturmian measure.*

Remark 3.1.5. The study of Sturmian words originates with Johannes Bernoulli, but the first comprehensive treatment, and the first use of the term "Sturmian" to describe them, was by Hedlund and Morse in 1940 [28]. Initially they were studied primarily as words in the context of the combinatorics of the shift map, but more recently their applications to billiards theory [35] and dynamical systems [29] have been more studied.

The latter family of applications has led to increased consideration being given to Sturmians as probability measures on the unit interval, rather than as words of the shift map. These measures are very "clumped up" - they are the only measures which can be supported on a semi-circle - and they turn out to be the maximising or minimising invariant measures for many classes of functions on the unit interval (degree-one trigonometric functions, concave functions, sine waves, etc), see [2], [3], [5]. For a more comprehensive treatment of Sturmian measures, see e.g. [6], [36].

Given these, we can state and prove the following two theorems:

Theorem 3.1.6. *The set of possible barycentres of elements of $\mathcal{M}_{T_\beta}$ is the interval*

$$\varrho(\mathcal{M}_{T_\beta}) = \begin{cases} [1-\beta-\sigma(\beta), 1-\beta] & \text{if } \beta \leq \frac{1}{2} \\ [1-\beta, 1-\beta+\sigma(\beta)] & \text{if } \beta \geq \frac{1}{2} \end{cases}$$

where

$$\sigma(\beta) := \begin{cases} \max_{\mu \in \mathcal{M}_{T_\beta}} \mu([\beta, \frac{1}{2}]) & \text{if } \beta \leq \frac{1}{2} \\ \max_{\mu \in \mathcal{M}_{T_\beta}} \mu([\frac{1}{2}, \beta]) & \text{if } \beta \geq \frac{1}{2} \end{cases}$$

Theorem 3.1.7. *For each β, let $\varrho(\beta)$ denote the barycentre of the Sturmian measure supported on the (unique) semicircle which contains the point $\frac{1}{2}$ and has one endpoint*

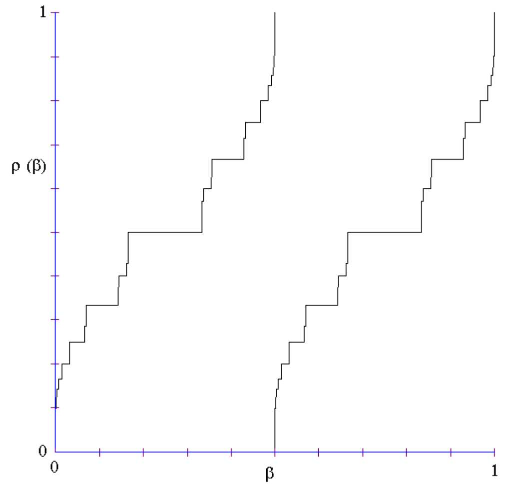

Figure 3.1: Centre of mass of the Sturmian measure supported on the connected interval $[\beta, \beta \pm \frac{1}{2}]$

at β.[1] $\varrho(\beta)$ *is illustrated as figure 3.1.*

Then for each β*, we have*

$$\sigma(\beta) = \begin{cases} 1 - \varrho(\beta) & \text{if } \beta \leq \frac{1}{2} \\ \varrho(\beta) & \text{if } \beta \geq \frac{1}{2}. \end{cases}$$

Theorems 3.1.6 and 3.1.7 together completely categorise the set $\varrho(\mathcal{M}_{T_\beta})$, which is illustrated as the shaded region in figure 2. Note that when $\beta = 0$, any value in the interval $[0, 1)$ can be the centre of mass of a T_β-invariant orbit.

Corollary 2. If $\varrho(\beta)$ is the centre of mass of the sturmian measure on the semicircle with an endpoint at β that does not contain 0 or 1, as above, then the set of possible values of ϱ is

$$\varrho \in \begin{cases} [\varrho(\beta) - \beta, 1 - \beta] & \text{if } \beta \leq \frac{1}{2} \\ [1 - \beta, \varrho(\beta) - \beta + 1] & \text{if } \beta \geq \frac{1}{2} \end{cases}$$

which is to say that $k - \beta$ is an admissible value for the centre of mass of a T_β-invariant measure if and only if the measure β-corresponding to σ_k is supported on a semicircle.

To aid in the proof of the above results, we define an equivalence relation on the set $\mathcal{M} = \bigcup_{\beta \in [0,1)} \mathcal{M}_{T_\beta}$ as follows:

[1] i.e. if $\beta \leq \frac{1}{2}$ then $\varrho(\beta)$ is the barycentre of the Sturmian measure supported on the interval $[\beta, \beta + \frac{1}{2}]$, whereas if $\beta \geq \frac{1}{2}$ then $\varrho(\beta)$ is the barycentre of the Sturmian measure supported on the semicircle $[\beta - \frac{1}{2}, \beta]$.

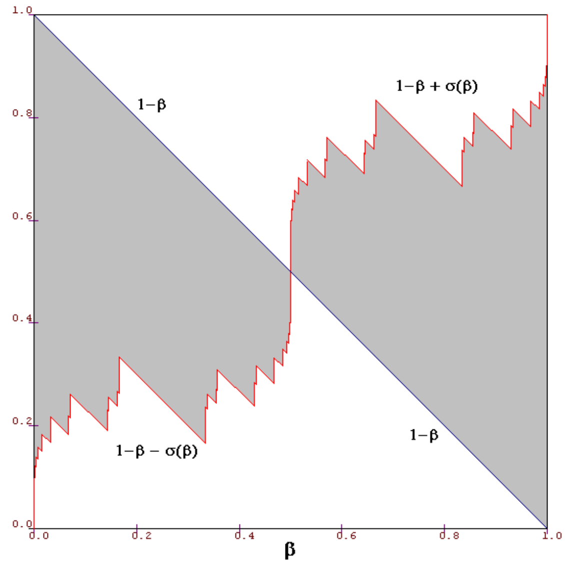

Figure 3.2: Range of possible barycentres of T_β-invariant measures

Definition 3.1.8. For any pair of measures $\mu \in \mathcal{M}_{T_\beta}$, $\nu \in \mathcal{M}_{T_\gamma}$, we say that μ *corresponds* to ν if $\mu(A - \beta \pmod 1) = \nu(A - \gamma \pmod 1)$ for all Borel sets A. For example, the set of fixed point measures $\delta_{1-\beta} \in \mathcal{M}_{T_\beta}$ correspond to one another, as do the set of measures supported on period-2 orbits, $\frac{1}{2}(\delta_{\frac{1}{3}-\beta \pmod 1} + \delta_{\frac{2}{3}-\beta \pmod 1}) \in \mathcal{M}_{T_\beta}$. Lebesgue measure is in $\mathcal{M}_{T_\beta}$ for all values of β, and corresponds to itself.

Remark 3.1.9. Note that if we allow measures supported on $[0,1]$ instead of $[0,1)$ then ambiguity arises. Although it can easily be resolved, for simplicity we shall restrict ourselves to the half-open interval.

Remark 3.1.10. Note that every element of $\mathcal{M}_{T_\beta}$ always corresponds to precisely one element of $\mathcal{M}_{T_\gamma}$, and so we can safely talk about *the* corresponding measure without needing to worry about either existance or uniqueness

Proof of Theorem 3.1.6. Let $\mu_\beta \in \mathcal{M}_{T_\beta}$ for some $\beta \neq 0, 1$, and let μ be the corresponding element of $\mathcal{M}_{T_0}$. By definition, $w_{\mu_\beta}(1) = \int_0^1 \mu_\beta[0,s)ds$. Since μ_β corresponds to μ,

$$w_{\mu_\beta}(1) = \int_0^1 \mu[\beta, \beta{+}s) \pmod 1 ds = \int_0^{1-\beta} \mu[\beta, \beta{+}s)ds + \int_{1-\beta}^1 \mu[\beta, 1] + \mu[0, \beta{+}s{-}1)ds\,,$$

and rewriting this in terms of w_μ gives us

$$w_{\mu_\beta}(1) = w_\mu(1) + \beta - \mu[0,\beta)\,. \tag{3.2}$$

Now μ is T-invariant, so taking preimages of $[0,s)$ under T we see that $w_\mu(1) =$

$\int_0^1 \mu[0,s)ds = \int_0^1 \mu\left[0,\frac{s}{2}\right) + \mu\left[\frac{1}{2},\frac{1}{2}+\frac{s}{2}\right) ds = \int_0^{\frac{1}{2}} 2\mu[0,s)ds + \int_{\frac{1}{2}}^1 2\mu[0,s) - 2\mu\left[0,\frac{1}{2}\right) ds =$ $2w_\mu(1) - \mu\left[0,\frac{1}{2}\right]$. Rearranging this gives

$$w_\mu(1) = \mu\left[0,\frac{1}{2}\right]. \tag{3.3}$$

Combining (3.3) and (3.2) gives us

$$w_{\mu_\beta}(1) = \begin{cases} \beta + \mu\left[\beta,\frac{1}{2}\right] & \text{if } \beta \le \frac{1}{2} \\ \beta - \mu\left[\frac{1}{2},\beta\right] & \text{if } \beta \ge \frac{1}{2}\,, \end{cases}$$

and because $\varrho(\mu) = 1 - w_\mu(1)$,

$$\varrho(\mu_\beta) = \begin{cases} 1 - \beta - \mu\left[\beta,\frac{1}{2}\right] & \text{if } \beta \le \frac{1}{2} \\ 1 - \beta + \mu\left[\frac{1}{2},\beta\right] & \text{if } \beta \ge \frac{1}{2}\,. \end{cases}$$

Now the Dirac measure invariant under T_β is $\delta_{1-\beta}$, and it gives $\varrho(\delta_{1-\beta}) = 1 - \beta$, which is one endpoint of $\varrho(\mathcal{M}_{T_\beta})$, while by the definition of $\sigma(\beta)$ the other endpoint must be $1 - \beta \pm \sigma(\beta)$.

□

Proof of Theorem 3.1.7. Without loss of generality, let $\beta \ge \frac{1}{2}$. We wish to find the maximum possible value of $\mu\left[\frac{1}{2},\beta\right]$ for a T-invariant probability measure μ (the case when $\beta < \frac{1}{2}$ is exactly analogous).

By the ergodic theorem, the maximum possible value a measure μ in $\mathcal{M}_T$ can assign to a set $A \subset I$ is equal to the maximum possible frequency (cf. [16]) with which the orbit under T of a point can hit that set,

$$\mu(A) = \sup_{x\in X} \limsup_{n\to\infty} \frac{\#\{i \le n : T^i(x) \in A\}}{n}$$

Let x be a point in I whose orbit hits $\left[\frac{1}{2},\beta\right]$ with maximal frequency, and let $x_1x_2x_3...$ be the binary expansion of x. The action of T on I is topologically semi-conjugate to the action of the shift map σ: $x_1x_2x_3... \mapsto x_2x_3x_4...$ on the binary expansion of x.

Now, if $T^i(x)$ is an element of the set $\left[\frac{1}{2},\beta\right]$ then x_{i+1} must be equal to 1, but not all x with $x_{i+1} = 1$ are less than β. So let us define another point X in terms of its binary expansion, as follows:

$$X_{i+1} = \begin{cases} 1 & \text{if } T^i(x) \in \left[\frac{1}{2},\beta\right] \\ 0 & \text{otherwise.} \end{cases}$$

This new point X has the properties that $T^i(X) \leq T^i(x)$ for all i, and $T^i(X) \in \left[\frac{1}{2}, \beta\right]$ whenever $T^i(x) \in \left[\frac{1}{2}, \beta\right]$. So X is a point whose orbit lands in $[\frac{1}{2}, \beta]$ with maximal frequency, and lands in $[\frac{1}{2}, \beta]$ whenever it lands in $[\frac{1}{2}, 1]$. Now define H to be the set of all such X:

$$H := \left\{ X \in \left[\frac{1}{2}, \beta\right] : X \text{ hits } [\frac{1}{2}, 1] \text{ with the greatest frequency possible, but never hits } [\beta, 1] \right\}.$$

We have shown that H is non-empty.

Now, let us introduce a new notion, the *gap sequence* of x, defined for values of x greater than $\frac{1}{2}$ in terms of the binary expansion of x as follows:

Binary expansion		Gap series
10010100101...	$\rightarrow$	3232...
100011101...	$\rightarrow$	4112...
$1101\dot{0}$	$\rightarrow$	12∞

Let β have gap sequence $g_1 g_2 g_3 ...$ and define β^1, β^2, β^3, etc to be the numbers that have the following gap sequences:

$$\begin{array}{lll}
\beta^1 & \rightarrow & g_1 + 1, g_1 + 1, ... \\
\beta^2 & \rightarrow & g_1, g_2 + 1, g_1, g_2 + 1, ... \\
\beta^3 & \rightarrow & g_1, g_2, g_3 + 1, g_1, g_2, g_3 + 1, ... \\
... & &
\end{array}$$

So for example if β has binary expansion 100110001... then β^1 would have binary expansion 1000 1000 1000..., β^2 would have binary expansion 10010 10010..., β^3 would have binary expansion 100110000 100110000... and so on.

If X is an element of H then $X_1 = 1 = \beta_1$. Define $k(X) := \inf\{i : X_i \neq \beta_i\}$, with $k(X) := \infty$ if $X = \beta$. We know that $X \leq \beta$, so we must have $X_{k(X)} = 0$ and $\beta_{k(X)} = 1$. That is to say, every value of X in H has a binary expansion $X_1, X_2, X_3 ... = \beta_1, \beta_2, \ldots, \beta_{k(X)-1}, 0, Y_1, Y_2 ...$ for some Y_i. We will say that a point whose binary expansion is of this form is *nice*.

Now Y, the point with binary expansion $Y_1, Y_2 ...$, is part of the orbit of X and hence also an element of H, and so it must also be nice. Hence every point X in H must have a gap sequence of form $g_1, ..., g_{n_1} + 1, g_1, ..., g_{n_2} + 1, g_1, ...$ So either at least one of $\beta^1, \beta^2, \beta^3 ...$, or β itself, must be in H.

We will now show that if β is Sturmian then β itself must be in H, rather than any of the β^i. Assume that β is an upper bound of a Sturmian orbit, of the form $x(\varrho, 1, -)$, again with gap sequence $g_1, g_2 ...$ Define the function

$$f(x) = \lim_{n\to\infty} \frac{\#\left\{i \le n : T^i(x) \in \left[\frac{1}{2}, 1\right]\right\}}{n}$$

Now, $f(\beta) = \varrho$ (see [19, footnote 2]) and $f(\beta^k) = \frac{k}{\left(\sum_{i=1}^k g_i\right)+1}$ but, by (3.1), taking $n = \sum_{i=1}^k g_i + 1$, we have the following inequalities:

$$\begin{aligned}
\left(\sum_{i=1}^{k} g_i + 1\right)\varrho \le k+1 &< \left(\sum_{i=1}^{n} g_i + 1\right)\varrho + 1, \\
k &< \left(\sum_{i=1}^{n} g_i + 1\right)\varrho, \\
\frac{k}{\left(\sum_{i=1}^k g_i + 1\right)} &< \varrho, \\
f(\beta^k) &< f(\beta).
\end{aligned}$$

So if β is of the form $x(\varrho, 1, -)$ - that is, if β is the maximum point of a Sturmian orbit - then β is itself a point of maximum hitting frequency for $[\frac{1}{2}, \beta]$, and its hitting frequency is precisely the barycentre of the associated Sturmian measure.

Now the barycentre of the Sturmian measure supported on the semi-circle $[\beta - \frac{1}{2}, \beta]$ is continuous and non-decreasing as a function of β (more precisely, it is a devil's staircase - see [6] for this result and for more details of this remarkable class of functions).

So if β is not the maximum point of a Sturmian orbit then there exists a unique $\varrho(\beta)$ such that the Sturmian orbit of barycentre ϱ, which we shall denote by $\sigma_{0,\varrho}$ (the reason for the additional 0 subscript will become apparent when we start dealing with measures invariant under shifted doubling maps corresponding to sturmian measures), is supported on $[\beta - \frac{1}{2}, \beta]$. For this to be the case, we must have $x(\varrho, 1, -) < \beta < x(\varrho, 0, +) + \frac{1}{2}$. Now since $x(\varrho, 1, -) < \beta$, obviously $\left[\frac{1}{2},\, x(\varrho, 1, -)\right]$ is a subset of $\left[\frac{1}{2},\, \beta\right]$, and so $\sigma_{0,\varrho}$ hits $[\frac{1}{2},\, \beta]$ with frequency ϱ.

Also, $x(\varrho + \epsilon, 1, -)$ tends to $x(\varrho, 0, +) + \frac{1}{2}$ as $\epsilon \searrow 0$, so if any orbit hits $[\frac{1}{2},\, \beta]$ with frequency more than ϱ then for sufficiently small positive ϵ it hits $[\frac{1}{2},\, x(\varrho + \epsilon, 1, -)]$ with frequency more than $\varrho + \epsilon$, which is a contradiction.

So the maximum value a measure in $\mathcal{M}_{T_0}$ can assign to $[\frac{1}{2},\, \beta]$ is precisely $\varrho(\beta)$, and this bound is achieved by the Sturmian measure $\sigma_{0,\varrho(\beta)}$.

□

These two results together completely categorise the range of possible barycentres of T_β-invariant measures (for the explicit formula for $\varrho(\beta)$, see e.g. [6]).

3.2 Majorisation of measures corresponding to Sturmian measures

We shall denote the Sturmian orbit of barycentre ϱ by $\sigma_{0,\varrho}$, and the element of $\mathcal{M}_{T_\beta}$ corresponding to $\sigma_{0,\varrho}$ by $\sigma_{\beta,\varrho-\beta}$.

Note that $\sigma_{\beta,\varrho}$ has barycentre $\varrho + \sigma_{0,\varrho+\beta}[0,\beta] \pmod 1$, which may or may not be equal to ϱ.

Every Sturmian measure (or shifted Sturmian measure) can be supported on a (not necessarily unique) semicircle (an interval of form $[\gamma,\gamma+\frac{1}{2}]$ or $[0,\gamma-\frac{1}{2}]\cup[\gamma,1]$).

Lemma 3.2.1. *The following conditions on β,ϱ are equivalent:*

(i) $\sigma_{0,\varrho+\beta}([0,\beta]) = 0$

(ii) $\sigma_{\beta,\varrho}$ *has centre of mass* ϱ

(iii) ϱ *is a permissible centre of mass for elements of* $\mathcal{M}_{T_\beta}$

(iv) The support of $\sigma_{0,\beta+\varrho}$ *can be contained in an interval of form*[2] $[\gamma,\gamma+\frac{1}{2}]$

(v) The support of $\sigma_{0,\beta+\varrho}$ *can be contained in an interval of form* $[\gamma,\gamma+\frac{1}{2}]$ *where* $\gamma < 1-\beta$ *if* $\beta < \frac{1}{2}$ *and* $\gamma > 1-\beta$ *if* $\beta > \frac{1}{2}$

Proof. The equivalence of (i) and (ii) follows directly from calculating the centre of mass of $\sigma_{\beta,\varrho}$. The equivalence of (ii) and (iii) comes from Theorems 3.1.6 and 3.1.7. The equivalence of these to (iv) and (v) comes from the fact (see Theorem 3.1.6) that one of the bounds (namely $1-\beta$) on the set of acceptable centres of mass is attained only by the fixed point measure $\delta_{1-\beta}$, and the fact that all semicircles not containing the fixed point contain the point $\beta+\frac{1}{2} \pmod 1$. $\square$

We can use this to prove the following theorem:

Theorem 3.2.2. *For $\beta \in [0,1)$, let ϱ be an admissible barycentre for T_β-invariant probability measures. If $\mu \in \mathcal{M}_{T_\beta}$ has barycentre ϱ, then $\sigma_{\beta,\varrho} \prec \mu$.*

To prove Theorem 3.2.2, first note that by Lemma 3.2.1, since ϱ is an admissible barycentre then $\sigma_{\beta,\varrho}$ does indeed have barycentre ϱ, so μ and $\sigma_{\beta,\varrho}$ are *potentially* comparable with respect to the majorisation partial order.

We will follow the method used by Jenkinson in [19] to prove this result for the unshifted doubling map. The definition of majorisation we will use is that $\mu \succ \nu$ if and only if $\int f d\mu \geq \int f d\nu$ for all convex f. Since C^2 functions are weakly dense in

[2] i.e. rather than requiring a 'wrapped' semi-circle of form $[0,\gamma-\frac{1}{2}]\cup[\gamma,1]$, though note that even when an interval of form $[\gamma,\gamma+\frac{1}{2}]$ exists, it may also be possible to find such a 'wrapped' interval, e.g this occurs for $\sigma_{\beta,1-\beta}$.

the set of convex functions, it is sufficient to show that for any given β, for all C^2 convex f, for all $\mu \in \mathcal{M}_{T_\beta}$, $\int f d\mu \geq \int f d\sigma_{\beta,\varrho}$.

For a given real value θ, define $f_\theta(x) = f(x) + \theta x$. Since $\int f_\theta d\mu = \int f d\mu + \theta\varrho$, if two measures have the same barycentre then $\int f d\mu \geq \int f d\nu$ if and only if $\int f_\theta d\mu \geq \int f_\theta d\nu$. This means that Theorem 3.2.2 is implied by the following result:

Theorem 3.2.3. *Let f be a C^2 convex function, let $\beta \in [0,1)$ and let $A = [1-\beta-\sigma(\beta), 1-\beta]$ or $[1-\beta, 1-\beta+\sigma(\beta)]$ be the set of allowed barycentres for T_β-invariant measures. Then for every $\varrho \in A$ there exists $\theta \in \mathbb{R}$ such that $\sigma_{\beta,\varrho}$ is a minimising measure for f_θ.*

Proof. Since ϱ is an allowable berycentre for T_β-invariant probability measures, Lemma 3.2.1 means that there exists $\gamma \in [0, \frac{1}{2} - \beta]$ such that the support of $\sigma_{\beta,\varrho}$ is contained in $[\gamma, \gamma + \frac{1}{2}] =: H_\gamma$.

Define the function $\tau : I \to I$ to be the preimage of x in H_γ - that is, if $\gamma \leq \frac{1-\beta}{2}$ then

$$\tau(x) = \begin{cases} (x+1-\beta)/2 & \text{if} \quad 0 \leq x < 2\gamma + \beta \\ (x-\beta)/2 & \text{if} \quad 2\gamma+\beta \leq x < 1 \end{cases}$$

while if $\gamma \geq \frac{1-\beta}{2}$ then

$$\tau(x) = \begin{cases} (x+2-\beta)/2 & \text{if} \quad 0 \leq x < 2\gamma + \beta - 1 \\ (x+1-\beta)/2 & \text{if} \quad 2\gamma+\beta-1 \leq x < 1\,. \end{cases}$$

Let $g(x) = \sum_{n=1}^{\infty} \frac{f'(\tau^n(x))}{2^n}$ be a (Lebesgue almost everywhere defined) weighted average of the gradient at the preimages under τ, and choose $\theta = -\int_0^1 g(x)dx$. Then $g_\theta = \sum_{n=1}^{\infty} 2^{-n} f'_\theta \circ \tau^n$ is a L^∞ function with Lebesgue integral 0, and so it will be the almost everywhere derivative of a Lipschitz function φ_θ, with $\varphi_\theta(0) = \varphi_\theta(1) = 0$.

From the definition of φ_θ we see that $(f + \varphi_\theta - \varphi_\theta \circ T_\beta)' = 0$ Lebesgue almost everywhere on H_γ, so applying the fundamental theorem of calculus (for absolutely continuous functions) we see that $f_\theta + \varphi_\theta - \varphi_\theta \circ T_\beta$ is constant on H_γ. We claim that this constant value is in fact the minimum value, that is that

$$(f_\theta + \varphi_\theta)(s) \leq (f_\theta + \varphi_\theta)(s + \frac{1}{2}) \quad \text{if} \quad s \in [\gamma, \frac{1}{2}] \tag{3.4}$$

$$(f_\theta + \varphi_\theta)(s) \leq (f_\theta + \varphi_\theta)(s - \frac{1}{2}) \quad \text{if} \quad s \in (\frac{1}{2}, \gamma + \frac{1}{2}] \tag{3.5}$$

This claim implies that $\sigma_{\beta,\varrho}$ is a minimising measure for $f_\theta + \varphi_\theta - \varphi_\theta \circ T_\beta$, and hence for f_θ.

To prove inequality (3.4), let $s \in [\gamma, \frac{1}{2}]$ and write $E = E_s = (\gamma, s]$ and $E' = E'_s = (\gamma + \frac{1}{2}, s + \frac{1}{2}]$. Note that $f_\theta + \varphi_\theta - \varphi_\theta \circ T_\beta$ is constant on H_γ, and so since $T_\beta(\gamma) = T_\beta(\gamma + \frac{1}{2})$ we must have $(f_\theta + \varphi_\theta)(\gamma) = (f_\theta + \varphi_\theta)(\gamma + \frac{1}{2})$, and so

$$\begin{aligned}
(f_\theta + \varphi_\theta)(s) - (f_\theta + \varphi_\theta)(s + \frac{1}{2}) &= \int_0^s (f_\theta + \varphi_\theta)' - \int_0^{s+\frac{1}{2}} (f_\theta + \varphi_\theta)' \\
&= \left(\int_0^\gamma + \int_\gamma^s - \int_0^\gamma - \int_\gamma^{\gamma+\frac{1}{2}} - \int_{\gamma+\frac{1}{2}}^{s+\frac{1}{2}} \right) (f_\theta + \varphi_\theta)' \\
&= \int_\gamma^s (f_\theta + \varphi_\theta)' - \int_{\gamma+\frac{1}{2}}^{s+\frac{1}{2}} (f_\theta + \varphi_\theta)' \\
&= \sum_{n=0}^\infty \left[\int_E 2^{-n} f'_\theta \circ \tau^n - \int_{E'} 2^{-n} f'_\theta \circ \tau^n \right] \\
&= \int C_s \cdot f'_\theta
\end{aligned}$$

where C_s is the sum of characteristic functions

$$C_s(x) = \sum_{n=0}^\infty \left[\chi_{\tau^n(E)}(x) - \chi_{\tau^n(E')}(x) \right] .$$

Now C_s is Lebesgue integrable, with $\int C_s = 0$, because $|\tau^n E| = |\tau^n E'| = 2^{-n}(s - \gamma)$. Defining $B_s(t) = \int_0^t C_s$, we have $B_s(0) = 0 = B_s(1)$, and so integrating by parts gives us $(f_\theta + \varphi_\theta)(s) - (f_\theta + \varphi_\theta)(s + \frac{1}{2}) = \int C_s \cdot f'_\theta = -\int B_s \cdot f''_\theta$.

Now f is convex, so $f''_\theta \geq 0$, and so all we have to show is that B_s is everywhere non-negative.

To prove this, observe that C_s is identically zero outside the interval $[\gamma, s + \frac{1}{2}]$, and hence so is B_s. The term $\sum_{n=0}^\infty \chi(\tau^n(E))$ is 0 except on the interval $[\gamma, \gamma + \frac{1}{2}]$, and is at least 1 on the interval $[\gamma, s] = E$, whereas the term $-\sum_{n=0}^\infty \chi(\tau^n(E'))$ is either 0 or -1 everywhere, because $\tau^m(E')$ and $\tau^n(E')$ are disjoint for all $m \neq n$. So on the interval $[\gamma + \frac{1}{2}, s + \frac{1}{2}]$ we have $C_s = -1$ and $B_s(t) = s + \frac{1}{2} - t \geq 0$, while on the interval $[\gamma, s]$ we have $C_s \geq 0$ and hence $B_s \geq 0$. Finally, on the interval $[s, \gamma + \frac{1}{2}]$ we have $B_s \geq |E| - \sum_{n=1}^\infty |\tau^n(E')| = |E| - \sum_{n=1}^\infty 2^{-n}|E| = 0$.

The proof of (3.5) is similar: if $s \in (\frac{1}{2}, \gamma + \frac{1}{2}]$ then

$$(f_\theta + \varphi_\theta)(s) - (f_\theta + \varphi_\theta)(s - \frac{1}{2}) = \int \hat{C}_s \cdot f'_\theta$$

where

$$\hat{C}_s(x) = \sum_{n=0}^{\infty} \left[\chi_{\tau^n(D)}(x) - \chi_{\tau^n(D')}(x) \right] ,$$

with $D = [s - \frac{1}{2}, \gamma)$, $D' = [s, \gamma + \frac{1}{2})$, and we can proceed as before. $\square$

Remark 3.2.4. This proof follows very closely the method used by Jenkinson in [19].

3.3 Majorisation of measures corresponding to given T-invariant measures

Theorem 3.3.1. *For any two measures $\mu, \nu \in \mathcal{M}_{T_0}$, the set of values of β such that for the corresponding $\mu_\beta, \nu_\beta \in \mathcal{M}_{T_\beta}$ we have $\mu_\beta \prec \nu_\beta$, is precisely the set of β that are global minimum points of $G_{\mu,\nu}(x)$ but not atoms of the signed measure $\mu - \nu$, where*

$$G_{\mu,\nu}(x) = w_\nu(x) - w_\mu(x) - x\big(w_\nu(1) - w_\mu(1)\big),$$

together with the left endpoints of any intervals contained in that set.

Example 3.3.2. If μ is the periodic measure $\frac{1}{3}(\delta_{\frac{1}{7}} + \delta_{\frac{2}{7}} + \delta_{\frac{4}{7}})$ and ν is Lebesgue measure (note that these measures do not have the same barycentre, hence are not comparable, but that μ_β and ν_β *will* be comparable) then

$$\begin{aligned} w_\mu(x) &= \begin{cases} 0 & 0 \le x \le \frac{1}{7} \\ \frac{(x-\frac{1}{7})}{3} & \frac{1}{7} \le x \le \frac{2}{7} \\ \frac{(2x-\frac{3}{7})}{3} & \frac{2}{7} \le x \le \frac{4}{7} \\ x - \frac{1}{3} & \frac{4}{7} \le x \le 1 \end{cases} \\ w_\nu(x) &= \tfrac{1}{2}x^2 \end{aligned}$$

$$G_{\mu,\nu}(x) = \begin{cases} \frac{1}{2}x^2 + \frac{x}{6} & 0 \le x \le \frac{1}{7} \\ \frac{1}{2}x^2 + \frac{-x}{6} + \frac{1}{21} & \frac{1}{7} \le x \le \frac{2}{7} \\ \frac{1}{2}x^2 + \frac{-3x}{6} + \frac{3}{21} & \frac{2}{7} \le x \le \frac{4}{7} \\ \frac{1}{2}x^2 + \frac{-5x}{6} + \frac{1}{3} & \frac{4}{7} \le x \le 1 \end{cases}$$

So the minimum value of $G_{\mu,\nu}$ is attained when $x = \frac{5}{6}$. If we set $\beta = \frac{5}{6}$ then μ_β is the periodic measure $\frac{1}{3}(\delta_{\frac{13}{42}} + \delta_{\frac{19}{42}} + \delta_{\frac{31}{42}})$ and ν_β is Lebesgue measure itself, which both have barycentre $\frac{1}{2}$, and we do indeed have $\mu_\beta \prec \nu_\beta$.

On the other hand, the *maximum* value of $G_{\nu,\mu}$ occurs at the value $\frac{2}{7}$, which is an atom of $\nu - \mu$, so no measure corresponding to μ can majorise the measure equivalently corresponding to ν (the measure $(\frac{9}{42}\delta_0 + \frac{1}{3}\delta_{\frac{2}{7}} + \frac{1}{3}\delta_{\frac{6}{7}} + \frac{5}{42}\delta_1)$ would do so, but it is not invariant under $T_{\frac{2}{7}}$).

Proof. Given measures $\mu, \nu \in \mathcal{M}_{T_0}, \mu_\beta, \nu_\beta \in \mathcal{M}_{T_\beta}$, define the signed measures Δ, Δ_β by $\Delta = \nu - \mu, \Delta_\beta = \nu_\beta - \mu_\beta$. The functions w_Δ and w_{Δ_β} can be defined exactly as for unsigned measures: $w_\Delta(x) = w_\nu(x) - w_\mu(x), w_{\Delta_\beta}(x) = w_{\nu_\beta}(x) - w_{\mu_\beta}(x)$, and $G_{\mu,\nu}$ as defined above can thus be expressed as

$$G_{\mu,\nu}(x) = w_\Delta(x) - x w_\Delta(1)\,. \tag{3.6}$$

Now, since

$$w_{\mu_\beta}(x) = \begin{cases} w_\mu(x+\beta) - w_\mu(\beta) - x\mu[0,\beta) & \text{if } x \le 1-\beta \\ w_\mu(x+\beta-1) - w_\mu(\beta) - x\mu[0,\beta) + w_\mu(1) + x + \beta - 1 & \text{if } x \ge 1-\beta \end{cases}$$

we have

$$w_{\Delta_\beta}(x) = \begin{cases} w_\Delta(x+\beta) - w_\Delta(\beta) - x\Delta([0,\beta)) & \text{if } x \le 1-\beta \\ w_\Delta(x+\beta-1) - w_\Delta(\beta) - x\Delta([0,\beta)) + w_\Delta(1) & \text{if } x \ge 1-\beta. \end{cases} \tag{3.7}$$

So, taking $x = 1$, we see that $w_{\Delta_\beta}(1) = w_\Delta(\beta) - w_\Delta(\beta) - \Delta([0,\beta)) + w_\Delta(1)$, and hence

$$\Delta([0,\beta)) = w_\Delta(1) - w_{\Delta_\beta}(1). \tag{3.8}$$

Rewriting (3.7) using this new formulation of $\Delta([0,\beta))$ gives

$$w_{\Delta_\beta}(x) = \begin{cases} w_\Delta(x+\beta) - w_\Delta(\beta) - x(w_\Delta(1) - w_{\Delta_\beta}(1)) & \text{if } x \le 1-\beta \\ w_\Delta(x+\beta-1) - w_\Delta(\beta) - x(w_\Delta(1) - w_{\Delta_\beta}(1)) + w_\Delta(1) & \text{if } x \ge 1-\beta \end{cases}$$

which can be rewritten in terms of $G_{\mu,\nu}$ as

$$w_{\Delta_\beta}(x) - x w_{\Delta_\beta}(1) = \begin{cases} G_{\mu,\nu}(x+\beta) - G_{\mu,\nu}(\beta) & \text{if } x \le 1-\beta \\ G_{\mu,\nu}(x+\beta-1) - G_{\mu,\nu}(\beta) & \text{if } x \ge 1-\beta \end{cases}$$

so if β is a minimum point for $G_{\mu,\nu}$ then $w_{\Delta_\beta}(x) - x w_{\Delta_\beta}(1) \ge 0$ for all x.

Also, $w_\Delta(x) = \int_0^x \Delta([0,s))ds$, and so if Δ is non-atomic at β then $\Delta([0,x))$ is continuous at β and $\frac{d}{dx} w_\Delta(x)\big|_\beta = \Delta([0,\beta))$. Now, by (3.6), $\frac{d}{dx} G_{\mu,\nu}(x)\big|_\beta = \frac{d}{dx} w_\Delta(x)\big|_\beta - w_\Delta(1)$, and hence if β is a local minimum for $G_{\mu,\nu}$ then $0 = \Delta([0,\beta)) - w_\Delta(1)$, and using this to rearrange (3.8) we get $w_{\Delta_\beta}(1) = 0$ and hence $w_{\Delta_\beta}(x) \ge 0$ for all x. And this is precisely the definition of $\mu_\beta \prec \nu_\beta$.

If a minimum β of $G_{\mu,\nu}$ is an atom of Δ then - since $G_{\mu,\nu}$ must be convex there - we must have $\mu(\{\beta\}) < \nu(\{\beta\})$. Here, we will not be able to differentiate w_Δ or $G_{\mu,\nu}$ at β. Instead of a corresponding T-invariant measure, by default there will be a measure supported on the closed interval $[0,1]$, of the form indicated in Example 3.3.2, that majorises μ. However, if there is an interval $(\beta, \beta+\epsilon)$ on which $\mu = \nu$ then that measure will in fact be T-invariant as it will assign no weight to the point 1, and hence be supported on $[0,1)$. $\square$

3.4 Majorisation of $T_{\frac{1}{2}}$-invariant measures

Since one measure can majorise another only if they have the same barycentre, we *might* expect that as the range of possible centres of mass of T_β-invariant measures gets smaller (as β gets closer to $\frac{1}{2}$), we would see more and more majorisation relations among them, as the range of possible barycentres gets smaller, and at $\beta = \frac{1}{2}$ all measures have centre of mass $\frac{1}{2}$ and are thus potential candidates for majorising one another. However, in fact we have the following rather surprising result:

Theorem 3.4.1. *Let μ, ν be distinct ergodic $T_{\frac{1}{2}}$-invariant Borel probability measures*[3]*. Then every $T_{\frac{1}{2}}$-invariant probability measure has barycentre equal to $1/2$, and $\mu \prec \nu$ if and only if $\mu = \delta_{\frac{1}{2}}$*

Proof of Theorem 3.4.1. Since μ, ν are ergodic and $\frac{1}{2}$ is a fixed point, either $\mu(\{\frac{1}{2}\}) = 0$ or $\mu = \delta_{\frac{1}{2}}$. Assume that $\mu(\{\frac{1}{2}\}) = \nu(\{\frac{1}{2}\}) = 0$.

Once again, recall that $\mu \prec \nu$ if and only if $w_\mu(x) \le w_\nu(x)$ for all x, and $w_\mu(1) = w_\nu(1)$. Now, since μ, ν are $T_{\frac{1}{2}}$-invariant, we have

$$\mu[0,s) = \begin{cases} \mu[\frac{1}{4}, \frac{1}{4} + \frac{s}{2}) + \mu[\frac{3}{4}, \frac{3}{4} + \frac{s}{2}) & \text{if } s \le \frac{1}{2} \\ \mu[\frac{1}{4}, \frac{1}{4} + \frac{s}{2}) + \mu[\frac{3}{4}, 1] + \mu[0, \frac{s}{2} - \frac{1}{4}) & \text{if } s \ge \frac{1}{2} \end{cases}$$

and so

[3] Note that in this section we shall omit the implicit subscript $\frac{1}{2}$ on the measures.

$$
\begin{aligned}
w_\mu(x) &= \int_0^x \mu[0,s)ds \\
&= \begin{cases} \int_0^x \mu[\frac{1}{4},\frac{1}{4}+\frac{s}{2}) + \mu[\frac{3}{4},\frac{3}{4}+\frac{s}{2})ds & \text{if } x \le \frac{1}{2} \\ \int_0^x \mu[\frac{1}{4},\frac{1}{4}+\frac{s}{2})ds + \int_0^{\frac{1}{2}} \mu[\frac{3}{4},\frac{3}{4}+\frac{s}{2})ds + \int_0^{x-\frac{1}{2}} \mu[\frac{3}{4},1) + \mu[0,\frac{s}{2})ds & \text{if } x \ge \frac{1}{2} \end{cases} \\
&= \begin{cases} 2w_\mu(\frac{x}{2}+\frac{1}{4}) + 2w_\mu(\frac{x}{2}+\frac{3}{4}) - 2w_\mu(\frac{1}{4}) - 2w_\mu(\frac{3}{4}) - x\left(\mu[0,\frac{1}{4}) + \mu[0,\frac{3}{4})\right) \\ \\ 2w_\mu(\frac{x}{2}+\frac{1}{4}) + 2w_\mu(\frac{x}{2}-\frac{1}{4}) - 2w_\mu(\frac{1}{4}) - 2w_\mu(\frac{3}{4}) - x\left(\mu[0,\frac{1}{4}) + \mu[0,\frac{3}{4})\right) \\ +2w_\mu(1) + x - \frac{1}{2}\,. \end{cases}
\end{aligned} \tag{3.9}
$$

Now, by, taking preimages we see that $\mu[0,\frac{1}{2}) = \mu[\frac{1}{4},\frac{1}{2}) + \mu[\frac{3}{4},1)$, and hence $\mu[0,\frac{1}{4}) + \mu[0,\frac{3}{4}) = 1$, and so in (3.9) taking $x = 1$ and rearranging proves that $w_\mu(1) = \frac{1}{2}$, for any $\mu \in \mathcal{M}_{T_{\frac{1}{2}}}$. Now, taking $x = \frac{1}{2}$ in (3.9) we get

$$
\begin{aligned}
w_\mu(\frac{1}{2}) &= 2w_\mu(\frac{1}{2}) + 2w_\mu(1) - 2w_\mu(\frac{1}{4}) - 2w_\mu(\frac{3}{4}) - \frac{1}{2} \\
&= 2w_\mu(\frac{1}{4}) + 2w_\mu(\frac{3}{4}) - \frac{1}{2}
\end{aligned}
$$

and so we can rewrite (3.9) as

$$
w_\mu(x) = \begin{cases} 2w_\mu(\frac{x}{2}+\frac{1}{4}) & + & 2w_\mu(\frac{x}{2}+\frac{3}{4}) & - & w_\mu(\frac{1}{2}) - \frac{1}{2} - x \\ 2w_\mu(\frac{x}{2}+\frac{1}{4}) & + & 2w_\mu(\frac{x}{2}-\frac{1}{4}) & - & w_\mu(\frac{1}{2}). \end{cases} \tag{3.10}
$$

Now, as in Section 3.3, let Δ be the signed measure $\nu - \mu$ and $w_\Delta(x) = w_\nu(x) - w_\mu(x)$, and define the function g by

$$
g(x) = w_\Delta(x) - w_\Delta(\frac{1}{2}) - (x - \frac{1}{2})\Delta[0,\frac{1}{2}].
$$

Notice that we have already proved that $w_\mu(1)$ is always equal to $\frac{1}{2}$ for any element of $\mathcal{M}_{T_{\frac{1}{2}}}$, and so $\mu \prec \nu$ if and only if $w_\Delta(x) \ge 0$ for all values of x.

Now, rewriting (3.10) in terms of w_Δ gives

$$
w_\Delta(x) = 2w_\Delta(\frac{x}{2}+\frac{1}{4}) + 2w_\Delta(\frac{x}{2}-\frac{1}{4} \bmod 1) - w_\Delta(\frac{1}{2})
$$

And for $|\epsilon| \le \frac{1}{2}$, setting $x = \frac{1}{2}+\epsilon$ gives

$$
w_\Delta(\frac{1}{2}+\epsilon) - 2w_\Delta(\frac{1}{2}+\frac{\epsilon}{2}) + w_\Delta(\frac{1}{2}) = 2w_\Delta(\frac{\epsilon}{2} \bmod 1)
$$

or, in terms of g,

$$\begin{aligned} g(\frac{1}{2}+\epsilon) &= 2g(\frac{1}{2}+\frac{\epsilon}{2})+2w_\Delta(\frac{\epsilon}{2} \bmod 1) \\ &= \lim_{n\to\infty} 2^n g(\frac{1}{2}+\frac{\epsilon}{2^n})+\sum_{i=1}^{\infty} 2^i w_\Delta(\frac{\epsilon}{2^i} \bmod 1) \end{aligned}$$

Now

$$\begin{aligned} 2^n g(\frac{1}{2}+\frac{\epsilon}{2^n}) &= 2^n \int_{\frac{1}{2}}^{\frac{1}{2}+\frac{\epsilon}{2^n}} \Delta[\frac{1}{2},x)dx \\ |2^n g(\frac{1}{2}+\frac{\epsilon}{2^n})| &\le \frac{\epsilon}{2^n} \sup_{x\epsilon[\frac{1}{2},\frac{1}{2}+\frac{\epsilon}{2^n}]} |\Delta[\frac{1}{2},x)| \end{aligned} \tag{3.11}$$

and therefore, since we have assumed that $\mu(\{\frac{1}{2}\}) = \nu(\{\frac{1}{2}\}) = \delta(\{\frac{1}{2}\}) = 0$, we see that $\lim_{n\to\infty} 2^n g(\frac{1}{2}+\frac{\epsilon}{2^n}) = 0$. And so $g(\frac{1}{2}+\epsilon) = \sum_{i=1}^{\infty} 2^i w_\Delta(\frac{\epsilon}{2^i} \bmod 1)$, and so if w_Δ is everywhere non-negative then so is g. But $w_\Delta(x) = w_\Delta(\frac{1}{2}) + (x-\frac{1}{2})\Delta[0,\frac{1}{2}] + g(x)$. So, using the fact that $w_\Delta(0) = w_\Delta(1) = 0$, we see that $\frac{1}{2}\Delta[0,\frac{1}{2}] = w_\Delta(\frac{1}{2}) + g(0) = -w_\Delta(\frac{1}{2}) - g(1)$.

So if w_Δ and g are both everywhere non-negative, we must have $\Delta[0,\frac{1}{2}] = w_\Delta(\frac{1}{2}) = g(0) = g(1) = 0$, and so from the definition of g we have $w_\Delta(x) = g(x)$ for all x. So, for $-\frac{1}{2} \le \epsilon \le \frac{1}{2}$,

$$w_\Delta(\epsilon+\frac{1}{2}) = \sum_{i=1}^{\infty} 2^i w_\Delta(\frac{\epsilon}{2^i} \bmod 1)\,.$$

So, taking $\epsilon = \frac{1}{2}$, we see that $w_\Delta(\frac{1}{2^n}) = 0$, and likewise taking $\epsilon = -\frac{1}{2}$ gives us $w_\Delta(1-\frac{1}{2^n}) = 0$. Then we can use $\epsilon = \frac{1}{2}-\frac{1}{2^n}, \frac{1}{2^n}-\frac{1}{2}$ to prove that $w_\Delta(x) = 0$ whenever $x = \frac{1}{2}-\frac{1}{2^n}+\frac{1}{2^{n+m}}$ or $x = \frac{1}{2}+\frac{1}{2^n}-\frac{1}{2^{n+m}}$

Repeating this procedure infinitely often, we see that $w_\Delta(x) = 0$ for all x of the form $\frac{k}{2^n}$, and hence for all x, as w_Δ is continuous. So if w_Δ is non-negative then $w_\Delta(x) = 0$, Δ is the zero measure and $\mu \equiv \nu$. So no ergodic element of $\mathcal{M}_{T_{\frac{1}{2}}}$ majorises any other such measure other than the fixed point measure $\delta_{\frac{1}{2}}$.

Conversely, since we have shown that all measures in $\mathcal{M}_{T_{\frac{1}{2}}}$ have $w_\mu(1) = \frac{1}{2}$, all measures majorise the fixed point measure.

□

Chapter 4

Periodic orbits of the Doubling Map

In this chapter, we will demonstrate which possible cyclic orderings can arise as periodic cycles of the doubling map $T : x \mapsto 2x \pmod 1$ on the unit interval. We will then use this result to show the complete structure of the majorisation partial ordering on the sets of T-invariant measures supported on periodic orbits of weight[1] 2 and weight 3, of any given length.

In [19], Jenkinson proved that every probability measure on I invariant under T majorises a member of a class of measures known as *Sturmian measures*. A great deal has been written about this class of measures; see for example [6] for a fuller treatment of their properties. Each closed subinterval of I of length $\frac{1}{2}$ supports a unique T-invariant probability measure; these are the Sturmian measures. For every $\varrho \in I$ there exists a unique Sturmian measure with centre of mass ϱ. This means that for every concave function there exists a maximising Sturmian measure (there may also exist non-Sturmian maximising measures).

If we treat T as acting on $[0, 1]$ instead of on $[0, 1)$, with $T(1) = 1$, then the minimising measure must be one of the two fixed point measures at the two end-points. On $[0, 1)$, δ_0 is still a fixed point measure, while the fixed point at 1 can be approximated by the sequence of measures supported on periodic orbits with binary expansions $01^n01^n\ldots$, so a minimising measure for a concave function f exists if and only if $f(0) \leq f(1)$.

In this chapter, we will examine how majorisation ordering of T-invariant measures behaves between these two endpoints by constructing the complete majorisation partial orderings on the set of measures supported on periodic orbits of a given length and of weight 2 (that is to say, periodic orbits of length l whose points have precisely two 1s in the first l digits of their binary expansion), and the set of measures supported on periodic orbits of a given length and of weight 3. To do this, we will first

[1] The *weight* of a periodic orbit is the number of 1's in its symbolic representation, cf. Definition 4.2.1.

examine the possible cyclic orderings of periodic orbits under T.

4.1 Possible orderings of elements of periodic orbits of T

Definition 4.1.1. Let $x \in I$ have binary expansion $0.x_0x_1x_2\ldots$ (the ambiguity about $0.1\dot{0} = 0.0\dot{1}$ can be safely ignored). Then the *gap sequence* of x is defined to be the sequence of g_i (each of which is either a natural number or infinite) such that $x_k = 1$ if and only if $k = g_1 + g_2 \ldots + g_n$ for some n. The gap sequence is finite if x is dyadic (that is, of form $\frac{a}{2^k}$), recurring for any other rational x, and non-recurring if x is irrational.

Example 4.1.2. We list various binary expansions and their corresponding gap sequences. Note that the notation $a\dot{b}$ denotes a followed by b recurring, while $a\dot{b}c\dot{d}$ denotes a followed by bcd recurring.

Binary expansion of x	Gap sequence of x
$00101000011\dot{0}$	3251∞
$\dot{0}\dot{1}$	$\dot{2}$
$\dot{1}\dot{0}$	$1\dot{2}$
$110010101101\ldots$	$1232212\ldots$

Theorem 4.1.3. *Let $x_0 \in I$ be a periodic point of least period q, with gap sequence $\dot{g}_1, g_2, \ldots, \dot{g}_n$, and let $T^i(x_0) = x_i$ for $i = 0, 1, \ldots, p-1$. Then we can construct the cyclic ordering of $\{x_0, x_1 \ldots, x_{p-1}\}$ as follows:*

1. *Construct an ordering $\{\pi_1, \pi_2, \ldots, \pi_n\}$ of the numbers $\{1, 2, \ldots, n\}$ by following the rule $a < b$ if $g_{a+i} > g_{b+i}$, where i is the smallest value such that $g_{a+i} \neq g_{b+i}$*

2. *Construct a pair of functions $\lambda, \nu : \{0, 1 \ldots, p-1\} \to \mathbb{N}$ as follows: if $k = g_1 + g_2 + \ldots + g_j - i$, with $1 \le i \le g_j$, then $\lambda(k) = i$ and $\nu(k) = j$.*

   These functions have the properties that $\lambda(k)$ is the position of the first 1 in the binary expansion of x_k, while among those x_k whose binary expansion first has a 1 in position j, the ordering is given by $\pi_{\nu(k)}$.

3. *Let G be the largest of the g_i. Then the ordering of the x_k will be given by:*

   - *If $\lambda(a) < \lambda(b)$ then $x_b < x_a$.*
   - *If $\lambda(a) = \lambda(b)$ and $\pi_{\nu(a)} < \pi_{\nu(b)}$ then $x_a < x_b$.*

We illustrate Theorem 4.1.3 with the following example:

Example 4.1.4. Let x have binary expansion $\dot{0}0101000100001000\dot{1}$, and hence gap series $\dot{3}245\dot{4}$

The largest gap is $g_4 = 5$, so $\pi_1 = 4 - 1 = 3$. Then $g_3 = g_5 = 4$, but $g_4 > g_1$ so $\pi_2 = 3 - 1 = 2$ and $\pi_3 = 5 - 1 = 4$. The next largest gap is $g_1 = 3$, and so $\pi_3 = 5$ and $\pi_5 = 1$. So π is the order 32451.

Now the functions λ, ν are as follows:

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
$\lambda(k)$	3	2	1	2	1	4	3	2	1	5	4	3	2	1	4	3	2	1
$\nu(k)$	1	1	1	2	2	3	3	3	3	4	4	4	4	4	5	5	5	5

and the ordering of the x_k is as follows:

k	9	5	10	14	6	11	15	0	7	3	12	16	1	8	4	13	17	2
$\lambda(k)$	5	4	4	4	3	3	3	3	2	2	2	2	2	1	1	1	1	1
$\nu(k)$	4	3	4	5	3	4	5	1	3	2	4	5	1	3	2	4	5	1

And it is easy to check this - observe, for example, that e.g. $T^9(x)$, which has the binary expansion $\dot{0}0001000100101000\dot{1}$, is the smallest point in the orbit of x.

Definition 4.1.5. Let $\{\pi_0, \pi_1, \dots, \pi_{n-1}\}$ be a permutation of $\{0, 1, \dots, n-1\}$. Write $a \to b$ to mean that a comes before b in the ordering $\{\pi_0, \pi_1, \dots, \pi_{n-1}\}$. We say that $k, k+1, \dots, k+m$ is a *rising sequence* of $\{\pi_0, \pi_1, \dots, \pi_{n-1}\}$ if:

- $k \to k+1 \to \dots \to k+m$
- $k \to k-1$ or $k = 0$
- $k+m+1 \to k+m$ or $k+m = n-1$.

So, for example, in the sequence 0457923618, the rising sequences are 01, 23, 456, 78 and 9.

Theorem 4.1.6. *If $x \in I$ is a periodic point with gap sequence $\dot{g}_1, g_2, \dots, \dot{g}_n$, and $\{\pi_0, \pi_1, \dots, \pi_{n-1}\}$ is the ordering of $x, T(x), \dots, T^{p-1}(x)$, then the rising sequences of $\{\pi_0, \pi_1, \dots, \pi_{n-1}\}$ are precisely $\{0, 1, \dots, g_1 - 1\}; \{g_1, g_1 + 1, \dots, g_2 - 1\}, \dots, \{g_{n-1} \dots, g_n\}$.*

Proof. This follows directly from the construction given in Theorem 4.1.3; the rising sequences will be precisely the sets of constant ν. $\square$

Remark 4.1.7. It follows that it is easy to tell whether or not a given permutation of the numbers $\{0, 1, \ldots, n-1\}$ can arise as the ordering of the elements of a periodic orbit of T: simply split it up into rising sequences, and see whether or not it is the ordering given by applying the construction in Theorem 4.1.3 to the number x with gap sequence given by the lengths of the rising sequences.

Now, armed with these methods, we will go on to look at majorisation ordering of orbits.

Remark 4.1.8. For the set of words of any given length and weight it is possible to work out the majorisation ordering using a computer, although for higher lengths the amount of computing power required becomes problematic.

4.2 Majorisation ordering of periodic orbits of weight 2

Definition 4.2.1. Let x be a periodic point of T. We say that x is a periodic point of length l and weight w if $T^l(x) = x$ and the binary expansion of x has precisely w 1s in the first l digits.

Note that l need not be the least period of x, so that for example $\frac{1}{3}$, with binary expansion $\dot{0}\dot{1}$, is a periodic point of length 2 and weight 1, and also a periodic point of length 4 and weight 2, and so on.

From now on, let $\underline{g_1 g_2 \ldots g_n}$ denote the number with gap sequence $\dot{g}_1 g_2 \ldots \dot{g}_n$, a periodic point of length $g_1 + g_2 + \ldots + g_n$ and weight n.

Lemma 4.2.2. *The centre of mass of a periodic orbit of length l and weight w is precisely $\frac{w}{l}$. As such, a necessary condition for one periodic orbit to majorise another is that $\frac{l_1}{l_2} = \frac{w_1}{w_2}$, where l_i and w_i are the lengths and weights of the two measures.*

Proof. Note that if μ is the orbit supported on the orbit of x then $\mu[\frac{1}{2}, 1] = \frac{w}{l}$. Now by referring to the definition of majorisation, and taking preimages under T, we see that $w_\mu(1) = \int_0^1 \mu[0, s)ds = \int_0^1 \mu[0, \frac{s}{2}) + \mu[\frac{1}{2}, \frac{1}{2} + \frac{s}{2})ds = 2w_\mu 1 - \mu[0, \frac{1}{2}) = \mu[0, \frac{1}{2})$, and so $\varrho(\mu) = 1 - w_\mu(1) = \mu[\frac{1}{2}, 1) = w/l$. So one measure can majorise another only if they have the same centre of mass. $\square$

So if we are looking for a complete majorisation ordering on periodic orbits of the doubling map, it suffices to look at the ordering among orbits of each given weight and length (it is possible to compare e.g. an orbit of length 4, weight 2 with one of length 6, weight 3 by looking at them both as orbits of length 12, weight 6).

For high weights, this rapidly becomes extremely complicated, but for weights 2 and 3 we are able to give a complete picture of the ordering for any length.

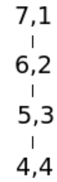

Figure 4.1: Hasse diagram of majorisation ordering on gap sequences of words of weight 2, length 8

Notation 4.2.3. For convenience, we will write x majorises y or $x \succ y$ to mean the (T-invariant Borel probability) measure supported upon the orbit of periodic point x majorises the measure supported upon the orbit of y.

Theorem 4.2.4.

$$\underline{l-1,1} \succ \underline{l-2,2} \succ \ldots \succ \underline{\lceil l/2 \rceil, \lfloor l/2 \rfloor}$$

Figure 4.1 illustrates this at length 8.

Example 4.2.5. For example, the gap sequence 31 corresponds to the binary expansion 0011 and the point $\frac{1}{5}$ with orbit $\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{3}{5} \ldots$, while the binary gap sequence 22 corresponds to the binary expansion 0101 and the point $\frac{1}{3}$ with orbit $\frac{1}{3}, \frac{2}{3}, \frac{1}{3}, \frac{2}{3} \ldots$. And $\frac{1}{4}\left(\delta_{\frac{1}{5}} + \delta_{\frac{2}{5}} + \delta_{\frac{3}{5}} + \delta_{\frac{4}{5}}\right) \succ \frac{1}{2}\left(\delta_{\frac{1}{3}} + \delta_{\frac{2}{3}}\right)$.

Proof. To prove this, we will use the following lemma:

Lemma 4.2.6. *If x and y are periodic points of (not necessarily least) period l then for $0 \leq i \leq l-1$, define $\alpha_i(x)$ to be the permutation of $0, 1, \ldots, l-1$ such that $T^{\alpha_0}(x) \leq T^{\alpha_1}(x) \leq \ldots \leq T^{\alpha_{l-1}}(x)$. Then $x \prec y$ if and only if $\sum_{i=0}^{n-1} T^{\alpha_i}(x) \leq \sum_{i=0}^{n-1} T^{\alpha_i}(y)$ for $0 \leq n \leq l-1$, with equality when $n = l$.*

Proof. To see this, simply note that if μ_x is the measure supported on the orbit of x then $w_{\mu_x}(z) = \sum_{T^{\alpha_i}(x) < z}(z - T^{\alpha_i}(x))$, and rewrite the definition of majorisation using this formula. □

Now, consider two integers such that $g_1 \geq g_2 \geq 2$.

If we let $x = \underline{g_1, g_2}$ then

$$x = \frac{2^{g_2} + 1}{2^{g_1+g_2} - 1}$$

and

$$T^n(x) = \begin{cases} \frac{2^{g_2+n} + 2^n}{2^{g_1+g_2}-1} & \text{if} \quad 0 \leq n \leq g_1 - 1 \\ \frac{2^{n-g_1} + 2^n}{2^{g_1+g_2}-1} & \text{if} \quad g_1 \leq n \leq g_1 + g_2. \end{cases}$$

So if we define $\alpha_i(x)$ as above then

$$\alpha_i = \begin{cases} i & \text{if} \quad i \leq g_1 - g_2 - 1 \\ \frac{i}{2} + \frac{g_1}{2} + \frac{g_2}{2} & \text{if} \quad \begin{cases} i \geq g_1 - g_2 + 1 \\ \text{and } i = g_1 - g_2 + 1 \pmod 2 \end{cases} \\ \frac{i}{2} + \frac{g_1}{2} - \frac{g_2}{2} - \frac{1}{2} & \text{if} \quad \begin{cases} i \geq g_1 - g_2 + 1 \\ \text{and } i = g_1 - g_2 \pmod 2 \end{cases} \end{cases}$$

$$(2^{g_1+g_2} - 1)T^{\alpha_i}(x) = \begin{cases} 2^{i+g_2} + 2^i & \text{if} \quad i \leq g_1 - g_2 \\ 2^{\frac{i+g_1+g_2}{2}} + 2^{\frac{i-g_1+g_2}{2}} & \text{if} \quad \begin{cases} i \geq g_1 - g_2 + 1 \\ i = g_1 - g_2 + 1 \pmod 2 \end{cases} \\ 2^{\frac{i+g_1+g_2-1}{2}} + 2^{\frac{i+g_1-g_2-1}{2}} & \text{if} \quad \begin{cases} i \geq g_1 - g_2 + 1 \\ i = g_1 - g_2 \pmod 2 \end{cases} \end{cases}$$

$$(2^{g_1+g_2}-1)\sum_{i=0}^{n-1} T^{\alpha_i}(x) = \begin{cases} (2^{g_2} + 1)(2^n - 1) & \text{if} \quad n \leq g_1 - g_2 \\ 2^{\frac{n+g_1+g_2}{2}} + 2^{\frac{n-g_1+g_2}{2}} & \text{if} \quad \begin{cases} n \geq g_1 - g_2 + 1 \\ n = g_1 - g_2 + 1 \pmod 2 \end{cases} \\ 2^{\frac{n+g_1+g_2-1}{2}} + 2^{\frac{n+g_1-g_2-1}{2}} & \text{if} \quad \begin{cases} n \geq g_1 - g_2 + 1 \\ n = g_1 - g_2 \pmod 2). \end{cases} \end{cases}$$

If we now denote $\underline{g_1 + 1, g_2 - 1}$ by $\hat{x}$ then

$$(2^{g_1+g_2}-1)\sum_{i=0}^{n-1} T^{\alpha_i}(\hat{x}) = \begin{cases} (2^{g_2-1} + 1)(2^n - 1) & \text{if} \quad n \leq g_1 - g_2 + 2 \\ 2^{\frac{n+g_1+g_2}{2}} + 2^{\frac{n-g_1+g_2-2}{2}} & \text{if} \quad \begin{cases} n \geq g_1 - g_2 + 3 \\ n = g_1 - g_2 + 1 \pmod 2 \end{cases} \\ 2^{\frac{n+g_1+g_2-1}{2}} + 2^{\frac{n+g_1-g_2+1}{2}} & \text{if} \quad \begin{cases} n \geq g_1 - g_2 + 3 \\ n = g_1 - g_2 \pmod 2), \end{cases} \end{cases}$$

and so, comparing termwise, we see that $\sum_{i=0}^{n-1} T^{\alpha_i}(\hat{x}) \leq \sum_{i=0}^{n-1} T^{\alpha_i}(x)$ with equality occuring when $n = g_1 + g_2$.

So $\hat{x} \succ x$. And so the complete ordering of measures supported on points of weight 2 follows directly.

$\square$

4.3 Words of Weight 3

Theorem 4.3.1. *Given a natural number $l > 3$, let $\mathcal{M}_l$ be the set of all ordered triples of integers (a, b, c) such that $x = \underline{a, b, c}$ is periodic under T with (not necessarily least) period l and weight 3, and is minimal in its own orbit.*

Then we can construct the poset of majorisation ordering on $\mathcal{M}_l$ from the diagram in Figure 4.2.

Remark 4.3.2. Precisely, the way in which Figure 4.2 yields the majorisation ordering on $\mathcal{M}_l$ is as follows: read each triple of numbers as the gap sequence of the binary expansion of a periodic point of period l - for example, if $l = 6$ then $l-3, 1, 2$ corresponds to the point with gap sequence $\dot{3}1\dot{2}$, binary expansion $\dot{0}0110\dot{1}$, which is to say 13/63. Restrict ourselves to the set of points that are minimal in their own orbits - so for example, at weight 6 we will ignore the entry $l-4, 3, 1$, because it corresponds to the orbit of $\underline{231} = \dot{0}1001\dot{1} = 19/63$, which contains the point $\underline{31} = \dot{0}0110\dot{1} = 13/63$. Then the lines in the diagram are precisely the majorisation ordering, with the entry higher up/further to the left majorising the other.

Note that we *do* include routes that go via points not in our diagram. For example, at $l = 6$ we have $\underline{312} \succ \underline{222}$, even though $\underline{231}$ is not minimal in its orbit.

Figure 4.3 illustrates the diagrams that this generates at $l = 6$, 7, 8, 9, 10 and 11.

Proof. To prove that the diagram in Figure 4.2 is both sufficient and necessary, it suffices to prove the following 7 lemmas:

Lemma 4.3.3. *Let $a \geq b > c$. Then $\underline{a, b, c} \succ \underline{a, c+1, b-1}$.*

This lemma gives us the horizontal lines in Figure 4.2.

Lemma 4.3.4. *Let $a \geq b \geq c+2$. Then $\underline{a, b, c} \succ \underline{a, b-1, c+1}$*

This lemma gives us the curved horizontal lines in Figure 4.2.

Lemma 4.3.5. *Let $a - 2 \geq b \geq c$. Then $\underline{a, b, c} \succ \underline{a-1, b+1, c}$*

This lemma gives us the vertical lines in the even columns of Figure 4.2.

Lemma 4.3.6. *Let $a - 2 \geq c \geq b$. Then $\underline{a, b, c} \succ \underline{a-1, b, c+1}$*

This lemma gives us the vertical lines in the odd columns of Figure 4.2.

Lemma 4.3.7. *Let $a - 2 \geq c \geq b$. Then $\underline{a, b, c} \succ \underline{a-1, c+1, b}$*

This lemma gives us the diagonal lines in Figure 4.2.

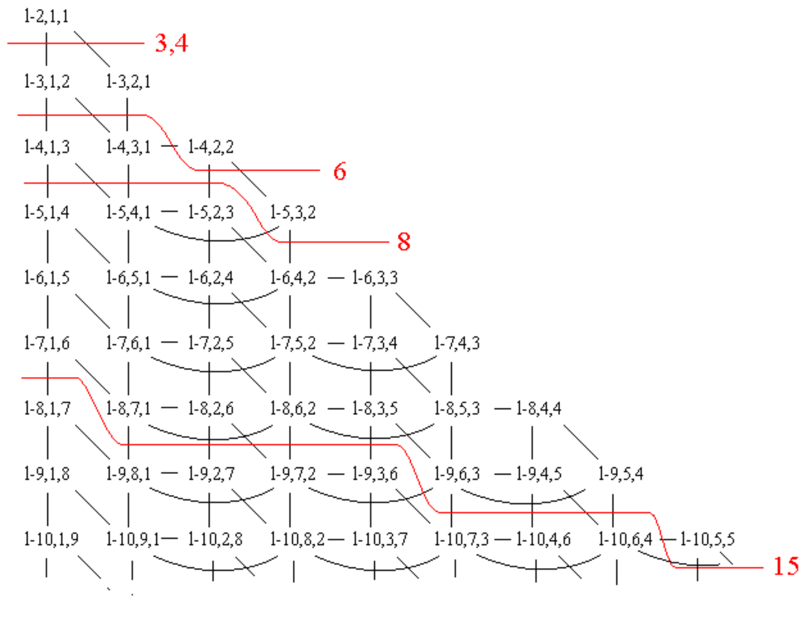

Figure 4.2: The Hasse diagram for majorisation ordering of words of weight 3 at a given length l, indexed by gap sequence as in Figure 4.3, can be constructed from this infinite diagram simply by restricting it to the set of gap sequences that correspond to points that are minimal in their own orbit at length l. For example, the Hasse diagrams for majorisation at weights 3,4,6,8 and 15 can be generated by taking the entries above the red lines indicated, and replacing "l" with the relevant value.

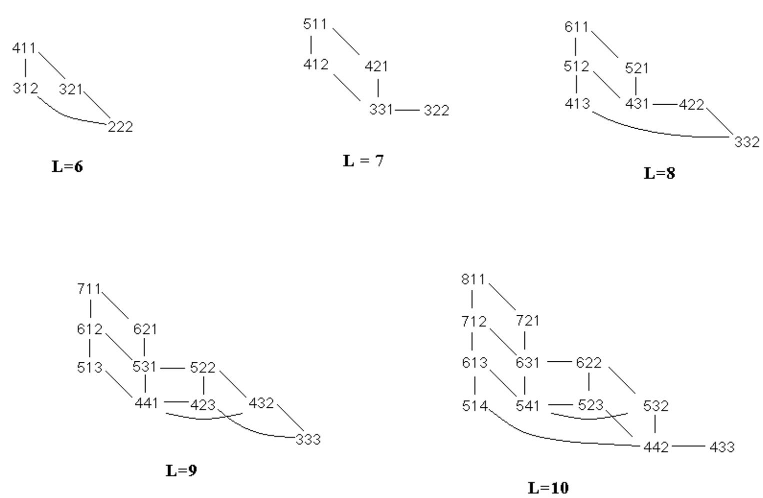

Figure 4.3: Hasse diagrams of majorisation ordering of periodic orbits of weight 3, lengths 6-10, denoted by the gap sequences of their minimal elements. Lines indicate majorisation; the entry above and/or to the left majorises the other. So, for example, the line linking 411 to 321 means that the orbit of $7/63$ (binary expansion $\dot{0}0011\dot{1}$) majorises the orbit of $11/63$ (binary expansion $\dot{0}0101\dot{1}$).

Lemma 4.3.8. *Let* $a_1 + b_1 + c_1 = a_2 + b_2 + c_2$ *with* $a_1 \geq b_1 \geq c_1, a_2 \geq b_2 \geq c_2$ *and* $a_1 > a_2, c_1 < c_2$. *Then*[2] $\left.\begin{array}{c} a_1b_1c_1 \\ a_1c_1b_1 \end{array}\right\} \begin{array}{c} \not\succ \\ \not\prec \end{array} \left\{\begin{array}{c} a_2b_2c_2 \\ a_2c_2b_2 \end{array}\right.$

Lemma 4.3.9. *Let* $a + b_1 + c_1 = a + b_2 + c_2$ *and* $a \geq c_1 \geq b_1; a \geq c_2 \geq b_2$, *and if* $a = c_i$ *then require that* $b_i < a, c_i$. *Then* $\underline{a, b_1, c_1} \overset{\not\succ}{\underset{\not\prec}{}} \underline{a, b_2, c_2}$.

Lemmas 4.3.3-4.3.7, between them, prove that all the orderings in the Hasse diagram occur, while Lemmas 4.3.8-4.3.9 prove that no others do.

Notation 4.3.10. For any three integers g_1, g_2, g_3 such that $g_1 \geq g_2, g_3$ and such that if $g_1 = g_3$ then $g_1 = g_2 = g_3$, and for $i = 1, 2, 3$, with $x = \underline{g_1, g_2, g_3}$, we define

$$\begin{aligned} l &= g_1 + g_2 + g_3 \\ \lambda_i &= 2^{g_{i+1}+g_{i+2}} + 2^{g_{i+2}} + 1 \\ d_i &= l - 3g_i \\ d_* &= \max(d_2, d_3) = l - 3\min(g_2, g_3) \\ \mu_i &= g_1 - g_i \\ \mu_* &= \min(\mu_2, \mu_3) = g_1 - \max(g_2, g_3) \end{aligned}$$

Where necessary, we will write these as e.g. $\lambda_1(x), d_2(y)$ to indicate whether we are refering to the gaps of x or of y as the g_i, but where possible we will omit this.

Example 4.3.11. For example, if x has the gap sequence $6, 3, 2$ then $l = 11, \lambda_1 = 2^{3+2} + 2^2 + 1 = 37, \lambda_2 = 2^{2+6} + 2^6 + 1 = 197, \lambda_3 = 2^{6+3} + 2^3 + 1 = 265, d_1 = 11 - 18 = -7, d_2 = 2, d_3 = 5, d_* = d_3 = 5, \mu_1 = 0, \mu_2 = 3, \mu_3 = 4$ and $\mu_* = 3$.

Now, as before, define $\alpha_k(x)$ to be the permutation of $0, 1 \ldots, n-1$ such that $T^{\alpha_0(x)}(x) \leq T^{\alpha_1(x)}(x) \ldots \leq T^{\alpha_{n-1}(x)}(x)$. The $\alpha_k(x)$ are uniquely defined if l is the least period of x, and $T^{\alpha_k(x)}(x)$ is always uniquely defined.

If $g_2 \leq g_3$ then we say that $x = \underline{g_1, g_2, g_3}$ is of type I, while if $g_2 > g_3$ we say that x is of type II.

Then, using the results of Section 4.2, we can construct the following tables, which will enable us to explicitly compute the values of $\sum_{i=0}^{k} T^{\alpha_i}(x)$ for a given x, and hence test explicitly whether $x \prec y$:

[2] The symbols $\not\succ$ and $\not\prec$ together denote that elements are incomparable with respect to the majorisation partial order.

i	α_i if x is type I	α_i if x is type II
$0 \le i < \mu_*$	i	i
$\mu_* \le i \le d_*$, $i \equiv \mu_* \pmod 2$	$\frac{i+g_1+2g_2+g_3}{2}$	$\frac{i+g_1-g_2}{2}$
$\mu_* \le i \le d_*$, $i \equiv \mu_*+1 \pmod 2$	$\frac{i+g_1-g_3-1}{2}$	$\frac{i+g_1+g_2-1}{2}$
$d_* \le i < l$, $i \equiv d_* \pmod 3$	$\frac{i+2g_1+2g_2+2g_3}{3}$	$\frac{i+2g_1+2g_2+2g_3}{3}$
$d_* \le i < l$, $i \equiv d_*+1 \pmod 3$	$\frac{i+2g_1+2g_2-g_3-1}{3}$	$\frac{i+2g_1-g_2-g_3-1}{3}$
$d_* \le i < l$, $i \equiv d_*+2 \pmod 3$	$\frac{i+2g_1-g_2-g_3-2}{3}$	$\frac{i+2g_1+2g_2-g_3-2}{3}$

i	$T^{\alpha_i}(x) \times (2^l-1)$ if x is type I	$T^{\alpha_i}(x) \times (2^l-1)$ if x is type II
$0 \le i < \mu_*$	$2^i\lambda_1$	$2^i\lambda_1$
$\mu_* \le i \le d_*$, $i \equiv \mu_* \pmod 2$	$2^{\frac{i-\mu_3}{2}}\lambda_3$	$2^{\frac{i+\mu_2}{2}}\lambda_1$
$\mu_* \le i \le d_*$, $i \equiv \mu_*+1 \pmod 2$	$2^{\frac{i+\mu_3-1}{2}}\lambda_1$	$2^{\frac{i-\mu_2-1}{2}}\lambda_2$
$d_* \le i < l$, $i \equiv d_* \pmod 3$	$2^{\frac{i-d_3}{3}}\lambda_3$	$2^{\frac{i-d_3}{3}}\lambda_3$
$d_* \le i < l$, $i \equiv d_*+1 \pmod 3$	$2^{\frac{i-d_2-1}{3}}\lambda_2$	$2^{\frac{i-d_1-1}{3}}\lambda_1$
$d_* \le i < l$, $i \equiv d_*+2 \pmod 3$	$2^{\frac{i-d_1-2}{3}}\lambda_1$	$2^{\frac{i-d_2-2}{3}}\lambda_2$

k	$\sum_{i=0}^{k} T^{\alpha_i}(x) \times (2^l - 1)$ if x is type I
$k < \mu_3$	$(2^{k+1} - 1)\lambda_1$
$\mu_3 \leq k \leq d_2$ $k \equiv \mu_3 \pmod 2$	$\left(2^{\frac{k+\mu_3}{2}} - 1\right)\lambda_1 + \left(2^{\frac{k-\mu_3+2}{2}} - 1\right)\lambda_3$
$\mu_3 \leq k \leq d_2$ $k \equiv \mu_3 + 1 \pmod 2$	$\left(2^{\frac{k+\mu_3+1}{2}} - 1\right)\lambda_1 + \left(2^{\frac{k-\mu_3+1}{2}} - 1\right)\lambda_3$
$d_2 \leq k < l$ $k \equiv d_2 \pmod 3$	$\left(2^{\frac{k-d_1}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+3}{3}} - 1\right)\lambda_3$
$d_2 \leq k < l$ $k \equiv d_2 + 1 \pmod 3$	$\left(2^{\frac{k-d_1-1}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2+2}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+2}{3}} - 1\right)\lambda_3$
$d_2 \leq k < l$ $k \equiv d_2 + 2 \pmod 3$	$\left(2^{\frac{k-d_1+1}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2+1}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+1}{3}} - 1\right)\lambda_3$

k	$\sum_{i=0}^{k} T^{\alpha_i}(x) \times (2^l - 1)$ if x is type II
$k < \mu_2$	$(2^{k+1} - 1)\lambda_1$
$\mu_2 \leq k \leq d_3$ $k \equiv \mu_2 \pmod 2$	$\left(2^{\frac{k+\mu_2}{2}} - 1\right)\lambda_1 + \left(2^{\frac{k-\mu_2}{2}} - 1\right)\lambda_2$
$\mu_2 \leq k \leq d_3$ $k \equiv \mu_2 + 1 \pmod 2$	$\left(2^{\frac{k+\mu_2+1}{2}} - 1\right)\lambda_1 + \left(2^{\frac{k-\mu_2+1}{2}} - 1\right)\lambda_2$
$d_3 \leq k < l$ $k \equiv d_3 \pmod 3$	$\left(2^{\frac{k-d_1}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+3}{3}} - 1\right)\lambda_3$
$d_3 \leq k < l$ $k \equiv d_3 + 1 \pmod 3$	$\left(2^{\frac{k-d_1+2}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2-1}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+2}{3}} - 1\right)\lambda_3$
$d_3 \leq k < l$ $k \equiv d_3 + 2 \pmod 3$	$\left(2^{\frac{k-d_1+1}{3}} - 1\right)\lambda_1 + \left(2^{\frac{k-d_2+1}{2}} - 1\right)\lambda_2 + \left(2^{\frac{k-d_3+1}{3}} - 1\right)\lambda_3$

Proof of lemma 4.3.3. Let $x = \underline{a, b, c}$ and $y = \underline{a, c+1, b-1}$. If $b = c+1$ then $x = y$, so we can assume that $b \geq c+2$. To avoid having to repeatedly divide all terms by $2^l - 1$, we will work with $\sum_{i=0}^{k-1} T^{\alpha_i(x)}(x) \times (2^l - 1)$, which we will write $S_k(x)$, and endeavour to show that $S_k(x) \leq S_k(y)$ for $k = 0, 1 \ldots, l-1$. Now x is type II and y is type I, and so the proof will follow from handling separately the following 10 cases:

Case (i): $k < a - b$:
$S_x(k) = (2^{k+1} - 1)\lambda_1(x)$, $S_y(k) = (2^{k+1} - 1)\lambda_1(y)$, and $\lambda_1(x) < \lambda_1(y)$ as $c < b - 1$.

Case (ii): $k = a - b$:
$S_x(k) = (2^{a-b+1} - 1)\lambda_1(x)$, $S_y(k) = (2^{a-b+1} - 1)\lambda_1(y)$, and $\lambda_1(x) < \lambda_1(y)$ as $c < b - 1$.

Case (iii): $a - b + 1 \leq k < a + b - 2c - 3$ and and $k = a - b \pmod 2$:
$S_x(k) = \left(2^{\frac{k+a-b+2}{2}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-a+b}{2}} - 1\right)\lambda_2(x)$
$S_y(k) = \left(2^{\frac{k+a-b+2}{2}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-a+b}{2}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^{\frac{k+a-b+2}{2}} - 1\right)\left(2^{b-1} - 2^c\right) + \left(2^{\frac{k-a+b}{2}} - 1\right)\left(2^{a+c} + 2^{c+1} - 2^a\right) > 0$

Case (iv): $a - b + 1 \leq k < a + b - 2c - 3$ and $k = a - b + 1 \pmod 2$:
$S_x(k) = \left(2^{\frac{k+a-b+1}{2}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-a+b+1}{2}} - 1\right)\lambda_2(x)$
$S_y(k) = \left(2^{\frac{k+a-b+2}{2}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-a+b+1}{2}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^{\frac{k+a-b+1}{2}} - 1\right)\left(2^{b-1} - 2^c\right) + \left(2^{\frac{k-a+b+1}{2}} - 1\right)\left(2^{a+c} + 2^{c+1} - 2^a\right) > 0$

Case (v): $k = a + b - 2c - 3$:
$S_x(k) = \left(2^{a-c-1} - 1\right)\lambda_1(x) + \left(2^{b-c-1} - 1\right)\lambda_2(x)$
$S_y(k) = \left(2^{a-c-1} - 1\right)\lambda_1(y) + \left(2^{b-c-1} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^{a-c-1} - 1\right)\left(2^{b-1} - 2^c\right) + \left(2^{b-c-1} - 1\right)\left(2^{a+c} + 2^{c+1} - 2^a\right) > 0$

Case (vi): $k = a + b - 2c - 2$:
$S_x(k) = \left(2^{a-c} - 1\right)\lambda_1(x) + \left(2^{b-c-1} - 1\right)\lambda_2(x)$
$S_y(k) = \left(2^{a-c-1} - 1\right)\lambda_1(y) + \lambda_2(y) + \left(2^{b-c-1} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^{a-c-1} - 1\right)\left(2^{b-1} - 2^c\right) + \left(2^{b-c-1} - 1\right)\left(2^{a+c} + 2^{c+1} - 2^a\right) + 2^{a-1} + 1 - 2^{a-c-1} > 0$

Case (vii): $k = a + b - 2c - 1$:
$S_x(k) = \left(2^{a-c} - 1\right)\lambda_1(x) + \left(2^{b-c} - 1\right)\lambda_2(x)$
$S_y(k) = \left(2^{a-c} - 1\right)\lambda_1(y) + \lambda_2(y) + \left(2^{b-c-1} - 1\right)\lambda_3(y)$

$S_y(k) - S_x(k) = (2^{a-c}-1)\left(2^{b-1}-2^c\right) + \left(2^{b-c-1}-1\right)(2^{a+c}+2^{c+1}-2^a) + (2^a+1)\left(2^{b-c-1}-1\right) > 0$

Case $(viii)$: $a+b-2c \le k < l$ and $k = a+b-2c \pmod 3$:

$S_x(k) = \left(2^{\frac{k-b-c+2a}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c}{3}}-1\right)\lambda_3(x)$

$S_y(k) = \left(2^{\frac{k-b-c+2a}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c}{3}}-1\right)\lambda_3(y)$

$S_y(k) - S_x(k) = 2^{\frac{k-a+2b+2c}{3}} + 2^{\frac{k+2a-b+2c}{3}} + 2^{\frac{k+2a+2b-c}{3}} + 2^{a+b-1} - 2^{a+c} - 2^b - 2^c > 0$

Case (ix): $a+b-2c \le k < l$ and $k = a+b-2c+1 \pmod 3$:

$S_x(k) = \left(2^{\frac{k-b-c+2a+2}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b-1}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+2}{3}}-1\right)\lambda_3(x)$

$S_y(k) = \left(2^{\frac{k-b-c+2a-1}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b-1}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+5}{3}}-1\right)\lambda_3(y)$

$S_y(k) - S_x(k) = 2^{\frac{k-a-b+2c-1}{3}}\left(5.2^a + 3.2^{a+b-1} + 2^{a-c} + 2^{a+b-c-1} + 1\right) - 2^{a+b-1} + 2^{a+c} + 2^{a+b-1} - 2^{b-1} > 0$

Case (x): $a+b-2c \le k < l$ and $k = a+b-2c+2 \pmod 3$:

$S_x(k) = \left(2^{\frac{k-b-c+2a+1}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b+1}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+1}{3}}-1\right)\lambda_3(x)$

$S_y(k) = \left(2^{\frac{k-b-c+2a+1}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b-2}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+3}{3}}-1\right)\lambda_3(y)$

$S_y(k) - S_x(k) = 2^{\frac{k-a-b+2c+1}{3}}\left(2^{a-c} + 2^{b-c-1} + 2^b + 3.2^{a+b-1} - 1\right) > 0$

□

Proof of lemma 4.3.4. Let $a \ge b \ge c+2$ and write $x = \underline{a,b,c}$, $y = \underline{a,b-1,c+1}$. x is type II. If $b > c+2$ then y is type II; if $b = c+2$ then y is type I. Again, we need to show that $S_x(k) \le S_y(k)$. If $b > c+2$ then we must consider the following cases:

Case (Ai): $k < a-b$:

$S_x(k) = (2^{k+1}-1)\lambda_1(x)$, $S_y(k) = (2^{k+1}-1)\lambda_1(y)$, and $\lambda_1(y) - \lambda_1(x) = 2^c > 0$.

Case (Aii): $a-b+1 \le k < a+b-2c-3, k = a-b \pmod 2$:

$S_x(k) = \left(2^{\frac{k+a-b+2}{2}}-1\right)\lambda_1(x) + \left(2^{\frac{k-a+b}{2}}-1\right)\lambda_2(x)$

$S_y(k) = \left(2^{\frac{k+a-b+2}{2}}-1\right)\lambda_1(y) + \left(2^{\frac{k-a+b}{2}}-1\right)\lambda_2(y)$

$S_y(k) - S_x(k) = \left(2^{\frac{k+a-b+2}{2}}-1\right)2^c + \left(2^{\frac{k-a+b}{2}}-1\right)2^{a+c} > 0$

Case $(Aiii)$: $a-b+1 \le k < a+b-2c-3, k = a-b+1 \pmod 2$:

$S_x(k) = \left(2^{\frac{k+a-b+1}{2}}-1\right)\lambda_1(x) + \left(2^{\frac{k-a+b+1}{2}}-1\right)\lambda_2(x)$

$S_y(k) = \left(2^{\frac{k+a-b+3}{2}}-1\right)\lambda_1(y) + \left(2^{\frac{k-a+b-1}{2}}-1\right)\lambda_2(y)$

$S_y(k) - S_x(k) = \left(2^{\frac{k-a+b-1}{2}}-1\right)\left(3.2^{a-b+c+1} + 2^{a-b+1} - 1 + 2^{a+c+1} - 2^a\right) - 2^{a+c} - 2^c > 0$

Case (Aiv): $k = a+b-2c-3$:
$S_x(k) = (2^{a-c-1}-1)\,\lambda_1(x) + \left(2^{b-c-1}-1\right)\lambda_2(x)$
$S_y(k) = (2^{a-c-1}-1)\,\lambda_1(y) + \left(2^{b-c-1}-1\right)\lambda_2(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-1} - 2^{a+c} - 2^{a+b-c-2} + 2^{a-1} - 2^{b-c-2} + 2^{b-1} - 2^c + 1 > 0$

Case (Av): $k = a+b-2c-2$:
$S_x(k) = (2^{a-c}-1)\,\lambda_1(x) + \left(2^{b-c-1}-1\right)\lambda_2(x)$
$S_y(k) = (2^{a-c}-1)\,\lambda_1(y) + \left(2^{b-c-2}-1\right)\lambda_2(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-1} - 2^{a+c} - 2^{a+b-c-2} + 2^{a} - 2^{b-c-2} + 2^{b-1} - 2^c + 1 > 0$

Case (Avi): $k = a+b-2c-1$:
$S_x(k) = (2^{a-c}-1)\,\lambda_1(x) + \left(2^{b-c}-1\right)\lambda_2(x)$
$S_y(k) = (2^{a-c}-1)\,\lambda_1(y) + \left(2^{b-c-1}-1\right)\lambda_2(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-1} - 2^{a+b-c-1} - 2^{a+c} + 2^{a} - 2^{b-c-1} + 2^{b+1} + 1 - 2^c > 0$

Case $(Avii)$: $a+b-2c \le k < l$ and $k = a+b-2c \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+3}{3}}-1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b-3}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c}{3}}-1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c}{3}}\right)(2^a+1)\left(2^{b-1}-2^c\right) + 2^c > 0$

Case $(Aviii)$: $a+b-2c \le k < l$ and $k = a+b-2c+1 \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a+2}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b-1}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+2}{3}}-1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a+2}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b-4}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+5}{3}}-1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c-1}{3}}\right)(2^a+1)\left(2^{b-1}-2^{c+1}\right) + 2^c + 2^{a+c} > 0$

Case (Aix): $a+b-2c \le k < l$ and $k = a+b-2c+2 \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a+1}{3}}-1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b+1}{3}}-1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+1}{3}}-1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a+1}{3}}-1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b-2}{3}}-1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+4}{3}}-1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c+1}{3}}\right)(2^a+1)\left(2^{b-1}-2^c\right) > 0$
If $b = c+2$ and y is hence type I, we have the following cases:

Case (Bi): $k \le a-b$:
$S_x(k) = (2^{k+1}-1)\lambda_1(x)$, $S_y(k) = (2^{k+1}-1)\lambda_1(y)$, and $S_y(k) - S_x(k) = \left(2^{k+1}-1\right)2^{b-2} > 0$.
Case (Bii): $k = a-b+1$:
$S_x(k) = \left(2^{a-b+1}-1\right)\lambda_1(x) + \lambda_2(x)$
$S_y(k) = \left(2^{a-b+1}-1\right)\lambda_1(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-2} - 2^{a-1} - 2^{b-2} > 0$

Case $(Biii)$: $k = a - b + 2$:
$S_x(k) = \left(2^{a-b+2} - 1\right) \lambda_1(x) + \lambda_2(x)$
$S_y(k) = \left(2^{a-b+1} - 1\right) \lambda_1(y) + \lambda_2(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-2} + 2^{b-2} > 0$

Case (Biv): $k = a - b + 2$:
$S_x(k) = \left(2^{a-b+2} - 1\right) \lambda_1(x) + 3\lambda_2(x)$
$S_y(k) = \left(2^{a-b+2} - 1\right) \lambda_1(y) + \lambda_2(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-2} - 1 > 0$

Case (Bv): $a - b + 4 \leq k < l$ and $k = a - b + 1 \pmod 3$:
$S_x(k) = \left(2^{\frac{k+2a-2b+2}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-a+b+2}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a+b-1}{3}} - 1\right) \lambda_3(x)$
$S_y(k) = \left(2^{\frac{k+2a-2b+2}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-a+b-1}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a+b+2}{3}} - 1\right) \lambda_3(y)$
$S_y(k) - S_x(x) = -2^{\frac{k+2a+b-4}{3}} + 2^{a+b-2} + 2^{b-2} > 0$

Case (Bvi): $a - b + 4 \leq k < l$ and $k = a - b + 2 \pmod 3$:
$S_x(k) = \left(2^{\frac{k+2a-2b4}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-a+b+1}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a+b-2}{3}} - 1\right) \lambda_3(x)$
$S_y(k) = \left(2^{\frac{k+2a-2b+1}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-a+b+1}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a+b+1}{3}} - 1\right) \lambda_3(y)$
$S_y(k) - S_x(x) = -2^{\frac{k-a+b-2}{3}} + 2^{a+b-2} + 2^{b-2} > 0$

Case $(Bvii)$: $a - b + 4 \leq k < l$ and $k = a - b \pmod 3$:
$S_x(k) = \left(2^{\frac{k+2a-2b+3}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-a+b+3}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a+b-3}{3}} - 1\right) \lambda_3(x)$
$S_y(k) = \left(2^{\frac{k+2a-2b+3}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-a+b}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a+b}{3}} - 1\right) \lambda_3(y)$
$S_y(k) - S_x(x) = \left(2^{b-2} - 2^{\frac{k-a+b-3}{3}}\right) (2^a + 1) > 0$

□

Proof of lemma 4.3.5. We are given $a-2 \geq b > c$. We wish to show that $x + \underline{a, b, c} \succ \underline{a-1, b+1, c} = y$. Now x and y are both of type II, and so we need to consider the following cases:

Case (i): $k \leq a - b - 2$:

$S_x(k) = (2^{k+1} - 1)\lambda_1(x)$, $S_y(k) = (2^{k+1} - 1)\lambda_1(y)$, and $S_y(k) - S_x(k) = 2^{b+c}\left(2^{k+1} - 1\right) > 0$.

Case (ii): k = a-b-2:
$S_x(k) = \left(2^{a-b-1} - 1\right) \lambda_1(x)$
$S_y(k) = \left(2^{a-b-2} - 1\right) \lambda_1(y)$
$S_y(k) - S_x(k) = (2^{a-b-1} + 1)2^{b+c} > 0$

Case (iii): $k = a - b - 1$:

$$S_x(k) = \left(2^{a-b} - 1\right) \lambda_1(x)$$
$$S_y(k) = \left(2^{a-b-1} - 1\right) \lambda_1(y) + \lambda_2(y)$$
$$S_y(k) - S_x(k) = 2^{a+c-1} - 2^{b+c} + 2^{a-1} - 2^{a-b-1} - 2^{a-b+c-1} + 1 > 0$$

Case (iv): $a-b \leq k < a+b-2c, k = a-b \pmod 2$:
$$S_x(k) = \left(2^{\frac{k+a-b}{2}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-a+b}{2}} - 1\right) \lambda_2(x)$$
$$S_y(k) = \left(2^{\frac{k+a-b-2}{2}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-a+b+2}{2}} - 1\right) \lambda_2(y)$$
$$S_y(k) - S_x(k) = 2^{\frac{k+a-b+2c-2}{2}} + 2^{a+c-1} + 2^{a-1} - 2^{b+c} > 0$$

Case (v): $a-c \leq k < c+a-2b, k = a-b+1 \pmod 2$:
$$S_x(k) = \left(2^{\frac{k+a-b+1}{2}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-a+b+1}{2}} - 1\right) \lambda_2(x)$$
$$S_y(k) = \left(2^{\frac{k+a-b-1}{2}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-a+b+3}{2}} - 1\right) \lambda_2(y)$$
$$S_y(k) - S_x(k) = 2^{\frac{k+a-b+2c-1}{2}} + 2^{a+c-1} + 2^{a-1} - 2^{b+c} > 0$$

Case (vi): $c+a-2b \leq k < l, k = c+a-2b \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-c-a+2b}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a-b+2c+3}{3}} - 1\right) \lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-3}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-c-a+2b+3}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a-b+2c+3}{3}} - 1\right) \lambda_3(y)$$
$$S_y(k) - S_x(k) = 2^{a+c-1} - 2^{b+c} - 2^{a-1} + 2^b + 2^{\frac{k-a-b-c}{3}} \left(2^{a+c-1} - 2^{b+c+1}\right) > 0$$

Case (vii): $c+a-2b \leq k < l, k = c+a-2b+1 \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a+2}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-c-a+2b-1}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a-b+2c+2}{3}} - 1\right) \lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-1}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-c-a+2b+2}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a-b+2c+2}{3}} - 1\right) \lambda_3(y)$$
$$S_y(k) - S_x(k) = 2^{a+c-1} - 2^{b+c} - 2^{a-1} + 2^b + 2^{\frac{k-a-b-c-1}{3}} \left(2^{a+c} - 2^{b+c+1}\right) > 0$$

Case $(viii)$: $c+a-2b \leq k < l, k = c+a-2b+2 \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a+1}{3}} - 1\right) \lambda_1(x) + \left(2^{\frac{k-c-a+2b+1}{3}} - 1\right) \lambda_2(x) + \left(2^{\frac{k-a-b+2c+1}{3}} - 1\right) \lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-2}{3}} - 1\right) \lambda_1(y) + \left(2^{\frac{k-c-a+2b+4}{3}} - 1\right) \lambda_2(y) + \left(2^{\frac{k-a-b+2c+1}{3}} - 1\right) \lambda_3(y)$$
$$S_y(k) - S_x(k) = 2^{a+c-1} - 2^{b+c} - 2^{a-1} + 2^b + 2^{\frac{k-a-b-c-1}{3}} \left(2^{a+c-1} - 2^{b+c}\right) \geq 0 \qquad \square$$

Proof of lemma 4.3.6. We are given $a-2 \geq c \geq b$. We wish to show that $x = \underline{a, b, c} \succ \underline{a-1, b, c+1} = y$. Now x and y are both type I, so we need to consider the following cases:

Case (i): $k \leq a-c-3$:

$S_x(k) = (2^{k+1} - 1)\lambda_1(x)$, $S_y(k) = (2^{k+1} - 1)\lambda_1(y)$, and $S_y(k) - S_x(k) = 2^c \left(2^{k+1} - 1\right) \left(2^b + 1\right) > 0$.

Case (ii): $k = a-c-2$:

$S_x(k) = (2^{a-c-1} - 1)\,\lambda_1(x)$
$S_y(k) = (2^{a-c-2} - 1)\,\lambda_1(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-1} - 2^{b+c} - 2^{a-c-2} - 2^c + 2^b + 1 > 0$

Case (iii): $k = a - c - 1$:
$S_x(k) = (2^{a-c} - 1)\,\lambda_1(x)$
$S_y(k) = (2^{a-c-1} - 1)\,\lambda_1(y) + \lambda_3(y)$
$S_y(k) - S_x(k) = 2^{a+b-1} - 2^{b+c} - 2^{a-c-1} - 2^c + 2^b + 1 > 0$

Case (iv): $a - c \le k < c + a - 2b, k = a - c \pmod 2$:
$S_x(k) = \left(2^{\frac{k+a-c}{2}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-a+c+2}{2}} - 1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k+a-c-2}{2}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-a+c+4}{2}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^c - 2^{\frac{k-a+c-2}{2}}\right)\left(2^{a-c-2} - 2^b - 1\right) - 2^{a-2} + 2^{a+b-1} > 0$

Case (v): $a - c \le k < c + a - 2b, k = a - c + 1 \pmod 2$:
$S_x(k) = \left(2^{\frac{k+a-c+1}{2}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-a+c+1}{2}} - 1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k+a-c-1}{2}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-a+c+3}{2}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(2^c - 2^{\frac{k-a+c+1}{2}}\right)\left(2^{a-c-1} - 2^b - 1\right) - 2^{a-1} + 2^{a+b-1} > 0$

Case (vi): $a + b - 2c \le k < l, k = a + b - 2c \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a}{3}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b}{3}} - 1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+3}{3}} - 1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a-3}{3}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b}{3}} - 1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+6}{3}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c}{3}}\right)\left(2^b + 1\right)\left(2^{b-1} - 2^c\right) + 2^{\frac{k-a-b+2c}{3}}\left(2^b + 1\right) > 0$

Case (vii): $a + b - 2c \le k < l, k = a + b - 2c + 1 \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a-1}{3}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b+2}{3}} - 1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+2}{3}} - 1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a-4}{3}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b+2}{3}} - 1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+5}{3}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c+2}{3}}\right)\left(2^b + 1\right)\left(2^{a-1} - 2^c\right) + 2^{\frac{k+2a-b-c-4}{3}} > 0$

Case $(viii)$: $a + b - 2c \le k < l, k = a + b - 2c + 2 \pmod 3$:
$S_x(k) = \left(2^{\frac{k-b-c+2a+1}{3}} - 1\right)\lambda_1(x) + \left(2^{\frac{k-c-a+2b+1}{3}} - 1\right)\lambda_2(x) + \left(2^{\frac{k-a-b+2c+1}{3}} - 1\right)\lambda_3(x)$
$S_y(k) = \left(2^{\frac{k-b-c+2a-2}{3}} - 1\right)\lambda_1(y) + \left(2^{\frac{k-c-a+2b+1}{3}} - 1\right)\lambda_2(y) + \left(2^{\frac{k-a-b+2c+4}{3}} - 1\right)\lambda_3(y)$
$S_y(k) - S_x(k) = \left(1 - 2^{\frac{k-a-b-c+1}{3}}\right)\left(2^b + 1\right)\left(2^{a-1} - 2^c\right) \ge 0$ □

Proof of lemma 4.3.7. We are given $a-2 \ge c \ge b$. We wish to show that $x = \underline{a, b, c} \succ \underline{a-1, c+1, b} = y$. Now x is type I, y is type II, so we need to consider the following cases:

Case (i): $k \leq a-c-2$:
$S_x(k) = (2^{k+1}-1)\lambda_1(x)$, $S_y(k) = (2^{k+1}-1)\lambda_1(y)$, and $S_y(k)-S_x(k) = \left(2^{k+1}-1\right)\left(2^{b+c}+2^b-2^c\right) > 0$.

Case (ii): $k = a-c-2$:
$$S_x(k) = \left(2^{a-c}-1\right)\lambda_1(x)$$
$$S_y(k) = \left(2^{a-c-1}-1\right)\lambda_1(y)+\lambda_2(y)$$
$$S_y(k)-S_x(k) = 2^{a-c-1}\left(2^b-1\right)\left(2^c+1\right) > 0$$

Case (iii): $a-c \leq k < a-2b+c, k = a-c \pmod 2$:
$$S_x(k) = \left(2^{\frac{k+a-c}{2}}-1\right)\lambda_1(x)+\left(2^{\frac{k-a+c+2}{2}}-1\right)\lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k+a-c}{2}}-1\right)\lambda_1(y)+\left(2^{\frac{k-a+c+2}{2}}-1\right)\lambda_2(y)$$
$$S_y(k)-S_x(k) = \left(2^{a-1}-2^c\right)\left(2^{\frac{k-a+2b-c+2}{2}}+2^b-1\right) > 0$$

Case (iv): $a-c \leq k < c+a-2b, k = a-c+1 \pmod 2$:
$$S_x(k) = \left(2^{\frac{k+a-c+1}{2}}-1\right)\lambda_1(x)+\left(2^{\frac{k-a+c+1}{2}}-1\right)\lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k+a-c-1}{2}}-1\right)\lambda_1(y)+\left(2^{\frac{k-a+c+3}{2}}-1\right)\lambda_2(y)$$
$$S_y(k)-S_x(k) = \left(1+2^{\frac{k-a-c+1}{2}}\right)\left(2^b-1\right)\left(2^{a-1}-2^c\right)+2^{\frac{k+a+c+3}{2}} > 0$$

Case v: $a+b-2c \leq k < l, k = a+b-2c \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a}{3}}-1\right)\lambda_1(x)+\left(2^{\frac{k-c-a+2b}{3}}-1\right)\lambda_2(x)+\left(2^{\frac{k-a-b+2c+3}{3}}-1\right)\lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-3}{3}}-1\right)\lambda_1(y)+\left(2^{\frac{k-c-a+2b+3}{3}}-1\right)\lambda_2(y)+\left(2^{\frac{k-a-b+2c+3}{3}}-1\right)\lambda_3(y)$$
$$S_y(k)-S_x(k) = \left(2^b-1\right)\left(2^{a-1}\left(1-2^{\frac{k-a-b-c}{3}}\right)-2^c\right)+2^{b+c+1}-1 > 0$$

Case (vi): $a+b-2c \leq k < l, k = a+b-2c+1 \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a-1}{3}}-1\right)\lambda_1(x)+\left(2^{\frac{k-c-a+2b+2}{3}}-1\right)\lambda_2(x)+\left(2^{\frac{k-a-b+2c+2}{3}}-1\right)\lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-1}{3}}-1\right)\lambda_1(y)+\left(2^{\frac{k-c-a+2b+2}{3}}-1\right)\lambda_2(y)+\left(2^{\frac{k-a-b+2c+2}{3}}-1\right)\lambda_3(y)$$
$$S_y(k)-S_x(k) = \left[2^b\left(1-2^{\frac{k-a-b-c+2}{3}}\right)+1\right]\left(2^{a-1}-2^c\right) > 0$$

Case (vii): $a+b-2c \leq k < l, k = a+b-2c+2 \pmod 3$:
$$S_x(k) = \left(2^{\frac{k-b-c+2a+1}{3}}-1\right)\lambda_1(x)+\left(2^{\frac{k-c-a+2b+1}{3}}-1\right)\lambda_2(x)+\left(2^{\frac{k-a-b+2c+1}{3}}-1\right)\lambda_3(x)$$
$$S_y(k) = \left(2^{\frac{k-b-c+2a-2}{3}}-1\right)\lambda_1(y)+\left(2^{\frac{k-c-a+2b+4}{3}}-1\right)\lambda_2(y)+\left(2^{\frac{k-a-b+2c+1}{3}}-1\right)\lambda_3(y)$$
$$S_y(k)-S_x(k) = \left(1-2^{\frac{k-a-b-c+1}{3}}\right)\left(2^b+1\right)\left(2^{a-1}-2^c\right) \geq 0$$

□

Proof of lemma 4.3.8. Let $a_1+b_1+c_1 = a_2+b_2+c_2 = l$, with $a_1 \geq b_1 \geq c_1, a_2 \geq b_2 \geq c_2, a_1 > a_2, c_1 < c_2$. We wish to show that if x is either $\underline{a_1, b_1, c_1}$ or $\underline{a_1, c_1, b_1}$ and

y is either $\underline{a_2, b_2, c_2}$ or $\underline{a_2, c_2, b_2}$ then $x \not\succ y, y \not\succ x$. To prove this, it is sufficient to note that $T^{\alpha_0(x)}(x) = x < 2^{-a_1-1} < y = T^{\alpha_0(y)}(y)$ since $a_1 > a_2$, and $T^{\alpha_{l-1}(x)}(x) = x < 2^{-1} + 2^{-c_2+1} < y = T^{\alpha_{l-1}(y)}(y)$. $\square$

Proof of lemma 4.3.9. Let $a + b_1 + c_1 = a + b_2 + c_2$, with $a \geq c_i \geq b_i$, and require that if $a = c_i$ then $b_i = a = c_i$. Let $x = \underline{a, b_1, c_1}$ and $y = \underline{a, b_2, c_2}$. We wish to show that $x \not\succ y, y \not\succ x$ unless $x = y$. To do this, assume without loss of generality that $b_1 > b_2$ and $c_1 < c_2$. Then $T^{\alpha_0(x)}(x) = x < y = T^{\alpha_0(y)}(y)$ since $b_1 > b_2$ and $T^{\alpha_{l-1}(x)}(x) = T^{a-1}(x) < T^{a-1}(y) = T^{\alpha_{l-1}(y)}(y)$ since $b_1 > b_2$. $\square$

$\square$

Remark 4.3.12. While brute force is sufficient to completely solve the majorisation ordering at weights 2 and 3, for higher weights it becomes much harder. In part this is due to the the fact that at weights 2 and 3 we have the useful property that $\underline{a_1, b_1, c_1} \succ \underline{a_2, b_2, c_2}$ if and only if $\underline{a_1 + 1, b_1, c_1} \succ \underline{a_2 + 1, b_2, c_2}$. However, at higher weights this is not the case.

Chapter 5

Weakly Expanding Orientation-Reversing Maps

5.1 Introduction

Majorization is a way of making precise the notion that one measure is more *spread out* than another (see e.g. [4, 7, 9, 12, 24, 25, 26, 30]). If μ and ν are Borel probability measures on the unit interval $[0,1]$, we say ν *majorizes* μ, and write $\mu \prec \nu$, if $\mu(f) \leq \nu(f)$ for all convex functions $f : [0,1] \to \mathbb{R}$. Equivalently, $\mu \prec \nu$ if and only if

$$\int_0^t \mu[0,x]\,dx \leq \int_0^t \nu[0,x]\,dx \quad \text{for all } t \in [0,1] \tag{5.1}$$

and

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

If $T : [0,1] \to [0,1]$ is Borel, its set $\mathcal{M}_T$ of invariant Borel probability measures becomes a partially ordered set when equipped with $\prec$. For the doubling map $T(x) = 2x \pmod 1$, the poset $(\mathcal{M}_T, \prec)$ was investigated in [19] (see also [18, 20]), where its minimal and maximal elements were identified:

Theorem 5.1.1. *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*[1] *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

[1]Sturmian measures are symbolic versions of rotations (see e.g. [3, 5, 15, 18, 19] 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.

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 5.1.2. *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 5.1.1, the statement of Theorem 5.1.2 is robust under nonlinear perturbation. It is a particular case of the following result:

Theorem 5.1.3. *If $T : [0,1] \to [0,1]$ is the lift of a continuous orientation-reversing expanding circle map, with fixed points $x_1 < \ldots < x_k$, 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 $\lambda \in [0,1]$, $1 \le i \le k-1$.*

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

Theorem 5.1.4. *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-2 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 5.1.4 can be generalised as follows:

Theorem 5.1.5. *Suppose $T : [0,1] \to [0,1]$ is the lift of a continuous orientation-reversing expanding circle map, with $T(0) = 1$ and $T(1) = 0$, and with fixed points $x_1 < \ldots < x_k$. 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. [5, 10, 11, 17]), i.e. the study of the smallest and largest possible ergodic averages of a given function $f : [0,1] \to \mathbb{R}$, and of the invariant measures μ (so-called *minimizing* and *maximizing* measures, cf. Defn. 5.2.3) for which $\mu(f)$ attains these extrema. When $f : [0,1] \to \mathbb{R}$ is convex[2], we solve the ergodic optimization problem as follows:

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

Theorem 5.1.6. *Suppose $T : [0,1] \to [0,1]$ is the lift of a continuous orientation-reversing expanding circle map, with $T(0) = 1$ and $T(1) = 0$, and with fixed points $x_1 < \ldots < x_k$. If $f : [0,1] \to \mathbb{R}$ is convex then its minimum ergodic average is*

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

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

This chapter is organised as follows. In Section 5.2 we introduce the class of *orientation-reversing weakly expanding maps*, a suitable generalisation of the maps T considered above. For orientation-reversing weakly expanding maps T we identify (see Theorem 5.3.1 in Section 5.3) the minimal elements of $(\mathcal{M}_T, \prec)$, a result which implies Theorem 5.1.3. Under additional hypotheses we then identify (see Theorem 5.4.1 in Section 5.4) the maximal elements of $(\mathcal{M}_T, \prec)$, a result which implies Theorem 5.1.5. Theorem 5.1.6 is a consequence of Theorems 5.3.4 and 5.4.4, which also give more precise information in the case where $f : [0,1] \to \mathbb{R}$ is *strictly* convex. In Section 5.5, some of the fine structure of $(\mathcal{M}_T, \prec)$ is explicitly computed in the case where T is the reverse doubling map. In Section 5.6 we consider the extension of our results to interval maps with infinitely many branches, with Gauss's continued fraction map serving as an illustratative example.

5.2 Preliminaries

Notation 5.2.1. For a Borel probability measure μ on $[0,1]$, let $b(\mu) := \int x\, d\mu(x)$ denote its *barycentre*. If $T : [0,1] \to [0,1]$ 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* $\mathcal{M}_\varrho := \{\mu \in \mathcal{M}_T : b(\mu) = \varrho\}$.

Definition 5.2.2. Suppose $0 = a_0 < a_1 < \ldots < a_k = 1$, where $k \ge 2$. Let $J_1, \ldots, J_k$ be disjoint sub-intervals of $[0,1]$, with $\cup_{i=1}^k J_i = [0,1]$, such that the left (respectively right) endpoint of each J_i is a_{i-1} (respectively a_i). Suppose that, for each $1 \le i \le k$, 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 $1 \le i \le k$, and that T is *weakly expanding* if $|T(x) - T(y)| \ge |x - y|$ for all $x, y \in J_i$, $1 \le i \le k$.

Definition 5.2.3. Suppose $f : [0,1] \to \mathbb{R}$ is bounded and Borel measurable, and that $T : [0,1] \to [0,1]$ is Borel. For $\varrho \in [0,1]$, a measure $\mu \in \mathcal{M}_\varrho$ is called *f-minimizing in $\mathcal{M}_\varrho$* (respectively *f-maximizing in $\mathcal{M}_\varrho$*) 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*

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

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

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

Lemma 5.2.4. *If $T : [0,1] \to [0,1]$ is an orientation-reversing weakly expanding map, with fixed points $x_1 < \ldots < x_k$, 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 $x_1 \le b(\mu) \le x_k$ for every $\mu \in \mathcal{M}_T$.

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

Defining $\tau(x) = \max\{x_1, x\}$, a similar argument shows that $x - \tau(x) + \tau(Tx) \ge x_1$ for all $x \in [0,1]$, hence that $b(\mu) = \int x \, d\mu(x) = \int (x - \tau(x) + \tau(Tx)) \, d\mu(x) \ge x_1$ 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 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 Section 5.3 and Section 5.4 this question will be answered affirmatively[3], where the following measures m_ϱ and ν_ϱ will be identified as, respectively, the smallest and largest elements of $(\mathcal{M}_\varrho, \prec)$.

Notation 5.2.5. (The measures m_ϱ and ν_ϱ)

Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \ldots < x_k$. 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 $1 \le i \le k-1$ are such that $\varrho = \lambda x_i + (1-\lambda)x_{i+1}$.

If furthermore $T(0) = 1$ and $T(1) = 0$, with $\mu_{01} := (\delta_0 + \delta_1)/2$, and $x_1 < 1/2 < x_k$, then define ν_ϱ to be the unique member of $\mathcal{M}_\varrho$ which is a convex combination of μ_{01} with either δ_{x_1} or δ_{x_k}. Explicitly,

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

5.3 Minimal elements of $(\mathcal{M}_T, \prec)$, and minimizing measures for convex functions f

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

Theorem 5.3.1. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \ldots < x_k$. For every $\varrho \in [x_1, x_k]$, the ordered set $(\mathcal{M}_\varrho, \prec)$ has a smallest element, namely m_ϱ.*

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

$$m_\varrho(f) \le \mu(f) \quad \text{for all } \mu \in \mathcal{M}_\varrho\,, \tag{5.2}$$

where

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

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

We now claim that all measures in $\mathcal{N} = \{\varepsilon\delta_{x_i} + (1-\varepsilon)\delta_{x_{i+1}} : \varepsilon \in [0,1]\}$ are f-minimizing. This implies $m_\varrho(f) \le \mu(f)$ for all $\mu \in \mathcal{M}_T$, which in particular implies (5.2).

[3] We emphasise 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 Chapter 2 of this thesis.

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

$$(f + \varphi - \varphi \circ T)(x) \geq \alpha \,. \tag{5.3}$$

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 (5.3) holds.

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

$$(f + \varphi - \varphi \circ T)(x) = f(x) - \alpha + f(Tx) \,. \tag{5.4}$$

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

$$x_i - T(x) \geq x - x_i \,. \tag{5.5}$$

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 (5.5) gives

$$f(x_i) - f(Tx) = \frac{f(x_i) - f(Tx)}{x_i - T(x)} \left(x_i - T(x)\right) \leq \frac{f(x_i) - f(Tx)}{x_i - T(x)} (x - x_i) \,. \tag{5.6}$$

But f is convex, so (see e.g. [31, p. 113])

$$\frac{f(x_i) - f(Tx)}{x_i - T(x)} \leq \frac{f(x) - f(x_i)}{x - x_i} \,. \tag{5.7}$$

Combining (5.6) with (5.7) gives $f(x_i) - f(Tx) \leq f(x) - f(x_i)$, and substituting into (5.4) gives (5.3).

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

$$T(x) - x_{i+1} \geq x_{i+1} - x \,. \tag{5.8}$$

Since $f(x_i) = f(x_{i+1})$, the convexity of f implies it is non-decreasing on $[x_{i+1}, 1]$, so $f(Tx) - f(x_{i+1}) \geq 0$. Combining this with (5.8), 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 (5.3) follows from (5.4).

Having established (5.3) 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. □

Theorem 5.3.1 asserts that if $f : [0,1] \to \mathbb{R}$ is convex then the measure m_ϱ is minimizing in $\mathcal{M}_\varrho$ (cf. Defn 5.2.3). If f is *strictly* convex then this conclusion can be strengthened as follows:

Corollary 5.3.2. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \ldots < x_k$. If $f : [0,1] \to \mathbb{R}$ is strictly convex, and $\varrho \in [x_1, x_k]$, then m_ϱ is the unique f-minimizing measure in $\mathcal{M}_\varrho$.*

Proof. An alternative definition of majorization (see [4, 7]) is that $\mu \prec \nu$ if and only if ν is a *dilation* of μ, i.e. there exists a family of probability measures $(D_x)_{x\in[0,1]}$, with each $b(D_x) = x$, such that if $f : [0,1] \to \mathbb{R}$ 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 $m_\varrho \prec \mu$ by Theorem 5.3.1, 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, i.e. $D_x(f) > f(x)$, 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* $\mathrm{var}(\mu) = \int (x - b(\mu))^2\, d\mu(x)$ around the mean ϱ is minimized precisely when $\mu = m_\varrho$:

Corollary 5.3.3. *For every $\varrho \in [x_1, x_k]$, the measure m_ϱ is the unique measure with smallest variance in $\mathcal{M}_\varrho$.*

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

Theorem 5.3.4. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with fixed points $x_1 < \ldots < x_k$. If $f : [0,1] \to \mathbb{R}$ is convex then its minimum ergodic average is*

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

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

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

Proof. The map $F : \varrho \mapsto \inf_{\mu\in\mathcal{M}_\varrho} \mu(f)$ is clearly convex on $[x_1, x_k]$. Since $m_\varrho(f) = \inf_{\mu\in\mathcal{M}_\varrho} \mu(f)$ by Theorem 5.3.1, the map F is affine on each interval $[x_i, x_{i+1}]$, $1 \le i \le k-1$. Consequently, there exists $1 \le i \le k$ such that $f(x_i) = \min_{\varrho\in[x_1,x_k]} F(\varrho) = \alpha(f)$, and therefore δ_{x_i} is f-minimizing.

Now suppose that f is strictly convex. If there is a unique $1 \leq i \leq k$ such that $f(x_i) = \min_{\varrho \in [x_1, x_k]} F(\varrho) = \alpha(f)$, then the f-minimizing measures are precisely those which are minimizing in $\mathcal{M}_{x_i}$; Corollary 5.3.2 then implies that δ_{x_i} is the unique invariant measure which is minimizing in $\mathcal{M}_{x_i}$, hence it is the unique f-minimizing measure. If not then the strict convexity of f implies that $f(x_i) = f(x_{i+1}) = \min_{\varrho \in [x_1, x_k]} F(\varrho) = \alpha(f)$ for some $1 \leq i \leq k-1$, with $f(x_j) > \alpha(f)$ for $j \in \{1, \ldots, k\} \setminus \{i, i+1\}$. So F attains its minimum value $\alpha(f)$ precisely on the interval $[x_i, x_{i+1}]$, hence an invariant measure is f-minimizing if and only if it is minimizing in $\mathcal{M}_\varrho$ for some $\varrho \in [x_i, x_{i+1}]$. Corollary 5.3.2 then implies that the set of f-minimizing measures is precisely $\{\lambda \delta_{x_i} + (1-\lambda)\delta_{x_{i+1}} : \lambda \in [0,1]\}$. $\square$

When T is the reverse doubling map, the result is particularly explicit:

Corollary 5.3.5. *Let $T(x) = -2x \pmod 1$ for $x \in (0,1]$, and $T(0) = 1$. If $f : [0,1] \to \mathbb{R}$ is strictly convex then its minimum ergodic average is*

$$\alpha(f) = \min\left(f(1/3), f(2/3)\right).$$

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

5.4 Maximal elements of $(\mathcal{M}_T, \prec)$, and maximizing measures for convex f

We shall prove the following stronger version of Theorem 5.1.5.

Theorem 5.4.1. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with $T(0) = 1$ and $T(1) = 0$. Let $x_1 < \ldots < x_k$ denote the fixed points of T, and assume that $x_1 < 1/2 < x_k$. For every $\varrho \in [x_1, x_k]$, the ordered set $(\mathcal{M}_\varrho, \prec)$ has a largest element, namely ν_ϱ.*

Proof. Suppose that $\varrho \in [x_1, 1/2]$; the proof for $\varrho \in [1/2, x_k]$ is almost identical, and will be omitted. For all convex functions $f : [0,1] \to \mathbb{R}$, we wish to show that

$$\mu(f) \leq \nu_\varrho(f) \quad \text{for all } \mu \in \mathcal{M}_\varrho. \tag{5.9}$$

We may assume that $f(0) = 0 = f(1)$ (since $\tilde{f}(x) := f(x) + (f(0) - f(1))x - f(0)$ satisfies $\tilde{f}(0) = 0 = \tilde{f}(1)$, and (5.9) holds if and only if $\mu(\tilde{f}) \leq \nu_\varrho(\tilde{f})$ for all $\mu \in \mathcal{M}_\varrho$).

So, since $\nu_\varrho = \frac{1/2-\varrho}{1/2-x_1}\delta_{x_1} + \frac{\varrho - x_1}{1/2-x_1}\mu_{01}$, we wish to show that

$$\mu(f) \le \frac{1/2-\varrho}{1/2-x_1} f(x_1) \quad \text{for all } \mu \in \mathcal{M}_\varrho \,. \tag{5.10}$$

Now (5.10) is implied by the existence of a continuous function $\varphi : [0,1] \to \mathbb{R}$ satisfying

$$(f + \varphi - \varphi \circ T)(x) \le \frac{1/2 - x}{1/2 - x_1} f(x_1) \quad \text{for all } x \in [0,1] \,, \tag{5.11}$$

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

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

then (5.11) becomes

$$(\hat{f} + \varphi - \varphi \circ T)(x) \ \le \ \frac{1/2}{1/2 - x_1} f(x_1) \ = \ \hat{f}(x_1) \quad \text{for all } x \in [0,1] \,. \tag{5.12}$$

Defining φ by

$$\varphi(x) = \begin{cases} -\hat{f}(x_1) & \text{for } x \in [0, x_1] \,, \\ -\hat{f}(x) & \text{for } x \in [x_1, 1] \,, \end{cases} \tag{5.13}$$

we shall verify that (5.12) holds[4]. If $x \in [x_1, 1]$ then $(\hat{f} + \varphi - \varphi \circ T)(x) = -\varphi(Tx) = \hat{f}(x_1)$, 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 (5.12) holds if and only if

$$\hat{f}(x) + \hat{f}(Tx) \le 2\hat{f}(x_1) \quad \text{for all } x \in [0, x_1] \,. \tag{5.14}$$

To prove (5.14), first note that convexity of $\hat{f}$ implies

$$\hat{f}(x) \le \frac{x}{x_1}\hat{f}(x_1) + \left(1 - \frac{x}{x_1}\right)\hat{f}(0) = \frac{x}{x_1}\hat{f}(x_1) \quad \text{for all } x \in [0, x_1] \,, \tag{5.15}$$

and

$$\hat{f}(y) \le \frac{1-y}{1-x_1}\hat{f}(x_1) + \frac{y - x_1}{1 - x_1}\hat{f}(1) \quad \text{for all } y \in [x_1, 1] \,. \tag{5.16}$$

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

$$T(x) \ge 2x_1 - x \,. \tag{5.17}$$

[4]Note in particular that (5.12) implies that δ_{x_1} is (globally) $\hat{f}$-maximizing.

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

$$\hat{f}(1) = 2\hat{f}(x_1) = \frac{f(x_1)}{1/2 - x_1} < 0 = \hat{f}(0)\,, \tag{5.18}$$

so (5.17) implies that

$$\hat{f}(Tx) \leq \hat{f}(2x_1 - x)\,. \tag{5.19}$$

Setting $y = 2x_1 - x$ in (5.16), and combining with (5.19), gives

$$\hat{f}(Tx) \leq \frac{1 + x - 2x_1}{1 - x_1}\hat{f}(x_1) + \frac{x_1 - x}{1 - x_1}\hat{f}(1)\,. \tag{5.20}$$

Using (5.15) with (5.20), and the fact that $\hat{f}(1) = 2\hat{f}(x_1)$, we obtain

$$\hat{f}(x) + \hat{f}(Tx) \leq \hat{f}(x_1)\left(\frac{x}{x_1} + \frac{1-x}{1-x_1}\right) \quad \text{for all } x \in [0, x_1]\,. \tag{5.21}$$

Now $x_1 < 1/2$, so the function $x \mapsto \frac{x}{x_1} + \frac{1-x}{1-x_1}$ is increasing on $[0, x_1]$, attaining its maximum value 2 when $x = x_1$. Therefore (5.21) implies the desired inequality (5.14), and the proof is complete. $\square$

The proofs of the following two corollaries are similar to those of Corollaries 5.3.2 and 5.3.3, so will be omitted.

Corollary 5.4.2. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with $T(0) = 1$ and $T(1) = 0$. Let $x_1 < \ldots < x_k$ denote the fixed points of T, and assume that $x_1 < 1/2 < x_k$. If $f : [0,1] \to \mathbb{R}$ is strictly convex, and $\varrho \in [x_1, x_k]$, then ν_ϱ is the unique maximizing measure in $\mathcal{M}_\varrho$.*

Corollary 5.4.3. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with $T(0) = 1$ and $T(1) = 0$. Let $x_1 < \ldots < x_k$ denote the fixed points of T, and assume that $x_1 < 1/2 < x_k$. For every $\varrho \in [x_1, x_k]$, the measure ν_ϱ is the unique measure with largest variance in $\mathcal{M}_\varrho$.*

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

Theorem 5.4.4. *Let $T : [0,1] \to [0,1]$ be an orientation-reversing weakly expanding map, with $T(0) = 1$ and $T(1) = 0$. Let $x_1 < \ldots < x_k$ denote the fixed points of T, and assume that $x_1 < 1/2 < x_k$. If $f : [0,1] \to \mathbb{R}$ is convex then*

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

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

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

(i) δ_{x_1} is the unique f-maximizing measure,

(ii) μ_{01} is the unique f-maximizing measure,

(iii) δ_{x_k} is the unique f-maximizing measure,

(iv) The set of f-maximizing measures is $\{\lambda\delta_{x_1} + (1-\lambda)\mu_{01} : \lambda \in [0,1]\}$,

(v) The set of f-maximizing measures is $\{\lambda\mu_{01} + (1-\lambda)\delta_{x_k} : \lambda \in [0,1]\}$.

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 5.4.1 implies that G is affine on both $[x_1, 1/2]$ and $[1/2, x_k]$. Consequently, G attains its maximum value at either 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 $f : [0,1] \to \mathbb{R}$ is strictly convex. If G has a unique maximum then this must be attained at either x_1, $1/2$, or x_k, in which case Corollary 5.4.2 implies that the unique f-maximizing measure is, respectively, δ_{x_1}, μ_{01}, or δ_{x_k}.

We next show that $G : [x_1, x_k] \to \mathbb{R}$ 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 5.4.2 implies that the set of f-maximizing measures is, respectively, either $\{\lambda\delta_{x_1} + (1-\lambda)\mu_{01} : \lambda \in [0,1]\}$ or $\{\lambda\mu_{01} + (1-\lambda)\delta_{x_k} : \lambda \in [0,1]\}$. □

Remark 5.4.5. Theorem 5.4.1 implies Theorem 5.1.5, and Theorem 5.4.4 implies the corresponding part of Theorem 5.1.6, since the condition $x_1 < 1/2 < x_k$ in Theorems 5.4.1 and 5.4.4 is automatically satisfied if T is expanding.

5.5 Computations

The majorization criterion (5.1) can be used to compute some of the structure of $(\mathcal{M}_T, \prec)$. 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 (5.1) can be re-formulated in terms of a well known criterion of Hardy, Littlewood & Pólya (see [12, 13, 18, 19]); 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 $\mu_1 \le \ldots \le \mu_Q$ and $\nu_1 \le \ldots \le \nu_Q$, and $b(\mu) = b(\nu)$, then $\mu \prec \nu$ if and only if $\sum_{i=1}^{n} \mu_i \ge \sum_{i=1}^{n} \nu_i$ for all $1 \le n \le Q-1$.

Now let $T : [0,1] \to [0,1]$ 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

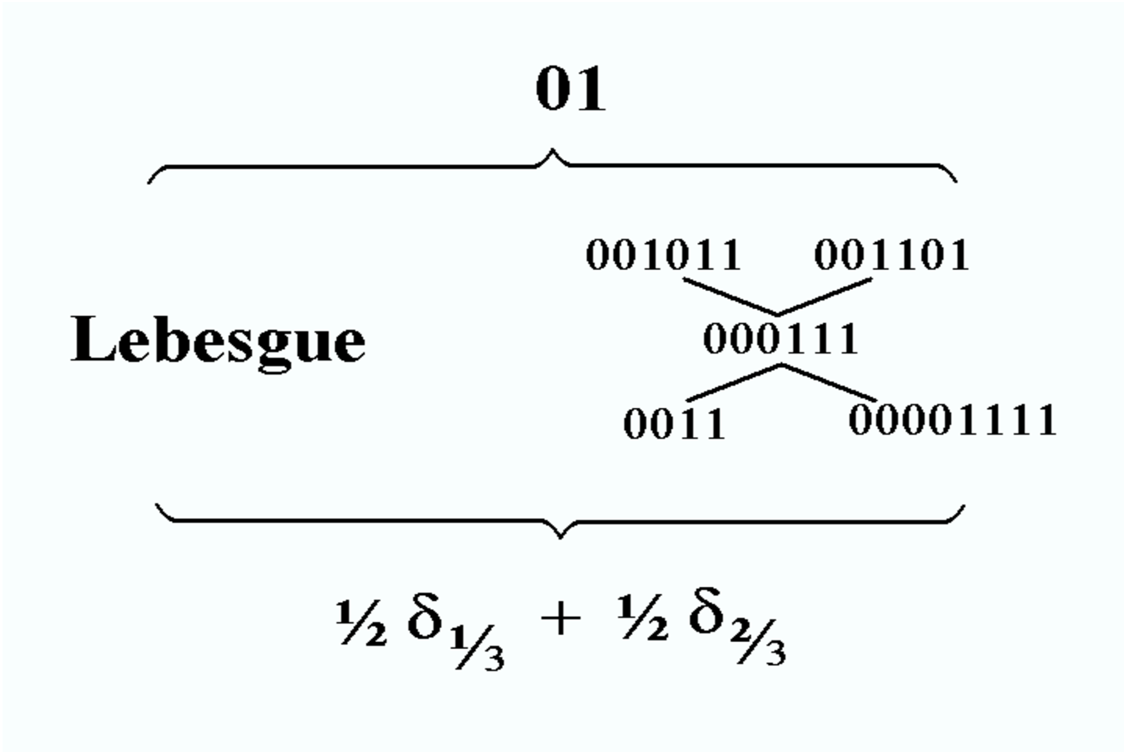

Figure 5.1: Hasse diagram for a portion of $(\mathcal{M}_{1/2}, \prec) \subset (\mathcal{M}_T, \prec)$*, where* T *is the reverse doubling map. Symbolic codes denote the corresponding periodic orbit measures. The least element in* $(\mathcal{M}_{1/2}, \prec)$ *is the non-ergodic measure* $(\delta_{1/3} + \delta_{2/3})/2$*, while the largest element is the invariant measure supported by the period-2 orbit* $\{0, 1\}$*.*

periodic orbit lies in $\mathcal{M}_\varrho$ 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 $\mathcal{M}_\varrho$. For barycentre $\varrho = 1/2$, the majorization relations between some of the periodic orbit measures in $\mathcal{M}_{1/2}$ are depicted in Figure 5.1. 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-4 orbit $\{1/5, 3/5, 4/5, 2/5\}$.

5.6 Infinitely many branches: the Gauss map

The results of this chapter can be extended, with some modification, to maps $T : [0, 1] \to [0, 1]$ which are orientation-reversing and weakly expanding, but with infinitely many branches. More precisely, the techniques of Section 5.3, and to a lesser extent those of Section 5.4, can be used to study the majorization structure of $\mathcal{M}_T$ for maps T satisfying the analogue of Definition 5.2.2 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 necessitate 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 shall 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 Section 5.3 and Section 5.4.

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 Section 5.4, 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 5.2.5, 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 5.2.4 and Theorem 5.3.1 give:

Theorem 5.6.1. *For $T : [0,1] \to [0,1]$ the Gauss map, $b(\mathcal{M}_T) = (0, (\sqrt{5}-1)/2]$. For every $\varrho \in (0,(\sqrt{5}-1)/2]$, the ordered set $(\mathcal{M}_\varrho, \prec)$ has a smallest element, namely m_ϱ.*

Using the same arguments as in Section 5.3, it can also be shown that m_ϱ is the unique f-minimizing measure in $\mathcal{M}_\varrho$ for strictly convex $f : [0,1] \to \mathbb{R}$, and in particular that m_ϱ is the unique measure in $\mathcal{M}_\varrho$ with smallest variance.

The following result on global minimization for convex functions is an analogue of Theorem 5.3.4, the significant difference being that f-minimizing measures need not exist (e.g. this occurs if the convex f is strictly increasing):

Theorem 5.6.2. *Let $T : [0,1] \to [0,1]$ be the Gauss map. If $f : [0,1] \to \mathbb{R}$ is convex then its minimum ergodic average is*

$$\alpha(f) = \inf_{i\geq 1} f(z_i)\,. \tag{5.22}$$

If the infimum (5.22) 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 (5.22) is attained by a unique* $i \geq 1$*, then the corresponding Dirac measure* δ_{z_i} *is the unique* f*-minimizing measure. If (5.22) 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* $\{\lambda\delta_{z_j} + (1-\lambda)\delta_{z_{j+1}} : \lambda \in [0,1]\}$.

Example 5.6.3. Let $T : [0,1] \to [0,1]$ be the Gauss map, and define $f : [0,1] \to \mathbb{R}$ by $f(x) = (x - 1/2)^2$. Since f is strictly convex, Theorem 5.6.2 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 5.6.4.

(a) A completely different approach to ergodic optimization for infinite branch maps, based on symbolic dynamics, has been considered in [14, 21, 22].

(b) The choice $T(0) = 1$ renders $\{0,1\}$ a period-2 orbit of T, so that $\mu_{01} = (\delta_0 + \delta_1)/2 \in \mathcal{M}_T$. Arguments similar to those of Theorem 5.4.1 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 $(\mathcal{M}_\varrho, \prec)$. For $\varrho \in (0, 1/2)$, however, the absence of a smallest fixed point of T can be used to show that $(\mathcal{M}_\varrho, \prec)$ has no largest element.

Bibliography

[1] C. D. Aliprantis & K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, 2007.

[2] V. Anagnostopoulou, Sturmian measures and stochastic dominance in ergodic optimization, *Ph.D. thesis*, Queen Mary, University of London, 2008.

[3] V. Anagnostopoulou, K. Díaz-Ordaz, O. Jenkinson & C. Richard, Sturmian maximising measures for the piecewise-linear family of cosine maps, *preprint*.

[4] D. Blackwell, Equivalent comparisons of experiments, *Ann. Math. Statist.*, **24** (1953), 265–272.

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

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

[7] P. Cartier, J. M. G. Fell, & P.-A. Meyer, Comparaison des mesures portées par un ensemble convexe compact, *Bull. Soc. Math. France*, **92** (1964), 435–445.

[8] K.-C. Chen & X. Dong, On the barycenter of the tent map, *Proc. Amer. Math. Soc.*, **138** (2010), 4025–4035.

[9] G. Choquet, Le théorème de représentation intégrale dans les ensembles convexes compacts, *Ann. Inst. Fourier (Grenoble)*, **10** (1960), 333–344.

[10] G. Contreras, A. Lopes, & Ph. Thieullen, Lyapunov minimizing measures for expanding maps of the circle, *Ergod. Th. & Dyn. Sys.*, **21** (2001), 1379–1409.

[11] J.-P. Conze & Y. Guivarc'h, Croissance des sommes ergodiques et principe variationnel, *manuscript*, 1993.

[12] G. H. Hardy, J. E. Littlewood, & G. Pólya, Some simple inequalities satisfied by convex functions, *Messenger Math.*, **58** (1929), 145–152.

[13] G. H. Hardy, J. E. Littlewood, & G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, (1st ed., 1934; 2nd ed., 1952).

[14] G. Iommi, Ergodic optimization for renewal type shifts, *Monatsh. Math.*, **150** (2007), 91–95.

[15] O. Jenkinson, Frequency locking on the boundary of the barycentre set, *Experimental Mathematics*, **9** (2000), 309–317.

[16] O. Jenkinson, Maximum hitting frequency and fastest mean return time, *Nonlinearity*, **18** (2005), 2305–2321.

[17] O. Jenkinson, Ergodic Optimization, *Discr. & Cont. Dyn. Sys.*, **15** (2006), 197-224

[18] O. Jenkinson, Optimization and majorization of invariant measures, *Elec. Res. Ann. Amer. Math. Soc.*, **13** (2007), 1–12.

[19] O. Jenkinson, A partial order on $\times 2$-invariant measures, *Math. Res. Lett.*, **15** (2008), 893-900.

[20] O. Jenkinson, Balanced words and majorization, *Discr. Math. Alg. & Appl.*, **1** (2009), 485–498.

[21] O. Jenkinson, R. D. Mauldin & M. Urbański, Ergodic optimization for countable alphabet subshifts of finite type, *Ergod. Th. & Dyn. Sys.*, **26** (2006), 1791–1803.

[22] O. Jenkinson, R. D. Mauldin & M. Urbański, Ergodic optimization for non-compact dynamical systems, *Dynamical Systems*, **22** (2007), 379–388.

[23] O. Jenkinson & J. Steel, Majorization of invariant measures for orientation-reversing maps, *Ergod. Th. & Dyn. Sys.*, **30** (2010), 1471–1483.

[24] J. Karamata, Sur une inégalité rélative aux fonctions convexes, *Publ. Math. Univ. Belgrade*, **1** (1932), 145–148.

[25] S. Karlin & A. Novikoff, Generalized convex inequalities, *Pacific J. Math.*, **13** (1963), 1251–1279.

[26] A. W. Marshall & I. Olkin, *Inequalities: theory of majorization and its applications*, Mathematics in science and engineering vol. 143, Academic Press, 1979.

[27] M. Misiurewicz, Rotation theory, *Proceedings RIMS Workshop on Dynamical Systems and Applications*, 2006. (`http://www.math.kyoto-u.ac.jp/~kokubu/RIMS2006/proc.html`).

[28] M. Morse & G. A. Hedlund, Symbolic Dynamics II. Sturmian Trajectories, *Amer. J. Math.*, **62** (1940), 1–42.

[29] N. Pytheas Fogg, *Substitutions in dynamics, arithmetics and combinatorics*, Springer Lecture Notes in Mathematics vol. 1794, 2002.

[30] M. Rothschild & J. Stiglitz, Increasing risk I: A definition, *J. of Economic Theory*, **2** (1970), 225–243.

[31] H. L. Royden, *Real analysis*, 3rd edn., MacMillan, New York, 1988.

[32] J.Steel, Concave unimodal maps have no majorisation relations between their ergodic measures *Proc. Amer. Math. Soc.*, to appear.

[33] J.Steel, Barycentres and majorisation for invariant measures of the shifted doubling map family, *preprint*, 2010.

[34] J. Steel, Majorisation ordering of periodic orbits of the doubling map with weights 2 and 3, *preprint*, 2010.

[35] S. Tabachnikov, *Billiards*, Panoramas et Synthèses 1, Soc. Math. France, 1995.

[36] L. Vuillon, Balanced words, *Bull. Belg. Math. Soc. Simon Stevin*, **10** (2003), 787–805.

[37] P. Walters, *An Introduction To Ergodic Theory*, Springer, 2000