Joint spectral radius, Sturmian measures and the finiteness conjecture

OLIVER JENKINSON† and MARK POLLICOTT‡

† *School of Mathematical Sciences, Queen Mary, University of London, Mile End Road,*
London E1 4NS, UK
(e-mail: omj@maths.qmul.ac.uk)

‡ *Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK*
(e-mail: mpollic@maths.warwick.ac.uk)

Abstract. The joint spectral radius of a pair of 2×2 real matrices $(A_0, A_1) \in M_2(\mathbb{R})^2$ is defined to be $r(A_0, A_1) = \limsup_{n \to \infty} \max\{\|A_{i_1} \cdots A_{i_n}\|^{1/n} : i_j \in \{0, 1\}\}$, the optimal growth rate of the norm of products of these matrices. The Lagarias–Wang finiteness conjecture [Lagarias and Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra Appl.* **214** (1995), 17–42], asserting that $r(A_0, A_1)$ is always the nth root of the spectral radius of some length-n product $A_{i_1} \cdots A_{i_n}$, has been refuted by Bousch and Mairesse [Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. *J. Amer. Math. Soc.* **15** (2002), 77–111], with subsequent counterexamples presented by Blondel *et al* [An elementary counterexample to the finiteness conjecture. *SIAM J. Matrix Anal.* **24** (2003), 963–970], Kozyakin [A dynamical systems construction of a counterexample to the finiteness conjecture. *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference (Seville, Spain, December 2005)*. IEEE, Piscataway, NJ, pp. 2338–2343] and Hare *et al* [An explicit counterexample to the Lagarias–Wang finiteness conjecture. *Adv. Math.* **226** (2011), 4667–4701]. In this article, we introduce a new approach to generating finiteness counterexamples, and use this to exhibit an open subset of $M_2(\mathbb{R})^2$ with the property that each member (A_0, A_1) of the subset generates uncountably many counterexamples of the form $(A_0, t A_1)$. Our methods employ ergodic theory; in particular, the analysis of Sturmian invariant measures. This approach allows a short proof that the relationship between the parameter t and the Sturmian parameter $\mathcal{P}(t)$ is a devil's staircase.

1. Introduction

1.1. *Problem and setting.* For a square matrix A with real entries, its *spectral radius* $r(A)$, defined as the maximum modulus of its eigenvalues, satisfies *Gelfand's formula*

$$r(A) = \lim_{n \to \infty} \|A^n\|^{1/n},$$

where $\| \cdot \|$ is a matrix norm. More generally, for a finite collection $\mathcal{A} = \{A_0, \ldots, A_l\}$ of real square matrices, all of the same size, the *joint spectral radius $r(\mathcal{A})$* is defined by

$$r(\mathcal{A}) = \limsup_{n \to \infty} \max\{\|A_{i_1} \cdots A_{i_n}\|^{1/n} : i_j \in \{0, \ldots, l\}\}, \tag{1}$$

or, equivalently (see, e.g., [10]), by

$$r(\mathcal{A}) = \lim_{n \to +\infty} \max\{r(A_{i_1} \cdots A_{i_n})^{1/n} : i_j \in \{0, \ldots, l\}\}. \tag{2}$$

The notion of joint spectral radius was introduced by Rota and Strang [19] and notably popularized by Daubechies and Lagarias [6] in their work on wavelets. Since the 1990s it has become an area of very active research interest, from both a pure and an applied perspective (see, e.g., [1, 10, 13, 21]).

The set $\mathcal{A}$ is said to have the *finiteness property* if $r(\mathcal{A}) = r(A_{i_1} \cdots A_{i_n})^{1/n}$ for some $i_1, \ldots, i_n \in \{0, \ldots, l\}$. It was conjectured by Lagarias and Wang [14] (see also Gurvits [7]) that every such $\mathcal{A}$ enjoys the finiteness property. This so-called finiteness conjecture was, however, refuted by Bousch and Mairesse [4], and a number of authors (see [2, 8, 12, 16]) have subsequently given examples of sets $\mathcal{A}$ for which the finiteness property fails. A common feature of these finiteness counterexamples has been a judicious choice of a pair of 2×2 matrices A_0, A_1, followed by an argument that, for certain $t > 0$, the finiteness property fails for the set $\mathcal{A}(t) = \{A_0^{(t)}, A_1^{(t)}\} = \{A_0, tA_1\}$.

In fact, for many of these examples, it has been observed that the family $(\mathcal{A}(t))_{t>0}$ can be associated with the class of *Sturmian* sequences of Morse and Hedlund [17]: for a given $t > 0$ an appropriate Sturmian sequence $(i_n)_{n=1}^{\infty} \in \{0, 1\}^{\mathbb{N}}$ turns out to give the optimal matrix product, in the sense that the joint spectral radius $r(\mathcal{A}(t))$ equals $\lim_{n \to \infty} r(A_{i_1}^{(t)} \cdots A_{i_n}^{(t)})^{1/n}$ (see [2, 4, 8, 12, 16] for further details). A Sturmian sequence $(i_n)_{n=1}^{\infty}$ has a well-defined *1-frequency* $\mathcal{P} = \lim_{N \to \infty} (1/N) \sum_{n=1}^{N} i_n$, and it is those sets $\mathcal{A}(t)$ whose associated Sturmian sequences† have *irrational* 1-frequency which yield counterexamples to the finiteness conjecture (see Proposition 9 below for a more precise description of the connection between finiteness counterexamples and Sturmian sequences with irrational 1-frequency). For certain such families $(\mathcal{A}(t))_{t>0}$ (which henceforth we refer to as *Sturmian families*), it has been proved by Morris and Sidorov [16] (see also [4, p. 109]) that if $\mathcal{P}(t)$ denotes the 1-frequency associated to $\mathcal{A}(t)$, then the parameter mapping $t \mapsto \mathcal{P}(t)$ is continuous and monotone, but *singular* in the sense that $\{t > 0 : \mathcal{P}(t) \notin \mathbb{Q}\}$ is nowhere dense: in other words, the uncountably many parameters t for which finiteness counterexamples occur only constitute a thin subset‡ of $\mathbb{R}^+$.

---

† We follow the definition of a Sturmian sequence given in [5], although note that some authors refer to these as *balanced* sequences, reserving the nomenclature *Sturmian* precisely for those balanced sequences with irrational 1-frequency.

‡ The belief that finiteness counterexamples are rare appears to be widespread; for example, Maesumi [15] conjectures that they constitute a set of (Lebesgue) measure zero in the space of matrices.

Examples of Sturmian families $(\mathcal{A}(t))_{t>0}$ have been given by Bousch and Mairesse [4], who considered the family generated by matrix pairs of the form

$$\mathcal{A} = \left(\begin{pmatrix} e^{\kappa h_0} + 1 & 0 \\ e^{\kappa} & 1 \end{pmatrix}, \begin{pmatrix} 1 & e^{\kappa} \\ 0 & e^{\kappa h_1} + 1 \end{pmatrix} \right), \quad \kappa > 0, h_0, h_1 > 0, h_0 + h_1 < 2, \quad (3)$$

by Kozyakin [12], who studied the family generated by pairs of the form

$$\mathcal{A} = \left(\begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \right), \quad 0 < a, d < 1 \leq bc, \quad (4)$$

and by various authors [2, 8, 16] focusing on the family generated by the particular pair

$$\mathcal{A} = \left(\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \right). \quad (5)$$

For an invertible matrix P, the simultaneous similarity $(A_0, A_1) \mapsto (P^{-1} A_0 P, P^{-1} A_1 P)$ leaves invariant the joint spectral radius, and does not change the sequences $(i_n)_{n=1}^{\infty}$ attaining the optimal matrix product, while, if $u, v > 0$, then (uA_0, vA_1) has the same optimizing sequences as $(A_0, (v/u)A_1)$. Therefore, declaring $\mathcal{A} = (A_0, A_1)$ and $\mathcal{A}' = (A_0', A_1')$ to be equivalent if $A_0' = u P^{-1} A_0 P$ and $A_1' = v P^{-1} A_1 P$ for some invertible P and $u, v > 0$, we see that the equivalence of $\mathcal{A}$ and $\mathcal{A}'$ implies that $(\mathcal{A}(t))_{t>0}$ is a Sturmian family if and only if $(\mathcal{A}'(t))_{t>0}$ is. In particular, $(\mathcal{A}'(t))_{t>0}$ is a Sturmian family whenever $\mathcal{A}'$ is equivalent to a matrix pair $\mathcal{A}$ of the form (3), (4), or (5).

The purpose of this article is to introduce an approach to studying the joint spectral radius and generating finiteness counterexamples, which, in particular, yields new examples of Sturmian families $(\mathcal{A}(t))_{t>0}$, i.e. where $\mathcal{A}$ is not equivalent to a matrix pair of the form (3), (4) or (5). Our method is conceptually different from that of previous authors, employing notions from dynamical systems, ergodic theory and, in particular, ergodic optimization (see, e.g., [9]). Specifically, we identify a dynamical system $T_{\mathcal{A}}$ with the matrix pair $\mathcal{A} = (A_0, A_1)$, and cast the problem of determining the joint spectral radius $r(\mathcal{A})$ in terms of ergodic optimization (see Theorem 2 below): it suffices to determine the $T_{\mathcal{A}}$-invariant probability measure which maximizes the integral of a certain auxiliary real-valued function $f_{\mathcal{A}}$. Working with the family of $\mathcal{A}$-*Sturmian measures* (certain probability measures invariant under $T_{\mathcal{A}}$) instead of Sturmian sequences, we exploit a characterization of these measures in terms of the smallness of their support to show that they give precisely the family of $f_{\mathcal{A}(t)}$-maximizing measures, $t > 0$. In particular, whenever the $f_{\mathcal{A}(t)}$-maximizing measure is Sturmian of *irrational* parameter $\mathcal{P}(t)$, then $\mathcal{A}(t)$ is a finiteness counterexample (cf. Proposition 9).

The $\mathcal{A}$-Sturmian measures are naturally identified with Sturmian measures on $\Omega = \{0, 1\}^{\mathbb{N}}$, the full shift on two symbols (see Notation 2). A notable feature of our approach is that the singularity of the parameter mapping $t \mapsto \mathcal{P}(t)$ (and, in particular, the fact that $\{t > 0 : \mathcal{P}(t) \notin \mathbb{Q}\}$ is nowhere dense in $\mathbb{R}^+$) is then readily deduced (see Theorem 9 in §11) as a consequence of classical facts about parameter dependence of Sturmian measures on Ω (i.e. rather than requiring the *ab initio* approach of [16]).

1.2. *Statement of results.*   We use $M_2(\mathbb{R})$ to denote the set of real 2×2 matrices, and we focus attention on certain of its open subsets.

Notation 1. $M_2(\mathbb{R}^+)$ will denote the set of *positive matrices*, i.e. matrices in $M_2(\mathbb{R})$ with entries in $\mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$, and $M_2^+(\mathbb{R}^+) = \{A \in M_2(\mathbb{R}^+) : \det A > 0\}$ will denote the set of *positive orientation-preserving matrices*.

Turning to *pairs* of matrices, we shall consider the following open subset of $M_2^+(\mathbb{R}^+)^2$.

Definition 1. Let $\mathfrak{C} \subset M_2^+(\mathbb{R}^+)^2$ denote the set of matrix pairs

$$(A_0, A_1) = \left(\begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix}, \begin{pmatrix} a_1 & b_1 \\ c_1 & d_1 \end{pmatrix} \right) \in M_2^+(\mathbb{R}^+)^2$$

satisfying

$$\frac{a_0}{c_0} < \frac{b_1}{d_1} \tag{6}$$

and

$$a_1 + c_1 - b_1 - d_1 < 0 < a_0 + c_0 - b_0 - d_0. \tag{7}$$

For reasons which will become apparent later (see Proposition 4), $\mathfrak{C}$ will be referred to as the set of *concave–convex matrix pairs*.

Finally, our counterexamples to the Lagarias–Wang finiteness conjecture will be drawn from a certain open subset $\mathfrak{D}$ (given by Definition 3 below) of $\mathfrak{C}$ which is conveniently described in terms of quantities ϱ_A and σ_A defined as follows.

Definition 2. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2^+(\mathbb{R}^+)$, we define

$$\varrho_A = \frac{2b}{a - d - 2b + \sqrt{(a-d)^2 + 4bc}},$$

and if $a + c \neq b + d$, then we define

$$\sigma_A = \frac{b - a}{a + c - b - d}.$$

It turns out (see Corollary 4) that if $(A_0, A_1) \in \mathfrak{C}$, then $\sigma_{A_0} < 0 < \varrho_{A_0}$ and $\varrho_{A_1} < -1$. The set $\mathfrak{D}$ is defined by imposing two inequalities.

Definition 3. Define

$$\mathfrak{D} = \{(A_0, A_1) \in \mathfrak{C} : \varrho_{A_1} < \sigma_{A_0} \text{ and } \sigma_{A_1} < \varrho_{A_0}\}.$$

Clearly, $\mathfrak{D}$ is an open subset of $\mathfrak{C}$, and hence also of $M_2(\mathbb{R})^2$. It is also non-empty: for example, it is readily verified that the two-parameter family

$$\mathfrak{D}_2 = \left\{ \left(\begin{pmatrix} 1 & b \\ c & 1 \end{pmatrix}, \begin{pmatrix} 1 & c \\ b & 1 \end{pmatrix} \right) : bc < 1 < c, (b, c) \in (\mathbb{R}^+)^2 \right\} \tag{8}$$

is a subset of $\mathfrak{D}$. Note that the pair (5) studied in [**2**, **8**, **16**], and corresponding to $(b, c) = (0, 1)$ in (8), lies on the boundary of both $\mathfrak{D}$ and $\mathfrak{D}_2$.

A version of our main result is the following theorem.

THEOREM 1. *The open subset $\mathfrak{D} \subset M_2(\mathbb{R})^2$ is such that if $\mathcal{A} = (A_0, A_1) \in \mathfrak{D}$, then, for uncountably many $t \in \mathbb{R}^+$, the matrix pair $(A_0, t A_1)$ is a finiteness counterexample.*

If we define $\mathfrak{E} \subset M_2(\mathbb{R})^2$ to be the set of matrix pairs which are equivalent to some pair in $\mathfrak{D}$ (recall that $\mathcal{A} = (A_0, A_1)$ and $\mathcal{A}' = (A_0', A_1')$ are equivalent if $A_0' = u P^{-1} A_0 P$ and $A_1' = v P^{-1} A_1 P$ for some invertible P and $u, v > 0$), then, clearly, we get the following corollary.

COROLLARY 1. *The open subset $\mathfrak{E} \subset M_2(\mathbb{R})^2$ is such that if $\mathcal{A} = (A_0, A_1) \in \mathfrak{E}$, then, for uncountably many $t \in \mathbb{R}^+$, the matrix pair $(A_0, t A_1)$ is a finiteness counterexample.*

Remark 1. Theorem 1 yields new finiteness counterexamples, in the sense that $\mathfrak{D}$ contains matrix pairs which are not equivalent to pairs satisfying (3), (4) or (5). To see this, note, for example, that

$$\mathcal{A} = (A_0, A_1) = \left(\begin{pmatrix} 5/8 & 3/112 \\ 7/8 & 15/16 \end{pmatrix}, \begin{pmatrix} 15/16 & 1 \\ 1/128 & 7/8 \end{pmatrix} \right) \tag{9}$$

belongs to $\mathfrak{D}$. Both A_0 and A_1 have their larger eigenvalue equal to one, and smaller eigenvalues given by $\lambda_0 = 9/16$ and $\lambda_1 = 13/16$, respectively. Now both matrices in (5) have the single eigenvalue one, so (9) cannot be equivalent to (5); moreover, (9) is not equivalent to any matrix pair satisfying (3), since both matrices in (3) have the property that the larger eigenvalue is more than double the smaller eigenvalue. Lastly, we show that $\mathcal{A}$ in (9) is not equivalent to any pair $\mathcal{A}'$ satisfying (4): i.e. $\mathcal{A}' = (A_0', A_1')$ with $A_0' = \begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}$, $A_1' = \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$, where $0 < a, d < 1 \leq bc$. Note that both A_0' and A_1' have their larger eigenvalue equal to one (as is the case for A_0 and A_1), and smaller eigenvalues equal to d and a, respectively. Thus, if $\mathcal{A}$ and $\mathcal{A}'$ were equivalent, then there would exist an invertible P such that $A_0' = P^{-1} A_0 P$ and $A_1' = P^{-1} A_1 P$ (i.e. the positive reals u, v in the above definition of equivalence must both equal one), so that $d = \lambda_0 = 9/16$, $a = \lambda_1 = 13/16$, and $\mathrm{trace}(A_0 A_1) = \mathrm{trace}(A_0' A_1')$. In particular, $1 \leq bc = \mathrm{trace}(A_0' A_1') - d - a = \mathrm{trace}(A_0 A_1) - \lambda_0 - \lambda_1 = 12\,995/14\,336 < 1$, which is a contradiction. It follows that $\mathcal{A} \in \mathfrak{D}$ given by (9) is not equivalent to any matrix pair satisfying (4).

A key tool in proving Theorem 1 is the following Theorem 2 (proved in §4 as Theorem 4) characterizing the joint spectral radius of $\mathcal{A} \in \mathfrak{C}$ in terms of maximizing the integral of a certain function $f_{\mathcal{A}}$ over the set $\mathcal{M}_{\mathcal{A}}$ of probability measures invariant under an associated mapping $T_{\mathcal{A}}$. More precisely, the action of any positive matrix A on $(\mathbb{R}^+)^2$ induces a projective map T_A (see §2.1), and if $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then the inverses $T_{A_0}^{-1}$, $T_{A_1}^{-1}$ together define a two-branch dynamical system $T_{\mathcal{A}}$ (see §4) on a subset of the unit interval X. Defining the real-valued function $f_{\mathcal{A}}$ in terms of the derivative $T_{\mathcal{A}}'$ and the characteristic functions of the images $T_{A_0}(X)$ and $T_{A_1}(X)$ by

$$f_{\mathcal{A}} = \tfrac{1}{2}(\log T_{\mathcal{A}}' + (\log \det A_0) \mathbb{1}_{T_{A_0}(X)} + (\log \det A_1) \mathbb{1}_{T_{A_1}(X)})$$

then gives the following theorem.

THEOREM 2. *If $\mathcal{A} \in \mathfrak{C}$, then*

$$\log r(\mathcal{A}) = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu. \tag{10}$$

In order to state a more precise version of Theorem 1, we first need some basic facts concerning ergodic theory, symbolic dynamics, and Sturmian measures.

Notation 2. Let $\Omega = \{0, 1\}^{\mathbb{N}}$ denote the set of one-sided sequences $\omega = (\omega_n)_{n=1}^{\infty}$, where $\omega_n \in \{0, 1\}$ for all $n \geq 1$. When equipped with the product topology, Ω becomes a compact space, and the shift map $\sigma : \Omega \to \Omega$ defined by $(\sigma\omega)_n = \omega_{n+1}$ for all $n \geq 1$ is then continuous. Let $\mathcal{M}$ denote the set of shift-invariant Borel probability measures on Ω; when equipped with the weak-$*$ topology, $\mathcal{M}$ is compact (see [**22**, Theorem 6.10]).

We equip Ω with the lexicographic order $<$ and write $[\omega^-, \omega^+] = \{\omega \in \Omega : \omega^- \leq \omega \leq \omega^+\}$. A *Sturmian interval* is one of the form $[0\omega, 1\omega]$, for some $\omega \in \Omega$. A measure $\mu \in \mathcal{M}$ is called *Sturmian* (see, e.g., [**4**, Proposition 1.5], [**5**]) if its support is contained in a Sturmian interval. Let $\mathcal{S} \subset \mathcal{M}$ denote the class of Sturmian measures on Ω. For a Sturmian measure $\mu \in \mathcal{S}$, the value $\mu([1])$, denoted by $\mathcal{P}(\mu)$, is called its *(Sturmian) parameter*†, where $[1]$ denotes the (cylinder) set $\{\omega \in \Omega : \omega_1 = 1\}$. A *Sturmian sequence* of parameter $\mathcal{P}$ is any point in the support of the Sturmian measure of parameter $\mathcal{P}$.

The following are classical facts about Sturmian measures (see, e.g., [**4**, §1.1] or [**5**]).

PROPOSITION 1.
(a) *For each Sturmian interval $[0\omega, 1\omega] \subset \Omega$, there exists a unique Sturmian measure whose support is contained in this interval.*
(b) *The mapping $\mathcal{P} : \mathcal{S} \to [0, 1]$ is a homeomorphism. If $\mu \in \mathcal{S}$ has $\mathcal{P}(\mu) \in \mathbb{Q}$, then its support is a single σ-periodic orbit, while if $\mathcal{P}(\mu) \notin \mathbb{Q}$, then its support is a Cantor subset of Ω which supports no other σ-invariant measure (and, in particular, contains no periodic orbit).*
(c) *If $d(\omega)$ denotes the Sturmian parameter of the Sturmian measure supported by the Sturmian interval $[0\omega, 1\omega] \subset \Omega$, then the map $d : \Omega \to [0, 1]$ is continuous, non-decreasing and surjective. The preimage $d^{-1}(\mathcal{P})$ is a singleton if $\mathcal{P}$ is irrational, and a positive-length closed interval if $\mathcal{P}$ is rational.*

For example, the Sturmian measures of parameter $1/2$, $1/3$, $2/5$, $3/8$ and $5/13$ are, respectively, supported by the σ-periodic orbits generated by the finite words

$$01, \quad 001, \quad 00101, \quad 00100101, \quad 0010010100101,$$

whereas the Sturmian measure of parameter $(3 - \sqrt{5})/2$ is supported by the smallest Cantor set containing the σ-orbit of

$$0010010100100101001010010010100101 \ldots .$$

In view of Theorem 2, for a matrix pair $\mathcal{A} \in \mathfrak{C}$ we are interested in measures $\nu \in \mathcal{M}_{\mathcal{A}}$ attaining the maximum in (10), i.e. satisfying $\int f_{\mathcal{A}} \, d\nu = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu$; such ν will be called $f_{\mathcal{A}}$-*maximizing*. There is a topological conjugacy between $T_{\mathcal{A}}$ and the shift map $\sigma : \Omega \to \Omega$, and this induces a natural homeomorphism between $\mathcal{M}_{\mathcal{A}}$ and $\mathcal{M}$; the image of any $f_{\mathcal{A}}$-maximizing measure under this homeomorphism will be called a *maximizing measure for $\mathcal{A}$*. We then say that $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$ *generates a full Sturmian*

† This corresponds to the *1-frequency* mentioned in §1.1, sometimes called the *1-ratio* (see, e.g., [**8**, **16**]), or the *rotation number* (see, e.g., [**5**]).

family if the set of maximizing measures for the family $\mathcal{A}(t) = (A_0, tA_1)$, $t \in \mathbb{R}^+$, is precisely the set $\mathcal{S}$ of all Sturmian measures on Ω.

A more precise version of our main result Theorem 1 is then the following.

THEOREM 3. *Every matrix pair in the open subset $\mathfrak{D} \subset M_2(\mathbb{R})^2$ (and hence the open subset $\mathfrak{E} \subset M_2(\mathbb{R})^2$) generates a full Sturmian family.*

Note that Theorem 3 will follow from a more detailed version, Theorem 9, which, in particular, incorporates the statement that the parameter map $t \mapsto \mathcal{P}(t)$ is a devil's staircase.

1.3. *Relationship with previous results.* The methods of this paper can also be used to give an alternative proof of some of the results mentioned above, namely, establishing the analogue of Theorem 3 in certain cases treated by Bousch and Mairesse [**4**] and Kozyakin [**12**], and the case considered by Blondel, Theys and Vladimirov [**2**], Hare, Morris, Sidorov and Theys [**8**], and Morris and Sidorov [**16**].

As already noted, the matrix pair (5) lies on the boundary of our open set $\mathfrak{D}$, and clearly it also lies on the boundary of the set $\mathfrak{K} \subset M_2^+(\mathbb{R}^+)^2$ defined by Kozyakin's conditions (4). It can be checked that $\mathfrak{K}$ itself lies in the boundary of our set $\mathfrak{E}$, but not in the boundary of $\mathfrak{D}$. However, the subset $\mathfrak{K}' \subset \mathfrak{K}$, defined by

$$\mathfrak{K}' = \left\{ \left(\begin{pmatrix} 1 & 0 \\ c & d \end{pmatrix}, \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \right) \in \mathfrak{K} : a \leq b \text{ and } d \leq c \right\}, \tag{11}$$

can be readily checked to lie in the boundary of $\mathfrak{D}$. Matrices in the Bousch–Mairesse family (3) do not all satisfy our condition (6), or indeed the corresponding weak inequality, so do not automatically belong to the boundary of $\mathfrak{E}$. However, imposing the additional condition

$$e^{2\kappa} \geq (e^{\kappa h_0} + 1)(e^{\kappa h_1} + 1) \tag{12}$$

ensures that a matrix pair satisfying (3) belongs to the boundary of $\mathfrak{E}$, and indeed also belongs to the boundary of $\mathfrak{D}$. In §7.2, we will indicate the minor modifications to our approach that are needed to handle the case of (5) and the sub-cases of (3) and (4) defined by (12) and (11), respectively.

1.4. *Organization of article.* The article is organized as follows. Section 2 consists of preliminaries: maps induced by matrices acting on projective space, Perron–Frobenius theory and some useful notation and identities. Section 3 develops the notions of projective convexity and projective concavity. Section 4 introduces the induced dynamical system $T_{\mathcal{A}}$ for concave–convex matrix pairs $\mathcal{A}$, the formulation of joint spectral radius in terms of ergodic optimization (Theorem 4) and the connection between the finiteness property and $T_{\mathcal{A}}$-periodic orbits. Section 5 introduces Sturmian measures and Sturmian intervals for the dynamical system $T_{\mathcal{A}}$ and makes the connection between finiteness counterexamples and unique maximizing measures which are Sturmian of irrational parameter. Section 6 establishes the existence of an important technical tool, the Sturmian transfer function. After deriving some explicit formulae for extremal Sturmian intervals in §7, the key §8 establishes the link between Sturmian intervals and the parameter t of the pair $\mathcal{A}(t)$.

Section 9 treats the case of those parameters t such that one matrix in the pair $\mathcal{A}(t)$ dominates the other, so that the joint spectral radius $r(\mathcal{A}(t))$ is simply the spectral radius of the dominating matrix. All other parameters are considered in §10, establishing that the joint spectral radius is always attained by a unique Sturmian measure. Finally, in §11, we show that the map taking parameter values t to the associated Sturmian parameter $\mathcal{P}(t)$ is a devil's staircase.

2. Preliminaries

2.1. The induced map for a positive matrix.

Notation 3. Throughout we use the notation $X = [0, 1]$.

A positive matrix $A \in M_2(\mathbb{R}^+)$ gives a self-map $v \mapsto Av$ of $(\mathbb{R}^+)^2$. This lifts to a self-map $\widetilde{A} : [v] \mapsto [Av]$ of projective space $(\mathbb{R}^+)^2 / \sim$, the equivalence relation $\sim$ being defined by $v \sim v'$ if $v = sv'$ for some $s > 0$, and where $[v]$ denotes the equivalence class containing $v \in (\mathbb{R}^+)^2$. It is convenient to identify the projective space with

$$\Sigma = \left\{ \begin{pmatrix} x \\ 1 - x \end{pmatrix} : x \in (0, 1) \right\},$$

so that the projection $\pi : (\mathbb{R}^+)^2 \to \Sigma$ takes the form

$$\pi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \frac{x}{x+y} \\ \frac{y}{x+y} \end{pmatrix},$$

and the projective map is represented as $\pi \circ A : \Sigma \to \Sigma$, taking the explicit form

$$\pi \circ A : \begin{pmatrix} x \\ 1 - x \end{pmatrix} \mapsto \begin{pmatrix} \frac{(a-b)x+b}{(a+c-b-d)x+b+d} \\ \frac{(c-d)x+d}{(a+c-b-d)x+b+d} \end{pmatrix}.$$

This projective mapping is completely determined by its first coordinate, thereby motivating the following definition of the self-map T_A of the unit interval $X = [0, 1]$.

Definition 4. For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2^+(\mathbb{R}^+)$, the *induced map* $T_A : X \to X$ is defined by

$$T_A(x) = \frac{(a - b)x + b}{(a + c - b - d)x + b + d},$$

the *induced image* X_A is defined by

$$X_A = T_A(X) = \left[\frac{b}{b + d}, \frac{a}{a + c} \right],$$

and the *induced inverse map* $S_A : X_A \to X$ is given by

$$S_A(x) = T_A^{-1}(x) = \frac{(b + d)x - b}{-(a + c - b - d)x + a - b}.$$

Remark 2. Defining $P = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, the Möbius maps T_A and S_A are represented, respectively, by the matrices PAP^{-1} and $PA^{-1}P^{-1}$.

Remark 3. The objects defined in Definition 4 do not change if the matrix A is multiplied by a positive real number: i.e. if $t > 0$, $A \in M_2^+(\mathbb{R}^+)$, then $T_{tA} = T_A$ (hence $S_{tA} = S_A$) and $X_{tA} = X_A$.

In view of (7), in the definition of $\mathfrak{C}$, it suffices to restrict attention to matrices of the following form.

Notation 4. Let $\mathbb{M}$ denote the set of matrices $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2^+(\mathbb{R}^+)$ such that $a + c \neq b + d$.

LEMMA 1. *For $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2^+(\mathbb{R}^+)$, the map T_A has a single fixed point $p_A = T_A(p_A)$ in X. If $A \in \mathbb{M}$, then*

$$p_A = \frac{a - d - 2b + \sqrt{(a-d)^2 + 4bc}}{2(a + c - b - d)}, \tag{13}$$

and if $A \notin \mathbb{M}$, then

$$p_A = \frac{b}{2b + d - a}. \tag{14}$$

Proof. Uniqueness follows from the fact that A has all entries strictly positive, and the formulae (13) and (14) are straightforward computations. □

2.2. *Notation and matrix preliminaries.* For a matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in M_2^+(\mathbb{R}^+)$, it will be useful to write

$$\alpha_A = a + c - b - d, \tag{15}$$
$$\beta_A = a - d - 2b, \tag{16}$$
$$\gamma_A = \sqrt{(a-d)^2 + 4bc}, \tag{17}$$

noting that these quantities are related by the following identity.

LEMMA 2. *For $A \in M_2^+(\mathbb{R}^+)$,*

$$\gamma_A^2 - \beta_A^2 = 4b\alpha_A. \tag{18}$$

Proof. This is a straightforward computation. □

For ease of reference, it will be convenient to collect together various previously defined objects expressed in terms of the above notation.

PROPOSITION 2. *For $A \in M_2^+(\mathbb{R}^+)$,*

$$\varrho_A = \frac{2b}{\beta_A + \gamma_A}, \tag{19}$$
$$T_A(x) = \frac{(a - b)x + b}{\alpha_A x + b + d},$$
$$S_A(x) = \frac{(b + d)x - b}{-\alpha_A(x + \sigma_A)},$$

and if, moreover, $A \in \mathbb{M}$, then

$$\sigma_A = \frac{b - a}{\alpha_A}, \tag{20}$$

$$p_A = \frac{\beta_A + \gamma_A}{2\alpha_A} = \frac{b\sigma_A}{(b - a)\varrho_A}. \tag{21}$$

The set $\mathfrak{C}$ can be written as

$$\mathfrak{C} = \left\{ (A_0, A_1) \in \mathbb{M}^2 : \frac{a_0}{c_0} < \frac{b_1}{d_1} \text{ and } \alpha_{A_1} < 0 < \alpha_{A_0} \right\}.$$

2.3. *Perron–Frobenius theory and the joint spectral radius.*

LEMMA 3. *The dominant (Perron–Frobenius) eigenvalue $\lambda_A > 0$ of the matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2^+(\mathbb{R}^+)$ is given by*

$$\lambda_A = \frac{1}{2}(a + d + \gamma_A) = \frac{b}{p_A} + a - b,$$

with corresponding left eigenvector

$$w_A = (a - d + \gamma_A, 2b)$$

and right eigenvector

$$v_A = \begin{pmatrix} p_A \\ 1 - p_A \end{pmatrix}.$$

For $A \in M_2^+(\mathbb{R}^+)$, the derivative of T_A at its fixed point p_A is related to the determinant and Perron–Frobenius eigenvalue of A as follows.

LEMMA 4. *If $A \in M_2^+(\mathbb{R}^+)$, then*

$$T_A'(p_A) = \frac{\det A}{\lambda_A^2}.$$

Proof. This is a straightforward computation. If $A \in \mathbb{M}$, we can use the expression $p_A = (\beta_A + \gamma_A)/2\alpha_A$ (see (21)), the derivative formula $T_A'(x) = \det A(\alpha_A x + b + d)^{-2}$ and the fact that $\lambda_A = \frac{1}{2}(a + d + \gamma_A) = \frac{1}{2}(\beta_A + \gamma_A) + b + d$ (see (3)). If $A \notin \mathbb{M}$, then $T_A' \equiv (a - b)/(b + d)$, $\lambda_A = b + d$, and the relation $a + c = b + d$ means that $\det A = (a - b)(b + d)$, so the result follows. □

Since the Perron–Frobenius eigenvalue λ_A is also the spectral radius $r(A)$, we obtain the following corollary.

COROLLARY 2. *If $A \in M_2^+(\mathbb{R}^+)$, then its spectral radius $r(A)$ satisfies*

$$r(A) = \left(\frac{\det A}{T_A'(p_A)} \right)^{1/2}. \tag{22}$$

Proof. This is immediate from Lemma 4. □

Notation 5. Let us write finite words using the alphabet $\{0, 1\}$ as $\underline{i} = (i_1, \ldots, i_n)$ and their length as $|\underline{i}| = n$. Let Ω^* denote the set of all such finite words: i.e. $\Omega^* = \bigcup_{n \geq 1} \{0, 1\}^n$. Given $\mathcal{A} = (A_0, A_1) \in M_2(\mathbb{R})^2$ and $\underline{i} \in \{0, 1\}^n$, let $A(\underline{i})$ denote the product

$$A(\underline{i}) = A_{i_1} \cdots A_{i_n}. \tag{23}$$

Corollary 2 then allows us to express the joint spectral radius of a matrix pair $\mathcal{A} = (A_0, A_1) \in M_2^+(\mathbb{R}^+)^2$ in terms of induced maps of the products $A(\underline{i})$ as follows.

PROPOSITION 3. *If $\mathcal{A} = (A_0, A_1) \in M_2^+(\mathbb{R}^+)^2$, then its joint spectral radius $r(\mathcal{A})$ satisfies*

$$r(\mathcal{A}) = \sup_{\underline{i} \in \Omega^*} \left(\frac{\det A(\underline{i})}{T_{A(\underline{i})}'(p_{A(\underline{i})})} \right)^{1/2|\underline{i}|}. \tag{24}$$

Proof. The expression (2) for the joint spectral radius can be written as

$$r(\mathcal{A}) = \sup\{r(A_{i_1} \cdots A_{i_n})^{1/n} : n \geq 1, i_j \in \{0, 1\}\} = \sup_{\underline{i} \in \Omega^*} r(A(\underline{i}))^{1/|\underline{i}|},$$

so applying Corollary 2 with A replaced by $A(\underline{i})$ yields the result. □

2.4. *Some useful formulae.* The purpose of this short subsection is to collect together various formulae which will prove useful in the subsequent work. Firstly, we have the following two expressions for the determinant of A involving α_A and σ_A.

LEMMA 5. *For $A \in \mathbb{M}$, its determinant can be expressed as*

$$\det A = -\alpha_A(a + c)\left(\frac{a}{a+c} + \sigma_A\right) \tag{25}$$

and

$$\det A = -\alpha_A(b + d)\left(\frac{b}{b+d} + \sigma_A\right). \tag{26}$$

Proof. This is a straightforward computation. □

There is a useful alternative way of expressing the quantity ϱ_A.

LEMMA 6. *For $A \in \mathbb{M}$,*

$$\varrho_A = \frac{\gamma_A - \beta_A}{2\alpha_A}, \tag{27}$$

and ϱ_A is the larger root of the quadratic polynomial q_A defined by

$$q_A(z) = \alpha_A z^2 + \beta_A z - b. \tag{28}$$

Proof. The expression (27) follows from (19) and the identity (18).

The larger root of q_A is computed to be

$$\frac{-\beta_A + \sqrt{\beta_A^2 + 4\alpha_A b}}{2\alpha_A} = \frac{-\beta_A + \gamma_A}{2\alpha_A},$$

again using (18). □

Clearly,

$$q_A(z) = \det\begin{pmatrix} 1 & z \\ -\alpha_A z & \beta_A z - b \end{pmatrix}, \tag{29}$$

although the following expression will prove to be more useful.

LEMMA 7. *For $A \in \mathbb{M}$,*

$$q_A(z) = \det\begin{pmatrix} 1 & z \\ b + d - \alpha_A z & (a - b)z - b \end{pmatrix}.$$

Proof. This is a straightforward computation. □

3. *Projective convexity and projective concavity*

Remark 4.

(a) For $A \in \mathbb{M}$ and $x \in X$, the derivative formula

$$T'_A(x) = \det A(\alpha_A x + b + d)^{-2} \tag{30}$$

implies that if $\mathcal{A} \in \mathbb{M}^2$, then T_{A_0} and T_{A_1} are orientation preserving.

(b) For $A \in \mathbb{M}$ and $x \in X$, the second derivative formula

$$T''_A(x) = -2\alpha_A \det A(\alpha_A x + b + d)^{-3} \tag{31}$$

implies that if $\mathcal{A} \in \mathfrak{C}$, then $T''_{A_0} < 0$ and $T''_{A_1} > 0$: i.e. T_{A_0} is strictly concave and T_{A_1} is strictly convex.

Part (b) of Remark 4 motivates the following definition, partitioning $\mathbb{M}$ into two subsets.

Definition 5. A matrix $A \in \mathbb{M}$ will be called *projectively convex* if the induced map T_A is strictly convex and *projectively concave* if the induced map T_A is strictly concave.

Remark 5. The set $\mathbb{M}$ is the disjoint union of the subset of projectively convex matrices and the subset of projectively concave matrices.

Recall that

$$w_A = (w_A^{(1)}, w_A^{(2)}) = (a - d + \gamma_A, 2b) \tag{32}$$

denotes the Perron–Frobenius left eigenvector of $A \in \mathbb{M}$ and that (consequently) the right eigenvector for the other eigenvalue of A is $\begin{pmatrix} w_A^{(2)} \\ -w_A^{(1)} \end{pmatrix}$. It is useful to record the following identity.

LEMMA 8. *For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}$,*

$$\varrho_A = \frac{w_A^{(2)}}{w_A^{(1)} - w_A^{(2)}}. \tag{33}$$

Proof. This is immediate from (19) and (32). □

COROLLARY 3. *For $A \in \mathbb{M}$, if $Q \in M_2(\mathbb{R})$ is non-singular, then $\varrho_{Q^{-1}AQ} = \varrho_A$: i.e. ϱ_A is invariant under similarities.*

Proof. This is immediate from Lemma 8 and the fact that the eigenvector w_A is invariant under similarities. □

There are various useful characterizations of projective convexity and projective concavity.

LEMMA 9. *For $A \in \mathbb{M}$, the following are equivalent:*

(i) *A is projectively concave;*

(ii) *$\alpha_A > 0$;*

(iii) *$\varrho_A > 0$;*

(iv) *$w_A^{(1)} > w_A^{(2)}$.*

Proof. As noted in Remark 4(b), the second derivative formula (31) yields the equivalence of (i) and (ii), since $\det A > 0$, and a function is strictly concave if and only if its second derivative is strictly negative.

To prove the equivalence of (ii) and (iii), we consider separately the cases where $\beta_A \geq 0$ and $\beta_A < 0$. If $\beta_A \geq 0$, then $\alpha_A = \beta_A + b + c > 0$, so we must simply show that $\varrho_A > 0$. But $\gamma_A > 0$, by definition, and hence $\beta_A + \gamma_A > 0$, and therefore (19) implies that $\varrho_A = 2b/(\beta_A + \gamma_A) > 0$, as required. If, on the other hand, $\beta_A < 0$, then $\gamma_A - \beta_A > 0$ is automatically true, again since $\gamma_A > 0$, by definition. Using (18) and (19), we see that

$$2\alpha_A \varrho_A = \gamma_A - \beta_A > 0,$$

so, indeed, $\alpha_A > 0$ if and only if $\varrho_A > 0$, as required.

Lastly, the equivalence of (iii) and (iv) is immediate from (33), since $w_A^{(2)} > 0$. □

LEMMA 10. *For $A \in \mathbb{M}$, the following are equivalent:*
(i)   *A is projectively convex;*
(ii) $\alpha_A < 0$;
(iii) $\varrho_A < -1$;
(iv) $w_A^{(1)} < w_A^{(2)}$.

Proof. A function is strictly convex if and only if its second derivative is strictly positive, so the equivalence of (i) and (ii) follows from (31) since $\det A > 0$ and $\alpha_A x + b + d = a + c + (b + d)(1 - x) > 0$ for all $x \in X$.

To prove that (iii) is equivalent to (iv), note that (33) gives $w_A^{(1)} = w_A^{(2)}(1 + \varrho_A^{-1})$; therefore $\varrho_A < -1$ if and only if $1 + \varrho_A^{-1} \in (0, 1)$, if and only if $w_A^{(1)} \in (0, w_A^{(2)})$.

Lastly, to prove the equivalence of (ii) and (iii), it follows from Lemma 9 that $\alpha_A < 0$ if and only if $\varrho_A < 0$, but this latter inequality, in fact, implies that $w_A^{(2)} - w_A^{(1)} > 0$ by (33), so

$$\varrho_A = -1 - \frac{w_A^{(1)}}{w_A^{(2)} - w_A^{(1)}} < -1,$$

as required. □

Note that, in Lemma 10, the assertion is not merely that $\varrho_A < 0$, but that $\varrho_A < -1$; this should be contrasted with the inequality $\varrho_A > 0$ in Lemma 9.

It is now clear why $\mathfrak{C}$ is described as the set of *concave–convex pairs*†.

PROPOSITION 4. *The set $\mathfrak{C}$ consists of those matrix pairs $(A_0, A_1) \in \mathbb{M}^2$ such that A_0 is projectively concave, A_1 is projectively convex and the induced image for A_0 is strictly to the left of the induced image of A_1.*

Proof. Lemmas 9 and 10 imply that the inequality $\alpha_{A_1} < 0 < \alpha_{A_0}$ in Definition 1 is equivalent to A_0 being projectively concave and A_1 being projectively convex. The inequality $a_0/c_0 < b_1/d_1$ in Definition 1 is equivalent to $T_{A_0}(1) = a_0/(a_0 + c_0) < b_1/(b_1 + d_1) = T_{A_1}(0)$, which asserts that the right endpoint of the induced image X_{A_0} is strictly to the left of the left endpoint of the induced image X_{A_1}. □

† Note, however, the restriction that the induced images be *disjoint*, with the concave image to the *left* of the convex one.

LEMMA 11. *If $A \in \mathbb{M}$ is projectively concave then $\sigma_A < 0$.*

Proof. Projective concavity of A means that $\alpha_A > 0$, so by (20) it suffices to show that $b < a$. Since $\det A = ad - bc > 0$ and $\alpha_A = a + c - b - d > 0$, we derive

$$a - b > d - c > \frac{bc}{a} - c = -\frac{c}{a}(a - b),$$

or, in other words,

$$(a - b)\left(1 + \frac{c}{a}\right) > 0,$$

and hence $a - b > 0$, as required. $\qquad\square$

We can now prove the following result mentioned in §1.2 (note, however, that there is no constraint on the sign of σ_{A_1} when $(A_0, A_1) \in \mathfrak{C}$).

COROLLARY 4. *If $(A_0, A_1) \in \mathfrak{C}$, then $\sigma_{A_0} < 0 < \varrho_{A_0}$ and $\varrho_{A_1} < -1$.*

Proof. This is immediate from Lemmas 9, 10 and 11 $\qquad\square$

An important result is the following lemma.

LEMMA 12. *If $A \in \mathbb{M}$, then*

$$-\alpha_A(x + \sigma_A) > 0 \quad \text{for all } x \in X_A. \tag{34}$$

In particular, if $A \in \mathbb{M}$ is projectively concave, then

$$x + \sigma_A < 0 \quad \text{for all } x \in X_A, \tag{35}$$

and if $A \in \mathbb{M}$ is projectively convex, then

$$x + \sigma_A > 0 \quad \text{for all } x \in X_A. \tag{36}$$

Proof. Clearly, (34) follows from (35) and (36) since α_A is positive if A is projectively concave and negative if A is projectively convex, by Lemmas 9 and 10.

To prove (35), note that $-\alpha_A(a + c)(a/(a + c) + \sigma_A) = \det A > 0$ by (25), and if A is projectively concave, then $\alpha_A > 0$, so $a/(a + c) + \sigma_A < 0$. But $a/(a + c)$ is the right-hand endpoint of X_A, so if $x \in X_A$, then $x \leq a/(a + c)$, and therefore $x + \sigma_A \leq a/(a + c) + \sigma_A < 0$, as required.

To prove (36), note that $-\alpha_A(b + d)(b/(b + d) + \sigma_A) = \det A > 0$ by (26), and if A is projectively convex, then $\alpha_A < 0$, so $b/(b + d) + \sigma_A > 0$. But $b/(b + d)$ is the left-hand endpoint of X_A, so if $x \in X_A$, then $x \geq b/(b + d)$, and therefore $x + \sigma_A \geq b/(b + d) + \sigma_A > 0$, as required. $\qquad\square$

COROLLARY 5. *If $(A_0, A_1) \in \mathfrak{C}$, then $x + \sigma_{A_0} < 0$ for $x \in X_{A_0}$, $x + \sigma_{A_1} > 0$ for $x \in X_{A_1}$ and $-\alpha_{A_i}(x + \sigma_{A_i}) > 0$ for all $x \in X_{A_i}$, $i \in \{0, 1\}$.*

Proof. This is immediate from Lemma 12. $\qquad\square$

LEMMA 13. *If $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then*

$$q_{A_1}(\varrho_{A_0}) < 0 < q_{A_0}(\varrho_{A_1}).$$

Proof. The larger root of q_{A_1} is ϱ_{A_1}, by Lemma 6. It follows that $q_{A_1}(z) = \alpha_{A_1} z^2 + \beta_{A_1} z - b_1 < 0$ for all $z > \varrho_{A_1}$, since the leading coefficient $\alpha_{A_1} < 0$, since A_1 is projectively convex. But, by Lemmas 9 and 10, we know that $\varrho_{A_0} > 0 > -1 > \varrho_{A_1}$, so, indeed, $q_{A_1}(\varrho_{A_0}) < 0$, as required.

The smaller root of q_{A_0}, which we shall denote by r_{A_0}, is given by

$$r_{A_0} = \frac{-(\gamma_{A_0} + \beta_{A_0})}{2\alpha_{A_0}}.$$

It follows that

$$q_{A_0}(z) = \alpha_{A_0} z^2 + \beta_{A_0} z - b_0 > 0 \quad \text{for all } z < r_{A_0}, \tag{37}$$

since the leading coefficient $\alpha_{A_0} > 0$, since A_0 is projectively concave. Now $\varrho_{A_1} < -1$ by Lemma 10, and if we can show that $r_{A_0} > -1$, then it follows that $\varrho_{A_1} < r_{A_0}$, and hence $q_{A_0}(\varrho_{A_1}) > 0$ by (37).

To show that, indeed, $r_{A_0} > -1$, note that this inequality is equivalent to $2\alpha_A - \beta_A > \gamma_A$. Both sides are positive, so this is equivalent to $(2\alpha_A - \beta_A)^2 > \gamma_A^2$, which, using (18), becomes $4\alpha_A(\alpha_A - \beta_A) > 4b\alpha_A$. This latter inequality is equivalent to $\alpha_A - \beta_A > b$, which is true because, in fact, $\alpha_A - \beta_A = b + c > b$. $\qquad\square$

We deduce the following technical lemma, which will be used in §9.

LEMMA 14. *If $\mathcal{A} \in \mathfrak{C}$, then the Möbius function*

$$x \mapsto \frac{x + \varrho_{A_0}}{(b_1 + d_1 - \alpha_{A_1}\varrho_{A_0})x + (a_1 - b_1)\varrho_{A_0} - b_1}$$

has strictly negative derivative, while the Möbius function

$$x \mapsto \frac{x + \varrho_{A_1}}{(b_0 + d_0 - \alpha_{A_0}\varrho_{A_1})x + (a_0 - b_0)\varrho_{A_1} - b_0}$$

has strictly positive derivative.

Proof. A general Möbius map $x \mapsto (Px + Q)/(Rx + S)$ has derivative $D(Rx + S)^{-2}$, where $D = PS - QR$, so the derivative is strictly negative if $D < 0$ and strictly positive if $D > 0$. For our first Möbius map,

$$D = \det\begin{pmatrix} 1 & \varrho_{A_0} \\ b_1 + d_1 - \alpha_{A_1}\varrho_{A_0} & (a_1 - b_1)\varrho_{A_0} - b_1 \end{pmatrix} = q_{A_1}(\varrho_{A_0})$$

by Lemma 7, and $q_{A_1}(\varrho_{A_0})$ is strictly negative by Lemma 13, so the derivative of the map is strictly negative, as required.

For our second Möbius map,

$$D = \det\begin{pmatrix} 1 & \varrho_{A_1} \\ b_0 + d_0 - \alpha_{A_0}\varrho_{A_1} & (a_0 - b_0)\varrho_{A_1} - b_0 \end{pmatrix} = q_{A_0}(\varrho_{A_1})$$

by Lemma 7, and $q_{A_0}(\varrho_{A_1})$ is strictly positive by Lemma 13, so the derivative of the map is strictly positive, as required. $\qquad\square$

4. *The induced dynamical system for a concave–convex matrix pair*

4.1. *The induced dynamical system and joint spectral radius.*

Definition 6. For a matrix pair $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, define the *induced space* $X_{\mathcal{A}}$ to be

$$X_{\mathcal{A}} = X_{A_0} \cup X_{A_1},$$

and define the *induced dynamical system* $T_{\mathcal{A}} : X_{\mathcal{A}} \to X$ by

$$T_{\mathcal{A}}(x) = \begin{cases} S_{A_0}(x) & \text{if } x \in X_{A_0}, \\ S_{A_1}(x) & \text{if } x \in X_{A_1}. \end{cases}$$

Remark 6.

(a) The map $T_{\mathcal{A}} : X_{\mathcal{A}} \to X$ is Lipschitz continuous since $X_{\mathcal{A}} = X_{A_0} \cup X_{A_1}$ is the union of disjoint intervals X_{A_0} and X_{A_1}, and the restriction of $T_{\mathcal{A}}$ to X_{A_i} is the Möbius mapping S_{A_i}, which is certainly Lipschitz continuous.

(b) Note that the (surjective) induced dynamical system $T_{\mathcal{A}}$ is naturally defined as a mapping from $X_{\mathcal{A}}$ to $X = [0, 1]$. To view it as a surjective *self*-mapping of some set (the natural setting for a dynamical system), we consider its restriction to the *induced Cantor set* $Y_{\mathcal{A}} := \bigcap_{n \geq 0} T_{\mathcal{A}}^{-n}(X_{\mathcal{A}})$ and note that $T_{\mathcal{A}} : Y_{\mathcal{A}} \to Y_{\mathcal{A}}$ is topologically conjugate to the shift map σ on $\Omega = \{0, 1\}^{\mathbb{N}}$.

PROPOSITION 5. *If $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then its joint spectral radius $r(\mathcal{A})$ satisfies*

$$r(\mathcal{A}) = \sup_{\underline{i} \in \Omega^*} (\det A(\underline{i}) (T_{\mathcal{A}}^{|\underline{i}|})'(p_{A(\underline{i})}))^{1/2|\underline{i}|}. \tag{38}$$

Proof. From Definition 6, we see that $T_{\mathcal{A}} \circ T_{A_i}$ is the identity map on X, for $i \in \{0, 1\}$, since $T_{\mathcal{A}}$ is defined in terms of the inverses $S_{A_i} = T_{A_i}^{-1}$. Similarly, for any $\underline{i} \in \{0, 1\}^n$, we see that $T_{\mathcal{A}}^n \circ T_{A(\underline{i})}$ is also the identity map on X, so $(T_{\mathcal{A}}^n)'(T_{A(\underline{i})}(x)) T_{A(\underline{i})}'(x) = 1$ for all $x \in X$, by the chain rule. Setting $x = p_{A(\underline{i})} = T_{\mathcal{A}}(p_{A(\underline{i})})$,

$$(T_{\mathcal{A}}^n)'(p_{A(\underline{i})}) = \frac{1}{T_{A(\underline{i})}'(p_{A(\underline{i})})}, \tag{39}$$

and combining this with (24) gives the required formula (38). □

4.2. *Invariant measures for the induced dynamical system.*

Definition 7. For $\mathcal{A} \in \mathfrak{C}$, let $\mathcal{M}_{\mathcal{A}}$ denote the set of $T_{\mathcal{A}}$-invariant Borel probability measures on $X = [0, 1]$; the support of any such measure is contained in $Y_{\mathcal{A}} = \bigcap_{n \geq 0} T_{\mathcal{A}}^{-n}(X_{\mathcal{A}})$.

The following is a well-known consequence of the compactness of $Y_{\mathcal{A}}$ and continuity of $T_{\mathcal{A}}$ (see, e.g., [**22**, Theorem 6.10]).

LEMMA 15. *The set $\mathcal{M}_{\mathcal{A}}$ is compact with respect to the weak* topology.*

Definition 8. For any $p \in X_{\mathcal{A}}$ that is a periodic point for $T_{\mathcal{A}}$, with $T_{\mathcal{A}}^n(p) = p$, we say that the probability measure $\mu \in \mathcal{M}_{\mathcal{A}}$ defined by

$$\mu = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T_{\mathcal{A}}^j(p)} \tag{40}$$

is the corresponding *periodic orbit measure* (or *$T_{\mathcal{A}}$-periodic orbit measure*).

Remark 7. The topological conjugacy $h_{\mathcal{A}} : \Omega \to Y_{\mathcal{A}}$ between the shift map $\sigma : \Omega \to \Omega$ and $T_{\mathcal{A}} : Y_{\mathcal{A}} \to Y_{\mathcal{A}}$ (cf. Remark 6(b)) induces a one-to-one correspondence $h_{\mathcal{A}}^* : \mathcal{M} \to \mathcal{M}_{\mathcal{A}}$ between invariant measures.

Definition 9. For a bounded Borel function $f : X_{\mathcal{A}} \to \mathbb{R}$, a measure $m \in \mathcal{M}_{\mathcal{A}}$ is called *f-maximizing* if

$$\int f \, dm = \sup_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f \, d\mu.$$

In the generality of Definition 9, the notion of an f-maximizing invariant measure is part of the wider field of so-called ergodic optimization (see, e.g., [9]).

Definition 10. For $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, define the *induced function* $f_{\mathcal{A}} : X_{\mathcal{A}} \to \mathbb{R}$ by

$$f_{\mathcal{A}} = \tfrac{1}{2}(\log T_{\mathcal{A}}' + (\log \det A_0)\mathbb{1}_{X_{A_0}} + (\log \det A_1)\mathbb{1}_{X_{A_1}}). \tag{41}$$

That is,

$$f_{\mathcal{A}}(x) = \begin{cases} \tfrac{1}{2}(\log S_{A_0}'(x) + \log \det A_0) & \text{if } x \in X_{A_0}, \\ \tfrac{1}{2}(\log S_{A_1}'(x) + \log \det A_1) & \text{if } x \in X_{A_1}, \end{cases} \tag{42}$$

so writing $A_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$ gives

$$f_{\mathcal{A}}(x) = \log\left(\frac{\det A_i}{-\alpha_{A_i}(x + \sigma_{A_i})}\right) \quad \text{for } x \in X_{A_i}, \tag{43}$$

where we recall from Corollary 5 that $-\det A_i / (\alpha_{A_i}(x + \sigma_{A_i})) > 0$ for all $x \in X_{A_i}$.

Remark 8. The function $f_{\mathcal{A}}$ is clearly Lipschitz continuous on each X_{A_i}, and hence is Lipschitz continuous on $X_{\mathcal{A}} = X_{A_0} \cup X_{A_1}$ since the intervals X_{A_0} and X_{A_1} are disjoint.

The reason for introducing the function $f_{\mathcal{A}}$ is provided by the following characterization of the joint spectral radius in terms of ergodic optimization.

THEOREM 4. *If $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then its joint spectral radius $r(\mathcal{A})$ satisfies*

$$\log r(\mathcal{A}) = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu. \tag{44}$$

Proof. From Proposition 5,

$$\log r(\mathcal{A}) = \sup_{\underline{i} \in \Omega^*} \log(\det A(\underline{i})(T_{\mathcal{A}}^{|\underline{i}|})'(p_{A(\underline{i})}))^{1/2|\underline{i}|}. \tag{45}$$

If $\underline{i} \in \{0, 1\}^n$, then

$$\begin{aligned}
\log(\det A(\underline{i})(T_{\mathcal{A}}^n)'(p_{A(\underline{i})}))^{1/2|\underline{i}|} &= \frac{1}{2n}(\log \det(A_{i_1} \cdots A_{i_n}) + \log(T_{\mathcal{A}}^n)'(p_{A(\underline{i})})) \\
&= \frac{1}{2n}\left(\log \prod_{j=1}^{n} \det A_{i_j} + \log \prod_{j=0}^{n-1} T_{\mathcal{A}}'(T_{\mathcal{A}}^j(p_{A(\underline{i})}))\right) \\
&= \frac{1}{n}\sum_{j=0}^{n-1} \frac{1}{2}(\log \det A_{i_{j+1}} + \log T_{\mathcal{A}}'(T_{\mathcal{A}}^j(p_{A(\underline{i})}))) \\
&= \frac{1}{n}\sum_{j=0}^{n-1} f_{\mathcal{A}}(T_{\mathcal{A}}^j(p_{A(\underline{i})})), \tag{46}
\end{aligned}$$

where the last step uses (41) together with the fact that

$$\log \det A_{i_{j+1}} = ((\log \det A_0)\mathbb{1}_{X_{A_0}} + (\log \det A_1)\mathbb{1}_{X_{A_1}})(T^j_{\mathcal{A}}(p_{A(\underline{i})})),$$

because

$$\mathbb{1}_{X_{A_k}}(T^j_{\mathcal{A}}(p_{A(\underline{i})})) = \begin{cases} 1 & \text{if } i_{j+1} = k, \\ 0 & \text{if } i_{j+1} \neq k. \end{cases}$$

Combining (45) and (46) gives

$$\log r(\mathcal{A}) = \sup_{\underline{i} \in \Omega^*} \frac{1}{|\underline{i}|} \sum_{j=0}^{|\underline{i}|-1} f_{\mathcal{A}}(T^j_{\mathcal{A}}(p_{A(\underline{i})})) = \sup_{\underline{i} \in \Omega^*} \int f_{\mathcal{A}} \, d\mu_{\underline{i}}, \tag{47}$$

where

$$\mu_{\underline{i}} = \frac{1}{|\underline{i}|} \sum_{j=0}^{|\underline{i}|-1} \delta_{T^j_{\mathcal{A}}(p_{A(\underline{i})})}$$

is the periodic orbit measure (see Definition 8) for the period-$|\underline{i}|$ point $p_{A(\underline{i})} = T^{|\underline{i}|}_{\mathcal{A}}(p_{A(\underline{i})})$: i.e. the unique measure in $\mathcal{M}_{\mathcal{A}}$ whose support equals the periodic orbit $\{T^j_{\mathcal{A}}(p_{A(\underline{i})})\}_{j=0}^{|\underline{i}|-1}$. By a result of Parthasarathy [18] (see also Sigmund [20]), the collection of periodic orbit measures $\{\mu_{\underline{i}} : \underline{i} \in \Omega^*\}$ is weak* dense in the weak* compact space $\mathcal{M}_{\mathcal{A}}$, so

$$\sup_{\underline{i} \in \Omega^*} \int f_{\mathcal{A}} \, d\mu_{\underline{i}} = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu, \tag{48}$$

and combining with (47) gives the required equality (44). □

4.3. *The finiteness property and periodic orbits.* In view of Theorem 4, we shall be interested in those measures $m \in \mathcal{M}_{\mathcal{A}}$ which are $f_{\mathcal{A}}$-maximizing, in the sense of Definition 9, i.e. m attains the maximum in (44): $\log r(\mathcal{A}) = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu = \int f_{\mathcal{A}} \, dm$. The finiteness property for $\mathcal{A}$ (which we recall means that $r(\mathcal{A}) = r(A_{i_1} \cdots A_{i_n})^{1/n}$ for some $i_1, \ldots, i_n \in \{0, 1\}$) corresponds to the existence of a periodic orbit measure which is $f_{\mathcal{A}}$-maximizing.

PROPOSITION 6. *$\mathcal{A} \in \mathfrak{C}$ has the finiteness property if and only if some $T_{\mathcal{A}}$-periodic orbit measure is $f_{\mathcal{A}}$-maximizing.*

Proof. If $\mathcal{A} \in \mathfrak{C}$ has the finiteness property and $\underline{i} \in \{0, 1\}^n$ satisfies $r(\mathcal{A}) = r(A(\underline{i}))^{1/n}$, then we claim that the corresponding periodic orbit measure $\mu_{\underline{i}} = (1/n) \sum_{j=0}^{n-1} \delta_{T^j_{\mathcal{A}}(p_{A(\underline{i})})}$ is $f_{\mathcal{A}}$-maximizing. To see this, first note that (44) gives

$$\log r(A(\underline{i}))^{1/n} = \log r(\mathcal{A}) = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu, \tag{49}$$

and the left-hand side of (49) can be written as

$$\log r(A(\underline{i}))^{1/n} = \log \left(\frac{\det A(\underline{i})}{T'_{A(\underline{i})}(p_{A(\underline{i})})} \right)^{1/2n} = \log(\det A(\underline{i})(T^n_{\mathcal{A}})'(p_{A(\underline{i})}))^{1/2n} \tag{50}$$

using Corollary 2 and (39). Therefore (46) gives

$$\log r(A(\underline{i}))^{1/n} = \frac{1}{n} \sum_{j=0}^{n-1} f_{\mathcal{A}}(T_{\mathcal{A}}^j(p_{A(\underline{i})})) = \int f_{\mathcal{A}} \, d\mu_{\underline{i}}, \tag{51}$$

so (49) implies that $\mu_{\underline{i}}$ is, indeed, $f_{\mathcal{A}}$-maximizing.

Conversely, if some $T_{\mathcal{A}}$-periodic orbit measure is $f_{\mathcal{A}}$-maximizing, then this measure is necessarily of the form $\mu_{\underline{i}} = (1/n) \sum_{j=0}^{n-1} \delta_{T_{\mathcal{A}}^j(p_{A(\underline{i})})}$ for some $n \in \mathbb{N}$ and $\underline{i} \in \{0, 1\}^n$, and it satisfies $\int f_{\mathcal{A}} \, d\mu_{\underline{i}} = \max_{\mu \in \mathcal{M}_{\mathcal{A}}} \int f_{\mathcal{A}} \, d\mu$. Combining (44) with (50) and (51), gives $\log r(A(\underline{i}))^{1/n} = \log r(\mathcal{A})$, so $\mathcal{A}$ has the finiteness property, as required. □

Note that the above proof has also established the following proposition.

PROPOSITION 7. *Suppose that $\mathcal{A} \in \mathfrak{C}$, and $\underline{i} \in \{0, 1\}^n$ for some $n \in \mathbb{N}$. Then $r(\mathcal{A}) = r(A(\underline{i}))^{1/n}$ if and only if the periodic orbit measure $\mu_{\underline{i}} = (1/n) \sum_{j=0}^{n-1} \delta_{T_{\mathcal{A}}^j(p_{A(\underline{i})})}$ is $f_{\mathcal{A}}$-maximizing.*

Recall that we say that $\mathcal{A} \in \mathfrak{C}$ is a *finiteness counterexample* if $r(\mathcal{A}) > r(A_{i_1} \cdots A_{i_n})^{1/n}$ for all $n \in \mathbb{N}$ and all choices $i_1, \ldots, i_n \in \{0, 1\}$.

PROPOSITION 8. *$\mathcal{A} \in \mathfrak{C}$ is a finiteness counterexample if and only if no $T_{\mathcal{A}}$-periodic orbit measure is $f_{\mathcal{A}}$-maximizing. In this case, there exists at least one measure $\mu \in \mathcal{M}_{\mathcal{A}}$ that is $f_{\mathcal{A}}$-maximizing, and there exist uncountably many sequences $\omega \in \Omega$ such that*

$$r(\mathcal{A}) = \lim_{n \to \infty} r(A_{\omega_1} \cdots A_{\omega_n})^{1/n}. \tag{52}$$

Proof. The first statement is equivalent to that of Proposition 6, while the existence of an $f_{\mathcal{A}}$-maximizing measure μ is a consequence (see, e.g., [**9**, Proposition 2.4(i)]) of the continuity of $f_{\mathcal{A}}$ and the weak* compactness of $\mathcal{M}_{\mathcal{A}}$ (see Lemma 15). In fact, μ may be chosen to be an ergodic measure since it is readily shown that the set of $f_{\mathcal{A}}$-maximizing measures is convex, and any of its extremal points is ergodic (see, e.g., [**9**, Proposition 2.4]). The ergodic theorem (see, e.g., [**22**, Theorem 1.14]) then implies that

$$\log r(\mathcal{A}) = \int f_{\mathcal{A}} \, d\mu = \lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f_{\mathcal{A}}(T_{\mathcal{A}}^j(p)) \tag{53}$$

for μ-almost every $p \in Y_{\mathcal{A}}$. Since no periodic orbit measure is $f_{\mathcal{A}}$-maximizing, the measure μ must have uncountable support, and therefore (53) holds for an uncountable set of points $p \in Y_{\mathcal{A}}$.

We may use the topological conjugacy $h_{\mathcal{A}} : \Omega \to Y_{\mathcal{A}}$ to define the image measure $m = (h_{\mathcal{A}}^{-1})^*(\mu)$, which also has uncountable support, and if we write $h_{\mathcal{A}}(\omega) = p$, then

$$r(A_{\omega_1} \cdots A_{\omega_n})^{1/n} = \exp\left(\frac{1}{n} \sum_{j=0}^{n-1} f_{\mathcal{A}}(T_{\mathcal{A}}^j(p))\right), \tag{54}$$

so (53) implies that $r(\mathcal{A}) = \lim_{n \to \infty} r(A_{\omega_1} \cdots A_{\omega_n})^{1/n}$ for m-almost every $\omega \in \Omega$, and hence for uncountably many $\omega \in \Omega$. □

4.4. *Monotonicity properties and formulae.* The following simple lemma records that, for $\mathcal{A} \in \mathfrak{C}$, the induced dynamical system $T_{\mathcal{A}(t)}$ is independent of t and that the induced function $f_{\mathcal{A}(t)}$ differs from $f_{\mathcal{A}}$ only by the addition of a scalar multiple of the characteristic function for the image X_{A_1}.

LEMMA 16. *For $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$ and all $t > 0$:*

(i) $T_{\mathcal{A}(t)} = T_{\mathcal{A}}$;

(ii) $f_{\mathcal{A}(t)} = f_{\mathcal{A}} + (\log t) \mathbb{1}_{X_{A_1}}$;

(iii) $f_{\mathcal{A}(t)}(T_{A_0}(1)) - f_{\mathcal{A}(t)}(T_{A_1}(0)) = f_{\mathcal{A}}(T_{A_0}(1)) - f_{\mathcal{A}}(T_{A_1}(0)) - \log t$; *and*

(iv) $f'_{\mathcal{A}(t)} = f'_{\mathcal{A}}$, *with*

$$f'_{\mathcal{A}}(x) = -(x + \sigma_{A_i})^{-1} \quad for \ x \in X_{A_i}, \ i \in \{0, 1\}. \tag{55}$$

Proof. (i) From Remark 3, we see that if $t > 0$, then $T_{tA_1} = T_{A_1}$, and hence $T_{\mathcal{A}(t)} = T_{\mathcal{A}}$.

(ii) Formula (42) gives $f_{\mathcal{A}} = f_{\mathcal{A}(t)}$ on X_{A_0}, while, for $x \in X_{A_1}$,

$$f_{\mathcal{A}(t)}(x) = \tfrac{1}{2}(\log S'_{A_1}(x) + \log \det t A_1) = \log t + f_{\mathcal{A}}(x)$$

since $\log \det t A_1 = \log(t^2 \det A_1) = 2 \log t + \log \det A_1$, and thus $f_{\mathcal{A}(t)} = f_{\mathcal{A}} + (\log t) \mathbb{1}_{X_{A_1}}$.

(iii) This is immediate from part (ii).

(iv) The formula for $f'_{\mathcal{A}}$ follows readily from the explicit formula (43) for $f_{\mathcal{A}}$ and is equal to $f'_{\mathcal{A}(t)}$ by (ii) above. $\square$

LEMMA 17. *If $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then:*

(i) $f'_{\mathcal{A}}$ *is strictly positive on X_{A_0} and strictly negative on X_{A_1};*

(ii) $f_{\mathcal{A}}$ *is strictly increasing on X_{A_0} and strictly decreasing on X_{A_1}; and*

(iii) $(f_{\mathcal{A}} \circ T^i_{A_0})'(x) > 0$ *and* $(f_{\mathcal{A}} \circ T^i_{A_1})'(x) < 0$ *for all $i \geq 1$, $x \in X$.*

Proof. (i) In view of formula (55), it suffices to note that, by Corollary 5, $x + \sigma_{A_0} < 0$ for $x \in X_{A_0}$ and $x + \sigma_{A_1} > 0$ for $x \in X_{A_1}$.

(ii) This is an immediate consequence of (i).

(iii) By the chain rule,

$$(f_{\mathcal{A}} \circ T^i_{A_j})'(x) = f'_{\mathcal{A}}(T^i_{A_j}(x))(T^i_{A_j})'(x). \tag{56}$$

The second factor $(T^i_{A_j})'(x)$ on the right-hand side of (56) is strictly positive for all $x \in X$, $i \geq 1$, $j \in \{0, 1\}$, since T_{A_j} is orientation preserving, as noted in Remark 4.

Regarding the sign of the first factor $f'_{\mathcal{A}}(T^i_{A_j}(x))$ on the right-hand side of (56), note that, since $i \geq 1$, $T^i_{A_j}(x) \in X_{A_j} = T_{A_j}(X)$ for all $x \in X$. Part (i) above then implies that $f'_{\mathcal{A}}(T^i_{A_j}(x))$ is strictly positive when $j = 0$ and strictly negative when $j = 1$. It follows that $(f_{\mathcal{A}} \circ T^i_{A_j})'(x)$ is strictly positive when $j = 0$ and strictly negative when $j = 1$, as required. $\square$

For the purposes of the following Lemma 18, it will be convenient to introduce the following notation.

Notation 6. For a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{M}$, define

$$\delta_A = \frac{b+d}{\alpha_A} = \frac{b+d}{a+c-b-d}.$$

We can now give another characterization of ϱ_A.

LEMMA 18. *For $A \in \mathbb{M}$,*

$$\varrho_A = \lim_{k \to \infty} \delta_{A^k}.$$

Proof. Perron–Frobenius theory (see, e.g., [22, Theorem 0.17]) gives

$$\lim_{k \to \infty} \lambda_A^{-k} A^k = v w_A = \begin{pmatrix} v^{(1)} w_A^{(1)} & v^{(1)} w_A^{(2)} \\ v^{(2)} w_A^{(1)} & v^{(2)} w_A^{(2)} \end{pmatrix},$$

where the positive dominant eigenvalue $\lambda_A > 0$ and corresponding left eigenvector w_A are as in Lemma 3, and v is a corresponding right eigenvector (a suitable multiple of v_A from Lemma 3), normalized so that $w_A v = 1$.

It follows that

$$\lim_{k \to \infty} \delta_{A^k} = \frac{v^{(1)} w_A^{(2)} + v^{(2)} w_A^{(2)}}{v^{(1)} w_A^{(1)} + v^{(2)} w_A^{(1)} - v^{(1)} w_A^{(2)} - v^{(2)} w_A^{(2)}} = \frac{w_A^{(2)}}{w_A^{(1)} - w_A^{(2)}}, \qquad (57)$$

so the formula $\varrho_A = w_A^{(2)} / (w_A^{(1)} - w_A^{(2)})$ from Lemma 8 concludes the proof. $\qquad \square$

COROLLARY 6. *For $A \in \mathbb{M}$, $x \in X$,*

$$\sum_{n=1}^{\infty} (\log S_A' \circ T_A^n)'(x) = \frac{2}{x + \varrho_A}. \qquad (58)$$

Proof. A simple calculation using the chain rule yields

$$\sum_{n=1}^{k} (\log S_A' \circ T_A^n)'(x) = -(\log T_{A^k}')'(x) = \frac{2}{x + \delta_{A^k}} \qquad (59)$$

for all $k \geq 1$, so, letting $k \to \infty$, we see that the result follows from Lemma 18. $\qquad \square$

Recalling from (42) that $f_{\mathcal{A}} = \frac{1}{2}(\log S_{A_i}' + \log \det A_i)$ on X_{A_i}, the following result is an immediate consequence of Corollary 6.

COROLLARY 7. *If $\mathcal{A} = (A_0, A_1) \in \mathbb{M}^2$, then, for $i \in \{0, 1\}$,*

$$\sum_{n=1}^{\infty} (f_{\mathcal{A}} \circ T_{A_i}^n)'(x) = \frac{1}{x + \varrho_{A_i}} \quad \text{for } x \in X_{A_i}. \qquad (60)$$

COROLLARY 8. *If $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$, then, for all $x \in X$,*

$$\sum_{n=1}^{\infty} (f_{\mathcal{A}} \circ T_{A_0}^n)'(x) > 0 \quad \text{and} \quad \sum_{n=1}^{\infty} (f_{\mathcal{A}} \circ T_{A_1}^n)'(x) < 0, \qquad (61)$$

and

$$x + \varrho_{A_0} > 0 \quad \text{and} \quad x + \varrho_{A_1} < 0. \qquad (62)$$

Proof. The inequalities in (61) follow from Lemma 17(iii), while (62) is an immediate consequence of (60) and (61). □

Remark 9. The inequality $x + \varrho_{A_0} > 0$ in (62) can also be deduced from the fact that $\varrho_{A_0} > 0$ (by Corollary 4) and $x \geq 0$.

5. *Sturmian measures associated to a concave–convex matrix pair*

For $\mathcal{A} \in \mathfrak{C}$, the induced space $X_{\mathcal{A}}$ becomes an ordered set when equipped with the usual order on $X = [0, 1]$. In particular, by a *sub-interval* of $X_{\mathcal{A}}$ we mean any subset of $X_{\mathcal{A}}$ of the form $I \cap X_{\mathcal{A}}$, where I is some sub-interval of X. Note that a sub-interval of $X_{\mathcal{A}}$ is a sub-interval of X if it is contained in either X_{A_0} or X_{A_1}; otherwise, it is a union of two disjoint intervals in X.

Definition 11. Given a matrix pair $\mathcal{A} \in \mathfrak{C}$, a closed interval $\Gamma \subset X_{\mathcal{A}}$ is called $\mathcal{A}$-*Sturmian* (or simply *Sturmian*) if $T_{\mathcal{A}}(\min \Gamma) = T_{\mathcal{A}}(\max \Gamma)$: i.e. its two endpoints $\min \Gamma$ and $\max \Gamma$ have the same image under the induced dynamical system $T_{\mathcal{A}}$.

Remark 10.
(a) The topological conjugacy $h_{\mathcal{A}} : \Omega \to Y_{\mathcal{A}}$ (cf. Remark 7) is order preserving, so if $\Gamma \subset X_{\mathcal{A}}$ is an $\mathcal{A}$-Sturmian interval, then $h_{\mathcal{A}}^{-1}(\Gamma \cap Y_{\mathcal{A}})$ is a Sturmian interval as defined in Notation 2 (i.e. of the form $[0\omega, 1\omega]$ for some $\omega \in \Omega$).
(b) For all $t > 0$, an interval is $\mathcal{A}$-Sturmian if and only if it is $\mathcal{A}(t)$-Sturmian.

Definition 12. Let $\mathcal{I}_{\mathcal{A}}$ denote the collection of all $\mathcal{A}$-Sturmian intervals. Note that $\mathcal{I}_{\mathcal{A}}$ is naturally parametrized by $X = [0, 1]$: for each $c \in X$, there is a unique $\Gamma \in \mathcal{I}_{\mathcal{A}}$ such that $T_{\mathcal{A}}(\min \Gamma) = T_{\mathcal{A}}(\max \Gamma) = c$. Henceforth, we shall write $c_{\mathcal{A}}(\Gamma)$ to denote the common value $T_{\mathcal{A}}(\min \Gamma) = T_{\mathcal{A}}(\max \Gamma)$ for an $\mathcal{A}$-Sturmian interval $\Gamma \in \mathcal{I}_{\mathcal{A}}$, noting that

$$c_{\mathcal{A}} : \mathcal{I}_{\mathcal{A}} \to X$$

is a bijection. As a subset of X, we can express $\Gamma \in \mathcal{I}_{\mathcal{A}}$ as

$$\Gamma = [T_{A_0}(c_{\mathcal{A}}(\Gamma)), T_{A_0}(1)] \cup [T_{A_1}(0), T_{A_1}(c_{\mathcal{A}}(\Gamma))]. \tag{63}$$

Remark 11. It is apparent from (63) that, viewed as a subset of $X = [0, 1]$, an $\mathcal{A}$-Sturmian interval Γ is *always* a disjoint union of two closed intervals. Note, however, that for the two extremal cases where $c_{\mathcal{A}}(\Gamma) = 0$ or 1, one of the intervals in the disjoint union is a singleton set (and the other interval is, respectively, either X_{A_0} or X_{A_1}). These extremal cases are particularly significant, and in the calculations of §7 onwards, it is convenient to neglect the singleton set, thereby identifying the extremal $\mathcal{A}$-Sturmian interval with either X_{A_0} or X_{A_1}.

Definition 13. We say that a $T_{\mathcal{A}}$-invariant Borel probability measure on $X_{\mathcal{A}}$ is $\mathcal{A}$-*Sturmian* if its support is contained in some $\mathcal{A}$-Sturmian interval. Let $\mathcal{S}_{\mathcal{A}}$ denote the collection of $\mathcal{A}$-Sturmian measures.

Remark 12.

(a) In view of Remarks 7 and 10, the class of $\mathcal{A}$-Sturmian measures on $X_{\mathcal{A}}$ is just the $h_{\mathcal{A}}^*$-image of the class of Sturmian measures on the shift space Ω: i.e. $\mathcal{S}_{\mathcal{A}} = h_{\mathcal{A}}^*(\mathcal{S})$. In particular (cf. Proposition 1(b)), $\mathcal{S}_{\mathcal{A}}$ is also naturally parametrized by $X = [0, 1]$: the map $\mathcal{P} \circ (h_{\mathcal{A}}^*)^{-1} : \mathcal{S}_{\mathcal{A}} \to [0, 1]$ is a homeomorphism and, for $\mu \in \mathcal{S}_{\mathcal{A}}$, we refer to $\mathcal{P} \circ (h_{\mathcal{A}}^*)^{-1}(\mu) = \mu(X_{A_1})$ as its *(Sturmian) parameter*.

(b) For all $t > 0$, a measure is $\mathcal{A}$-Sturmian if and only if it is $\mathcal{A}(t)$-Sturmian.

In §10, we shall identify cases where $\mathcal{A}$-Sturmian measures arise as unique maximizing measures for $f_{\mathcal{A}(t)}, t > 0$. In particular, for certain t, the unique $f_{\mathcal{A}(t)}$-maximizing measure is a Sturmian measure of *irrational* parameter, and such $\mathcal{A}(t)$ turn out to be *finiteness counterexamples*.

PROPOSITION 9. *If $\mathcal{A} \in \mathfrak{C}$ is such that there is a unique $f_{\mathcal{A}}$-maximizing measure, and if this measure is an $\mathcal{A}$-Sturmian measure with irrational parameter $\mathcal{P}$, then $\mathcal{A}$ is a finiteness counterexample (i.e. $r(\mathcal{A}) > r(A_{i_1} \cdots A_{i_n})^{1/n}$ for all $n \in \mathbb{N}$ and all choices $i_1, \ldots, i_n \in \{0, 1\}$). In this case,*

$$r(\mathcal{A}) = \lim_{n \to \infty} r(A_{\omega_1} \cdots A_{\omega_n})^{1/n} \tag{64}$$

holds for the uncountably many Sturmian sequences $\omega = (\omega_n)_{n=1}^{\infty}$ of parameter $\mathcal{P}$.

Proof. By assumption, there is a unique $f_{\mathcal{A}}$-maximizing measure μ, and this measure is an $\mathcal{A}$-Sturmian measure with irrational parameter $\mathcal{P}$, which, in particular, is not a periodic orbit measure. It follows that no $T_{\mathcal{A}}$-periodic orbit measure is $f_{\mathcal{A}}$-maximizing, so Proposition 8 implies that $\mathcal{A}$ is a finiteness counterexample and that there exist uncountably many sequences $\omega \in \Omega$ such that (64) holds. In fact, the support of any $\mathcal{A}$-Sturmian measure μ is uniquely ergodic (see, e.g., [**4**, Corollary 1.6]), so the ergodic theorem holds for the uncountably many points in the support of μ (see, e.g., [**22**, Theorem 6.19]), and therefore $\lim_{n \to \infty} (1/n) \sum_{j=0}^{n-1} f_{\mathcal{A}}(T_{\mathcal{A}}^j(p)) = \int f_{\mathcal{A}} \, d\mu = \log r(\mathcal{A})$ for all p in the support of μ. Writing $p = h_{\mathcal{A}}(\omega)$, as in the proof of Proposition 8, the relationship (54) then implies that (64) holds for all points in the support of the Sturmian measure m, i.e. for all Sturmian sequences of parameter $\mathcal{P}$. □

6. *The Sturmian transfer function*

In order to show that the maximizing measure for $f_{\mathcal{A}(t)}$ is supported in some $\mathcal{A}$-Sturmian interval $\Gamma \in \mathcal{I}_{\mathcal{A}}$, our strategy will be to add a coboundary $\varphi_{\Gamma} - \varphi_{\Gamma} \circ T_{\mathcal{A}}$, where the corresponding *Sturmian transfer function* φ_{Γ} is introduced below, so that the new function $f_{\mathcal{A}} + \varphi_{\Gamma} - \varphi_{\Gamma} \circ T_{\mathcal{A}}$ takes a constant value on all of Γ and is strictly smaller than this constant value on the complement of Γ. This approach is patterned on ideas of Bousch [**3**] in the setting of the angle-doubling map and degree-one trigonometric polynomials.

To proceed, it is convenient to introduce the following definition.

Definition 14. For $\mathcal{A} \in \mathfrak{C}$, to each $\mathcal{A}$-Sturmian interval Γ we associate the *hybrid contraction* $\tau_{\Gamma} : X \to X_{\mathcal{A}}$, defined by

$$\tau_{\Gamma}(x) = \begin{cases} T_{A_1}(x) & \text{if } x \in [0, c(\Gamma)), \\ T_{A_0}(x) & \text{if } x \in [c(\Gamma), 1]. \end{cases} \tag{65}$$

Remark 13. The hybrid contraction τ_Γ satisfies $\tau_\Gamma(X) = \Gamma$ and is piecewise Lipschitz continuous. More precisely, its restriction to $[0, c_{\mathcal{A}}(\Gamma))$ is Lipschitz, as is its restriction to $[c_{\mathcal{A}}(\Gamma), 1]$.

LEMMA 19. *Given $\mathcal{A} \in \mathfrak{C}$ and an $\mathcal{A}$-Sturmian interval $\Gamma \in \mathcal{I}_{\mathcal{A}}$, there exists a unique Lipschitz continuous function $\varphi_{\mathcal{A},\Gamma} : X \to \mathbb{R}$ which simultaneously satisfies†*

$$\varphi'_{\mathcal{A},\Gamma} = \sum_{n=1}^{\infty} (f_{\mathcal{A}} \circ \tau_\Gamma^n)' \quad \text{Lebesgue almost everywhere,} \tag{66}$$

and

$$\varphi_{\mathcal{A},\Gamma}(0) = 0. \tag{67}$$

Proof. The function $f_{\mathcal{A}}$ is Lipschitz and τ_Γ is piecewise Lipschitz (cf. Remark 13), so each τ_Γ^n is piecewise Lipschitz, and thus by Rademacher's Theorem is differentiable Lebesgue almost everywhere, with L^∞ derivative. Now $\|(\tau_\Gamma^n)'\|_\infty = O(\theta^n)$ as $n \to \infty$ for some $\theta \in (0, 1)$, so the sum

$$\sum_{n=1}^{\infty} (f_{\mathcal{A}} \circ \tau_\Gamma^n)' = \sum_{n=1}^{\infty} f'_{\mathcal{A}} \circ \tau_\Gamma^n . (\tau_\Gamma^n)'$$

is Lebesgue almost everywhere convergent (as its nth term is $O(\theta^n)$), and it defines an L^∞ function with respect to Lebesgue measure on X. In particular, it has a Lipschitz antiderivative φ_Γ, which is the unique Lipschitz antiderivative up to an additive constant, and hence is uniquely defined if it satisfies the additional condition $\varphi_{\mathcal{A},\Gamma}(0) = 0$. □

Notation 7. For $\mathcal{A} \in \mathfrak{C}$, $\Gamma \in \mathcal{I}_{\mathcal{A}}$, the function $\varphi_\Gamma = \varphi_{\mathcal{A},\Gamma}$, whose existence and uniqueness is guaranteed by Lemma 19, will be referred to as the corresponding *Sturmian transfer function*.

Remark 14. Note that, although the induced function $f_{\mathcal{A}}$ is only defined on $X_{\mathcal{A}}$, the Sturmian transfer function φ_Γ is actually defined on all of $X = [0, 1]$. For the most part, however, we shall only be interested in the restriction of φ_Γ to $X_{\mathcal{A}}$. More precisely, we shall be interested in certain properties of $f_{\mathcal{A}} + \varphi_\Gamma$ or of $f_{\mathcal{A}} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$, considered as functions defined on $X_{\mathcal{A}}$, beginning with the following Corollary 9.

COROLLARY 9. *If $\mathcal{A} \in \mathfrak{C}$, and Γ is any $\mathcal{A}$-Sturmian interval, then both $f_{\mathcal{A}} + \varphi_\Gamma$ and $f_{\mathcal{A}} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ are Lipschitz continuous functions on $X_{\mathcal{A}}$.*

Proof. Both $T_{\mathcal{A}}$ and $f_{\mathcal{A}}$ are Lipschitz continuous on $X_{\mathcal{A}}$, as noted in Remarks 6 and 8, and φ_Γ is Lipschitz continuous on X, as noted in Lemma 19, and hence Lipschitz continuous on $X_{\mathcal{A}}$. It follows that both $f_{\mathcal{A}} + \varphi_\Gamma$ and $f_{\mathcal{A}} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ are Lipschitz continuous on $X_{\mathcal{A}}$. □

LEMMA 20. *Suppose that $\mathcal{A} \in \mathfrak{C}$, $t > 0$, and that Γ is any $\mathcal{A}$-Sturmian interval. The Lipschitz continuous function $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}} : X_{\mathcal{A}} \to \mathbb{R}$ has the property that its restriction to $\Gamma \cap X_{A_0}$ is a constant function and its restriction to $\Gamma \cap X_{A_1}$ is a constant function.*

---

† The substantial condition is (66), which determines $\varphi_{\mathcal{A},\Gamma}$ up to an additive constant. The extra condition (67) is useful in that it removes any ambiguity when discussing $\varphi_{\mathcal{A},\Gamma}$.

Proof. By Corollary 9, the function $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ is Lipschitz continuous on $X_{\mathcal{A}}$, because $\mathcal{A}(t) \in \mathfrak{C}$. So, by the fundamental theorem of calculus for Lipschitz functions (see, e.g., [**11**, Theorem 7.1.15]), the required result will follow if it can be shown that

$$(f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}})' = 0 \quad \text{Lebesgue almost everywhere on } \Gamma. \tag{68}$$

But $f'_{\mathcal{A}(t)} = f'_{\mathcal{A}}$, so (68) is equivalent to proving that

$$(f_{\mathcal{A}} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}})' = 0 \quad \text{Lebesgue almost everywhere on } \Gamma. \tag{69}$$

To establish this almost everywhere equality, note that

$$f'_{\mathcal{A}} + \varphi'_\Gamma = f'_{\mathcal{A}} + \sum_{n=1}^\infty (f_{\mathcal{A}} \circ \tau_\Gamma^n)' = \sum_{n=0}^\infty (f_{\mathcal{A}} \circ \tau_\Gamma^n)' = \sum_{n=0}^\infty f'_{\mathcal{A}} \circ \tau_\Gamma^n .(\tau_\Gamma^n)'$$

and

$$(\varphi_\Gamma \circ T_{\mathcal{A}})' = \sum_{n=1}^\infty f'_{\mathcal{A}} \circ \tau_\Gamma^n \circ T_{\mathcal{A}} .(\tau_\Gamma^n)' \circ T_{\mathcal{A}} . T'_{\mathcal{A}} = \sum_{n=1}^\infty f'_{\mathcal{A}} \circ \tau_\Gamma^{n-1} .(\tau_\Gamma^{n-1})'$$

since $(\tau_\Gamma \circ T_{\mathcal{A}})' = \tau_\Gamma^{n-1}$, so, indeed, (69) holds. □

Remark 15. In the generality of Lemma 20, the constant values assumed by $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ on $\Gamma \cap X_{A_0}$ and $\Gamma \cap X_{A_1}$ do not coincide. However, we shall shortly give (see Lemma 23) an extra condition which does ensure that $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ takes the *same* constant value on the whole of Γ. Indeed, this possibility is a key tool in our strategy.

7. *The extremal Sturmian intervals*

7.1. *Formulae involving extremal intervals.* As noted in Remark 11, an $\mathcal{A}$-Sturmian interval is the disjoint union of two closed intervals when viewed as a subset of $X = [0, 1]$. However, the two *extremal* cases yield a leftmost $\mathcal{A}$-Sturmian interval equal to $X_{A_0} \cup \{T_{A_1}(0)\}$ and a rightmost $\mathcal{A}$-Sturmian interval equal to $\{T_{A_0}(1)\} \cup X_{A_1}$. The presence of singleton sets in these expressions is notationally inconvenient, and unnecessary for our purposes, so henceforth we neglect them.

More precisely, henceforth, the leftmost $\mathcal{A}$-Sturmian interval is taken to be $X_{A_0} = T_{A_0}(X)$, and denoted by Γ_0, so that $\tau_{\Gamma_0} = T_{A_0}$; the rightmost $\mathcal{A}$-Sturmian interval is taken to be $X_{A_1} = T_{A_1}(X)$, and denoted by Γ_1, so that $\tau_{\Gamma_1} = T_{A_1}$.

When the $\mathcal{A}$-Sturmian interval Γ is either Γ_0 or Γ_1, there is an explicit formula for the Sturmian transfer function φ_Γ.

LEMMA 21. *Suppose that $\mathcal{A} \in \mathfrak{C}$. For $i \in \{0, 1\}$ and all $x \in X$,*

$$\varphi_{\Gamma_i}(x) = \log\left(\frac{x + \varrho_{A_i}}{\varrho_{A_i}}\right). \tag{70}$$

Proof. Now $\tau_{\Gamma_i} = T_{A_i}$, so the defining formula (66) becomes

$$\varphi'_{\Gamma_i}(x) = \sum_{n=1}^\infty (f_{\mathcal{A}} \circ T_{A_i}^n)'(x), \tag{71}$$

and then (60) implies that

$$\varphi'_{\Gamma_i}(x) = \frac{1}{x + \varrho_{A_i}}.$$

Noting that the sign of $x + \varrho_{A_i}$ is positive when $i = 0$ and negative when $i = 1$ (see Corollary 8), as well as the convention that $\varphi_{\Gamma_i}(0) = 0$ (see Definition 19), we deduce the required expression (70). □

Definition 15. Given $\mathcal{A} = (A_0, A_1) \in \mathfrak{C}$ and $\Gamma \in \mathcal{I}_\mathcal{A}$, define $\Delta_\mathcal{A}(\Gamma) \in \mathbb{R}$ by

$$\Delta_\mathcal{A}(\Gamma) = (\varphi_\Gamma(1) - \varphi_\Gamma(0)) - (\varphi_\Gamma(T_{A_0}(1)) - \varphi_\Gamma(T_{A_1}(0))), \tag{72}$$

noting the equivalent expression

$$\Delta_\mathcal{A}(\Gamma) = \varphi_\Gamma(1) - (\varphi_\Gamma(T_{A_0}(1)) - \varphi_\Gamma(T_{A_1}(0))) \tag{73}$$

as a consequence of the convention that $\varphi_\Gamma(0) = 0$ (see Lemma 19).

The values $\Delta_\mathcal{A}(\Gamma_i)$ play an important role, so it will be useful to record the following explicit formulae.

LEMMA 22. *Suppose that $\mathcal{A} \in \mathfrak{C}$. For $i \in \{0, 1\}$,*

$$\Delta_\mathcal{A}(\Gamma_i) = \log\left(\frac{(1 + \varrho_{A_i})(b_1/(b_1 + d_1) + \varrho_{A_i})}{\varrho_{A_i}(a_0/(a_0 + c_0) + \varrho_{A_i})}\right). \tag{74}$$

Proof. This is immediate from the defining formula (72) (or (73)) for $\Delta_\mathcal{A}(\Gamma_i)$, together with formula (70) for φ_{Γ_i}, and the fact that $T_{A_0}(1) = a_0/(a_0 + c_0)$ and $T_{A_1}(0) = b_1/(b_1 + d_1)$. □

COROLLARY 10. *If $\mathcal{A} \in \mathfrak{C}$, then*

$$\Delta_\mathcal{A}(\Gamma_1) < 0 < \Delta_\mathcal{A}(\Gamma_0).$$

Proof. The four terms ϱ_{A_i}, $1 + \varrho_{A_i}$, $a_0/(a_0 + c_0) + \varrho_{A_i}$ and $b_1/(b_1 + d_1) + \varrho_{A_i}$ in (74) are all positive if $i = 0$ and all negative if $i = 1$, by the inequalities (62) in Corollary 8. Now $\mathcal{A} \in \mathfrak{C}$ implies that (6) holds, so $b_1/(b_1 + d_1) + \varrho_{A_i} > a_0/(a_0 + c_0) + \varrho_{A_i}$ and, clearly, $1 + \varrho_{A_i} > \varrho_{A_i}$ Consequently,

$$\frac{(1 + \varrho_{A_i})(b_1/(b_1 + d_1) + \varrho_{A_i})}{\varrho_{A_i}(a_0/(a_0 + c_0) + \varrho_{A_i})}$$

is strictly greater than one if $i = 0$, and strictly smaller than one if $i = 1$. The result then follows from Lemma 22. □

7.2. *Adaptations for other matrix pairs.* As mentioned in §1.3, the methods of this paper can be adapted so as to give alternative proofs of certain results (analogues of Theorem 3) mentioned in §1, namely, establishing that a full Sturmian family is generated by the matrix pair (5) and for matrix pairs corresponding to sub-cases of (3) and (4) which lie in the boundary of $\mathfrak{D}$†. In this subsection, we indicate the modifications necessary to handle these cases.

† Note that all of the matrix pairs in (3), (4) and (5) have the property that A_0 is projectively concave and A_1 is projectively convex.

Firstly, the induced space $X_{\mathcal{A}}$ may be the whole of $X = [0, 1]$ rather than a disjoint union of two closed intervals: this occurs if $a_0/c_0 = b_1/d_1$ (i.e. when (6) becomes an equality), which is the case for the pair (5) and for (4) if $bc = 1$.

Secondly, in each of the cases (3), (4) and (5), the induced maps T_{A_0} and T_{A_1} have fixed points at zero and one, respectively, so that the dynamical system $T_{\mathcal{A}}$ also fixes these points. For (5), both zero and one are indifferent fixed points: i.e. $T'_{A_0}(0) = 1 = T'_{A_1}(1)$. For (3) and (4), these fixed points are unstable for the induced maps T_{A_0} and T_{A_1}, i.e. $T'_{A_0}(0) > 1$ and $T'_{A_1}(1) > 1$, but both of these maps also have stable fixed points in the interior of $X = [0, 1]$. Consequently, for (5), the dynamical system $T_{\mathcal{A}} : X \to X$ has indifferent fixed points at zero and one, and no other fixed points, while for (3) and (4), the dynamical system $T_{\mathcal{A}}$ has stable fixed points at zero and one and two further unstable fixed points in the interior of X.

The potentially problematic stable fixed points for $T_{\mathcal{A}}$ can, in fact, be avoided by omitting to consider the two extremal $\mathcal{A}$-Sturmian intervals: this ensures the asymptotic $\|(\tau_\Gamma^n)'\|_\infty = O(\theta^n)$ as $n \to \infty$, $\theta \in (0, 1)$, and the existence of Sturmian transfer functions is proved as in Lemma 19. In the case where $T_{\mathcal{A}}$ has indifferent fixed points, it is even possible to consider extremal $\mathcal{A}$-Sturmian intervals, as the series defining the Sturmian transfer function is, nonetheless, convergent. The existence of Sturmian transfer functions then allows the remainder of the method of proof to proceed essentially as for matrix pairs in $\mathfrak{D}$, ultimately establishing analogues of the main result, Theorem 3.

8. *Associating $\mathcal{A}$-Sturmian intervals to parameter values*

Notation 8. For a Sturmian interval $\Gamma \in \mathcal{I}_{\mathcal{A}}$, let $s_\Gamma \in \mathcal{S}_{\mathcal{A}}$ denote the $\mathcal{A}$-Sturmian measure supported by Γ: i.e. s_Γ is the unique $T_{\mathcal{A}}$-invariant probability measure whose support is contained in Γ.

LEMMA 23. *Suppose that $\mathcal{A} \in \mathfrak{C}$. If $t \in \mathbb{R}^+$ and $\Gamma \in \mathcal{I}_{\mathcal{A}}$ are such that*

$$f_{\mathcal{A}(t)}(T_{A_0}(1)) - f_{\mathcal{A}(t)}(T_{A_1}(0)) = \Delta_{\mathcal{A}}(\Gamma), \tag{75}$$

then the Lipschitz continuous function $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ is equal to the constant value $\int f_{\mathcal{A}(t)} \, ds_\Gamma$ when restricted to Γ.

Proof. By Lemma 20, we know that $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ is constant when restricted to $\Gamma \cap X_{A_0}$, and it is also constant when restricted to $\Gamma \cap X_{A_1}$. To prove that these constant values are the *same*, it suffices to show that $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}}$ takes the same value at the point $T_{A_0}(1) \in X_{A_0}$ as it does at the point $T_{A_1}(0) \in X_{A_1}$. But the equality

$$(f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}})(T_{A_0}(1)) = (f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_{\mathcal{A}})(T_{A_1}(0))$$

holds if and only if

$$f_{\mathcal{A}(t)}(T_{A_0}(1)) - f_{\mathcal{A}(t)}(T_{A_1}(0)) = \varphi_\Gamma(1) - \varphi_\Gamma(0) - (\varphi_\Gamma(T_{A_0}(1)) - \varphi_\Gamma(T_{A_1}(0))) :$$

in other words, $f_{\mathcal{A}(t)}(T_{A_0}(1)) - f_{\mathcal{A}(t)}(T_{A_1}(0)) = \Delta_{\mathcal{A}}(\Gamma)$, which is precisely the hypothesis (75). □

COROLLARY 11. *Given $\mathcal{A} \in \mathfrak{C}$, if $t \in \mathbb{R}^+$ and $\Gamma \in \mathcal{I}_\mathcal{A}$ are such that*

$$\log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right)t^{-1}\right) = \Delta_\mathcal{A}(\Gamma), \tag{76}$$

then the Lipschitz continuous function $f_{\mathcal{A}(t)} + \varphi_\Gamma - \varphi_\Gamma \circ T_\mathcal{A}$ is equal to the constant value $\int f_{\mathcal{A}(t)}\, ds_\Gamma$ on Γ.

Proof. By Lemma 23, it suffices to show that

$$f_{\mathcal{A}(t)}(T_{A_0}(1)) - f_{\mathcal{A}(t)}(T_{A_1}(0)) = \log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right)t^{-1}\right),$$

and, by Lemma 16(iii), this is equivalent to showing that

$$f_\mathcal{A}(T_{A_0}(1)) - f_\mathcal{A}(T_{A_1}(0)) = \log\left(\frac{a_0 + c_0}{b_1 + d_1}\right).$$

Substituting $T_{A_0}(1) = a_0/(a_0 + c_0)$ and $T_{A_1}(0) = b_1/(b_1 + d_1)$ into, respectively, the formulae (43) for $f_\mathcal{A}$ on X_{A_0} and X_{A_1}, yields

$$f_\mathcal{A}(T_{A_0}(1)) = \log(a_0 + c_0) \tag{77}$$

and

$$f_\mathcal{A}(T_{A_1}(0)) = \log(b_1 + d_1), \tag{78}$$

so the result follows. □

In view of equation (76), we make the following definition.

Definition 16. For $\mathcal{A} \in \mathfrak{C}$ and $i \in \{0, 1\}$, define $t_i = t_i(\mathcal{A})$ by

$$t_i = t_i(\mathcal{A}) = \left(\frac{a_0 + c_0}{b_1 + d_1}\right)e^{-\Delta_\mathcal{A}(\Gamma_i)}, \tag{79}$$

so that

$$\log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right)t_i^{-1}\right) = \Delta_\mathcal{A}(\Gamma_i).$$

Remark 16. Since $e^{-\Delta_\mathcal{A}(\Gamma_0)} < 1 < e^{-\Delta_\mathcal{A}(\Gamma_1)}$ by (10), it follows that

$$t_0(\mathcal{A}) < t_1(\mathcal{A}).$$

LEMMA 24. *For $\mathcal{A} \in \mathfrak{C}$ and $i \in \{0, 1\}$,*

$$t_i(\mathcal{A}) = \frac{\varrho_{A_i}(a_0 + \varrho_{A_i}(a_0 + c_0))}{(1 + \varrho_{A_i})(b_1 + \varrho_{A_i}(b_1 + d_1))}. \tag{80}$$

Proof. From (74), we see that, for $i \in \{0, 1\}$,

$$e^{-\Delta_\mathcal{A}(\Gamma_i)} = \frac{\varrho_{A_i}(a_0/(a_0 + c_0) + \varrho_{A_i})}{(1 + \varrho_{A_i})(b_1/(b_1 + d_1) + \varrho_{A_i})},$$

so that (79) gives

$$t_i(\mathcal{A}) = \left(\frac{a_0 + c_0}{b_1 + d_1}\right)e^{-\Delta_\mathcal{A}(\Gamma_i)} = \frac{\varrho_{A_i}(a_0 + \varrho_{A_i}(a_0 + c_0))}{(1 + \varrho_{A_i})(b_1 + \varrho_{A_i}(b_1 + d_1))}, \tag{81}$$

which is the required expression (80). □

A consequence is the following property.

COROLLARY 12. *For $\mathcal{A} \in \mathfrak{C}$, $t \in \mathbb{R}^+$, and $i \in \{0, 1\}$,*

$$t_i(\mathcal{A}(t)) = \frac{t_i(\mathcal{A})}{t}. \tag{82}$$

Proof. This follows easily from (80), and the easily verified fact (used only in the proof of the $i = 1$ case) that $\varrho_{tA_1} = \varrho_{A_1}$. Specifically, for $i \in \{0, 1\}$,

$$t_i(\mathcal{A}(t)) = \frac{\varrho_{A_i}(a_0 + \varrho_{A_i}(a_0 + c_0))}{(1 + \varrho_{A_i})(tb_1 + \varrho_{A_i}(tb_1 + td_1))} = \frac{1}{t} \frac{\varrho_{A_i}(a_0 + \varrho_{A_i}(a_0 + c_0))}{(1 + \varrho_{A_i})(b_1 + \varrho_{A_i}(b_1 + d_1))} = \frac{t_i(\mathcal{A})}{t}.$$

$\square$

LEMMA 25. *For $\mathcal{A} \in \mathfrak{C}$, the quantities $t_0(\mathcal{A})$ and $t_1(\mathcal{A})$ admit the alternative expressions*

$$t_0(\mathcal{A}) = \frac{\det A_0}{(a_0 - b_0(1 + \varrho_{A_0}^{-1}))(d_1 + b_1(1 + \varrho_{A_0}^{-1}))} \tag{83}$$

and

$$t_1(\mathcal{A}) = \frac{(a_0 + c_0(1 + \varrho_{A_1}^{-1})^{-1})(a_1 - b_1(1 + \varrho_{A_1}^{-1}))}{\det A_1}. \tag{84}$$

Proof. Since (81) implies that

$$t_0(\mathcal{A}) = \frac{a_0 + \varrho_{A_0}(a_0 + c_0)}{(1 + \varrho_{A_0})(d_1 + b_1(1 + \varrho_{A_0}^{-1}))}, \tag{85}$$

we see that $t_0(\mathcal{A})$ is equal to (83) if and only if

$$\frac{a_0 + \varrho_{A_0}(a_0 + c_0)}{1 + \varrho_{A_0}} = \frac{a_0 d_0 - b_0 c_0}{a_0 - b_0(1 + \varrho_{A_0}^{-1})}. \tag{86}$$

Clearing fractions in (86) reveals it to be equivalent to the equation

$$q_{A_0}(\varrho_{A_0}) = \alpha_{A_0}\varrho_{A_0}^2 + \beta_{A_0}\varrho_{A_0} - b_0 = 0,$$

which is true by Lemma 6.

Since (81) implies that

$$t_1(\mathcal{A}) = \frac{\varrho_{A_1}(a_0 + c_0(1 + \varrho_{A_1}^{-1})^{-1})}{b_1 + \varrho_{A_1}(b_1 + d_1)}, \tag{87}$$

we see that $t_1(\mathcal{A})$ is equal to (84) if and only if

$$\frac{\varrho_{A_1}}{b_1 + \varrho_{A_1}(b_1 + d_1)} = \frac{a_1 - b_1(1 + \varrho_{A_1}^{-1})}{\det A_1}. \tag{88}$$

Clearing fractions in (88) reveals it to be equivalent to the equation

$$q_{A_1}(\varrho_{A_1}) = \alpha_{A_1}\varrho_{A_1}^2 + \beta_{A_1}\varrho_{A_1} - b_1 = 0,$$

which is true by Lemma 6.

$\square$

Notation 9. For $\mathcal{A} \in \mathfrak{C}$, let $\mathcal{T}_{\mathcal{A}}$ denote the open interval $(t_0(\mathcal{A}), t_1(\mathcal{A}))$.

PROPOSITION 10. *Let $\mathcal{A} \in \mathfrak{C}$. For each $t \in \mathcal{T}_\mathcal{A}$, there exists an $\mathcal{A}$-Sturmian interval $\Gamma_\mathcal{A}(t) \in \mathcal{I}_\mathcal{A}$ such that $f_{\mathcal{A}(t)} + \varphi_{\Gamma_t} - \varphi_{\Gamma_t} \circ T_\mathcal{A}$ is equal to the constant value $\int f_{\mathcal{A}(t)} \, ds_{\Gamma_\mathcal{A}(t)}$ on $\Gamma_\mathcal{A}(t)$.*

Proof. First, we show that $\Delta_\mathcal{A} : \Gamma \mapsto \Delta_\mathcal{A}(\Gamma)$ is continuous. The formula (73) defines

$$\Delta_\mathcal{A}(\Gamma) = \varphi_\Gamma(1) - \varphi_\Gamma(T_{A_0}(1)) + \varphi_\Gamma(T_{A_1}(0))$$
$$= \varphi_\Gamma(1) - \varphi_\Gamma\left(\frac{a_0}{a_0 + c_0}\right) + \varphi_\Gamma\left(\frac{b_1}{b_1 + d_1}\right),$$

so the continuity of $\Delta_\mathcal{A}$ will follow from the fact that $\Gamma \mapsto \varphi_\Gamma(z)$ is continuous for each $z \in X$. To see this, first note that Definition 19 gives

$$\varphi_\Gamma(z) = \varphi_\Gamma(z) - \varphi_\Gamma(0) = \int_0^z \varphi_\Gamma' = \int_0^z \sum_{n=1}^\infty (f_\mathcal{A} \circ \tau_\Gamma^n)',$$

and rewriting this integral as

$$\sum_{n=1}^\infty \int_0^z (f_\mathcal{A} \circ \tau_\Gamma^n)' = \sum_{n=1}^\infty \int_{\tau_\Gamma^n[0,z]} f_\mathcal{A}' = \sum_{n=1}^\infty \int \mathbb{1}_{\tau_\Gamma^n[0,z]} f_\mathcal{A}' = \int f_\mathcal{A}' \sum_{n=1}^\infty \mathbb{1}_{\tau_\Gamma^n[0,z]}$$

gives

$$\varphi_\Gamma(z) = \int f_\mathcal{A}' H_z(\Gamma), \tag{89}$$

where

$$H_z(\Gamma) = \sum_{n=1}^\infty \mathbb{1}_{\tau_\Gamma^n[0,z]}.$$

Now each map $H_{z,n} : \Gamma \mapsto \mathbb{1}_{\tau_\Gamma^n[0,z]}$ clearly belongs to $C([\Gamma_0, \Gamma_1], L^1)$, the space of continuous functions from $[\Gamma_0, \Gamma_1]$ to $L^1 = L^1(dx)$, and $\sum_{n=1}^\infty H_{z,n}$ is convergent in $C([\Gamma_0, \Gamma_1], L^1)$, so $H_z(\cdot) \in C([\Gamma_0, \Gamma_1], L^1)$. It then follows from (89) that $\Gamma \mapsto \varphi_\Gamma(z)$ is continuous, as required.

Now note that the function $G_\mathcal{A}$, defined by

$$G_\mathcal{A}(t) = \log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right) t^{-1}\right), \tag{90}$$

is strictly decreasing since $a_0, c_0, b_1, d_1 > 0$, so if $t \in \mathcal{T}_\mathcal{A} = (t_0(\mathcal{A}), t_1(\mathcal{A}))$, then

$$G_\mathcal{A}(t) \in (G_\mathcal{A}(t_1(\mathcal{A})), G_\mathcal{A}(t_0(\mathcal{A}))) = (\Delta_\mathcal{A}(\Gamma_1), \Delta_\mathcal{A}(\Gamma_0)). \tag{91}$$

Now $\Delta_\mathcal{A}$ is continuous, so, applying the intermediate value theorem to this function (defined on the interval $[\Gamma_0, \Gamma_1]$), we see that, in view of (91), there exists an $\mathcal{A}$-Sturmian interval, which we denote by $\Gamma_\mathcal{A}(t)$, such that $\Gamma_\mathcal{A}(t) \in (\Gamma_0, \Gamma_1)$ and

$$\Delta_\mathcal{A}(\Gamma_\mathcal{A}(t)) = G_\mathcal{A}(t). \tag{92}$$

In other words,

$$\log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right) t^{-1}\right) = \Delta_\mathcal{A}(\Gamma_\mathcal{A}(t)),$$

so that Corollary 11 implies that $f_{\mathcal{A}(t)} + \varphi_{\Gamma_t} - \varphi_{\Gamma_t} \circ T_\mathcal{A} = \int f_{\mathcal{A}(t)} \, ds_{\Gamma_\mathcal{A}(t)}$ on $\Gamma_\mathcal{A}(t)$, as required. $\qquad\square$

9. The case when one matrix dominates

It will be useful to record the value of the induced function $f_{\mathcal{A}}$ at the two fixed points of $T_{\mathcal{A}}$.

LEMMA 26. *For $\mathcal{A} \in \mathfrak{C}$ and $i \in \{0, 1\}$,*

$$f_{\mathcal{A}}(p_{A_i}) = \log\left(\frac{\det A_i}{a_i - b_i(1 + \varrho_{A_i}^{-1})}\right) = \log\left(\frac{\det A_i}{a_i - b_i - (1/2)(\beta_{A_i} + \gamma_{A_i})}\right).$$

Proof. This is a straightforward computation using (19), (21) and (43). ☐

We first consider a sufficient condition for the projectively concave matrix A_0 to be the dominant matrix of the pair $\mathcal{A} = (A_0, A_1)$.

THEOREM 5. *If $\mathcal{A} \in \mathfrak{C}$ is such that*

$$t_0(\mathcal{A}) \geq 1, \tag{93}$$

then the Dirac measure at the fixed point p_{A_0} is the unique $f_{\mathcal{A}}$-maximizing measure; in particular, the joint spectral radius of $\mathcal{A}$ is equal to the spectral radius of A_0.

Proof. Choosing $\varphi(x) = \varphi_{\Gamma_0}(x) = \log((x + \varrho_{A_0})/\varrho_{A_0})$ ensures, by Lemma 20, that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is constant when restricted to $X_{A_0} = \Gamma_0$, and the constant value assumed by this function is clearly $f_{\mathcal{A}}(p_{A_0})$. The result will follow if we can show that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is strictly decreasing on X_{A_1} and that the value $(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(b_1/(b_1 + d_1))$ at the left endpoint of X_{A_1} is no greater than the constant value $f_{\mathcal{A}}(p_{A_0})$. This is because the Dirac measure $\delta_{p_{A_0}}$ will then clearly be the unique maximizing measure for $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$, and hence the unique maximizing measure for $f_{\mathcal{A}}$.

To compute the value $(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(b_1/(b_1 + d_1))$, we recall from (78) that

$$f_{\mathcal{A}}\left(\frac{b_1}{b_1 + d_1}\right) = f_{\mathcal{A}}(T_{A_1}(0)) = \log(b_1 + d_1)$$

and note that

$$\varphi\left(T_{\mathcal{A}}\left(\frac{b_1}{b_1 + d_1}\right)\right) = \varphi(0) = 0$$

and

$$\varphi\left(\frac{b_1}{b_1 + d_1}\right) = \log\left(\frac{b_1/(b_1 + d_1) + \varrho_{A_0}}{\varrho_{A_0}}\right) = \log\left(\frac{(d_1 + b_1(1 + \varrho_{A_0}^{-1}))}{b_1 + d_1}\right).$$

Therefore

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})\left(\frac{b_1}{b_1 + d_1}\right) = \log(d_1 + b_1(1 + \varrho_{A_0}^{-1})). \tag{94}$$

By Lemma 26,

$$f_{\mathcal{A}}(p_{A_0}) = \log\left(\frac{\det A_0}{a_0 - b_0(1 + \varrho_{A_0}^{-1})}\right), \tag{95}$$

so (94) and (95) imply that the desired inequality

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})\left(\frac{b_1}{b_1 + d_1}\right) \leq f_{\mathcal{A}}(p_{A_0})$$

is precisely the hypothesis (93), since

$$t_0(\mathcal{A}) = \frac{\det A_0}{(a_0 - b_0(1 + \varrho_{A_0}^{-1}))(d_1 + b_1(1 + \varrho_{A_0}^{-1}))}$$

by (83).

It remains to show that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is strictly decreasing on X_{A_1}. Suppose that $x \in X_{A_1}$. We know by (43) that

$$f_{\mathcal{A}}(x) = \log\left(\frac{\det A_1}{-\alpha_{A_1}(x + \sigma_{A_1})}\right).$$

Now

$$\varphi(x) = \log\left(\frac{x + \varrho_{A_0}}{\varrho_{A_0}}\right),$$

so

$$\varphi(T_{\mathcal{A}}(x)) = \log\left(\frac{S_{A_1}(x) + \varrho_{A_0}}{\varrho_{A_0}}\right),$$

and therefore

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(x) = \log\left(\frac{\det A_1(x + \varrho_{A_0})}{-\alpha_{A_1}(x + \sigma_{A_1})(S_{A_1}(x) + \varrho_{A_0})}\right).$$

It, therefore, suffices to show that

$$x \mapsto \frac{x + \varrho_{A_0}}{-\alpha_{A_1}(x + \sigma_{A_1})(S_{A_1}(x) + \varrho_{A_0})} \tag{96}$$

is strictly decreasing. For this, note that

$$S_{A_1}(x) + \varrho_{A_0} = \frac{(b_1 + d_1)x - b_1}{-\alpha_{A_1}(x + \sigma_{A_1})} + \varrho_{A_0} = \frac{(b_1 + d_1 - \alpha_{A_1}\varrho_{A_0})x + (a_1 - b_1)\varrho_{A_0} - b_1}{-\alpha_{A_1}(x + \sigma_{A_1})},$$

so (96) is seen to be the Möbius function

$$x \mapsto \frac{x + \varrho_{A_0}}{(b_1 + d_1 - \alpha_{A_1}\varrho_{A_0})x + (a_1 - b_1)\varrho_{A_0} - b_1},$$

which is known to be strictly decreasing by Lemma 14. □

As a consequence of Theorem 5, we obtain the following corollary.

COROLLARY 13. *If $\mathcal{A} \in \mathfrak{C}$ and $t \in \mathbb{R}^+$ are such that*

$$t \le t_0(\mathcal{A}), \tag{97}$$

then the Dirac measure at the fixed point p_{A_0} is the unique $f_{\mathcal{A}(t)}$-maximizing measure; in particular, the joint spectral radius of $\mathcal{A}(t)$ is equal to the spectral radius of A_0.

Proof. The assumption (97) means, using (82), that $t_0(\mathcal{A}(t)) \ge 1$, so the result follows by applying Theorem 5 with $\mathcal{A}$ replaced by $\mathcal{A}(t)$. □

We now turn to an analogous sufficient condition for the projectively convex matrix A_1 to be dominant.

THEOREM 6. *If $\mathcal{A} \in \mathfrak{C}$ is such that*

$$t_1(\mathcal{A}) \le 1, \tag{98}$$

then the Dirac measure at the fixed point p_{A_1} is the unique $f_{\mathcal{A}}$-maximizing measure; in particular, the joint spectral radius of $\mathcal{A}$ is equal to the spectral radius of A_1.

Proof. Choosing $\varphi(x) = \varphi_{\Gamma_1}(x) = \log((x + \varrho_{A_1})/\varrho_{A_1})$ ensures, by Lemma 20, that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is constant when restricted to $X_{A_1} = \Gamma_1$, and the constant value assumed by this function is clearly $f_{\mathcal{A}}(p_{A_1})$. The result will follow if we can show that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is strictly increasing on X_{A_0} and that the value $(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(a_0/(a_0 + c_0))$ at the right endpoint of X_{A_0} is no greater than the constant value $f_{\mathcal{A}}(p_{A_1})$. This is because the Dirac measure $\delta_{p_{A_1}}$ will then clearly be the unique maximizing measure for $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$, and hence the unique maximizing measure for $f_{\mathcal{A}}$.

To compute the value $(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(a_0/(a_0 + c_0))$, we recall from (77) that

$$f_{\mathcal{A}}\left(\frac{a_0}{a_0 + c_0}\right) = f_{\mathcal{A}}(T_{A_0}(1)) = \log(a_0 + c_0),$$

and note that

$$\varphi\left(T\left(\frac{a_0}{a_0 + c_0}\right)\right) = \varphi(1) = \log\left(\frac{1 + \varrho_{A_1}}{\varrho_{A_1}}\right) = \log(1 + \varrho_{A_1}^{-1})$$

and

$$\varphi\left(\frac{a_0}{a_0 + c_0}\right) = \log\left(\frac{a_0/(a_0 + c_0) + \varrho_{A_1}}{\varrho_{A_1}}\right) = \log\left(\frac{(c_0 + a_0(1 + \varrho_{A_1}^{-1}))}{a_0 + c_0}\right).$$

Therefore

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})\left(\frac{a_0}{a_0 + c_0}\right) = \log(a_0 + c_0(1 + \varrho_{A_1}^{-1})^{-1}). \tag{99}$$

By Lemma 26,

$$f_{\mathcal{A}}(p_{A_1}) = \log\left(\frac{\det A_1}{a_1 - b_1(1 + \varrho_{A_1}^{-1})}\right), \tag{100}$$

so (99) and (100) imply that the desired inequality

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})\left(\frac{a_0}{a_0 + c_0}\right) \le f_{\mathcal{A}}(p_{A_1})$$

is precisely the hypothesis (98), since

$$t_1(\mathcal{A}) = \frac{(a_0 + c_0(1 + \varrho_{A_1}^{-1})^{-1})(a_1 - b_1(1 + \varrho_{A_1}^{-1}))}{\det A_1}$$

by (83).

It remains to show that $f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}}$ is strictly increasing on X_{A_0}. Suppose that $x \in X_{A_0}$. We know by (43) that

$$f_{\mathcal{A}}(x) = \log\left(\frac{\det A_0}{-\alpha_{A_0}(x + \sigma_{A_0})}\right).$$

Now

$$\varphi(x) = \log\left(\frac{x + \varrho_{A_1}}{\varrho_{A_1}}\right),$$

so

$$\varphi(T_{\mathcal{A}}(x)) = \log\left(\frac{S_{A_0}(x) + \varrho_{A_1}}{\varrho_{A_1}}\right),$$

and therefore

$$(f_{\mathcal{A}} + \varphi - \varphi \circ T_{\mathcal{A}})(x) = \log\left(\frac{\det A_0(x + \varrho_{A_1})}{-\alpha_{A_0}(x + \sigma_{A_0})(S_{A_0}(x) + \varrho_{A_1})}\right).$$

It, therefore, suffices to show that

$$x \mapsto \frac{x + \varrho_{A_1}}{-\alpha_{A_0}(x + \sigma_{A_0})(S_{A_0}(x) + \varrho_{A_1})} \tag{101}$$

is strictly increasing. For this, note that

$$S_{A_0}(x) + \varrho_{A_1} = \frac{(b_0 + d_0)x - b_0}{-\alpha_{A_0}(x + \sigma_{A_0})} + \varrho_{A_1} = \frac{(b_0 + d_0 - \alpha_{A_0}\varrho_{A_1})x + (a_0 - b_0)\varrho_{A_1} - b_0}{-\alpha_{A_0}(x + \sigma_{A_0})},$$

so (101) is seen to be the Möbius function

$$x \mapsto \frac{x + \varrho_{A_1}}{(b_0 + d_0 - \alpha_{A_0}\varrho_{A_1})x + (a_0 - b_0)\varrho_{A_1} - b_0},$$

which is known to be strictly increasing by Lemma 14. $\square$

As a consequence of Theorem 6 we obtain the following corollary.

COROLLARY 14. *If $\mathcal{A} \in \mathfrak{C}$ and $t \in \mathbb{R}^+$ are such that*

$$t \geq t_1(\mathcal{A}), \tag{102}$$

then the Dirac measure at the fixed point p_{A_1} is the unique $f_{\mathcal{A}(t)}$-maximizing measure; in particular, the joint spectral radius of $\mathcal{A}(t)$ is equal to the spectral radius of tA_1.

Proof. The assumption (102) means, using (82), that $t_1(\mathcal{A}(t)) \leq 1$, so the result follows by applying Theorem 6 with $\mathcal{A}$ replaced by $\mathcal{A}(t)$. $\square$

10. *Sturmian maximizing measures*

It is at this point that we make the extra hypothesis that the matrix pair $\mathcal{A}$ lies in the class $\mathfrak{D} \subset \mathfrak{C}$. By Lemma 17(ii), we know that if $\mathcal{A} \in \mathfrak{C}$, then $f_{\mathcal{A}}$ is strictly increasing on X_{A_0} and strictly decreasing on X_{A_1}; the following result asserts that if we make the stronger hypothesis that $\mathcal{A} \in \mathfrak{D}$, then these monotonicity properties are inherited by all functions formed by adding a Sturmian transfer function φ_Γ to $f_{\mathcal{A}}$.

PROPOSITION 11. *Let $\mathcal{A} \in \mathfrak{D}$. For each $\mathcal{A}$-Sturmian interval $\Gamma \in \mathcal{I}_{\mathcal{A}}$, the function $f_{\mathcal{A}} + \varphi_\Gamma : X_{\mathcal{A}} \to \mathbb{R}$ is strictly increasing on X_{A_0} and strictly decreasing on X_{A_1}.*

Proof. First, suppose that $x \in X_{A_0}$. Let $0 = i_0 < i_1 < i_2 < \cdots$ be the sequence of all integers such that $\tau_\Gamma^{i_k}(x) \in X_{A_0}$.

For $k \geq 0$, writing $z = \tau_\Gamma^{i_k}(x)$, we see that if $1 \leq i < i_{k+1} - i_k$, then $\tau_\Gamma^i(z) \in X_{A_1}$, and thus $\tau_\Gamma^i(z) = T_{A_1}^i(z)$, so that

$$\sum_{i=0}^{i_{k+1}-i_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) = f_\mathcal{A}'(z) + \sum_{i=1}^{i_{k+1}-i_k-1} (f_\mathcal{A} \circ T_{A_1}^i)'(z) > f_\mathcal{A}'(z) + \sum_{i=1}^{\infty} (f_\mathcal{A} \circ T_{A_1}^i)'(z),$$
(103)

where the inequality is because $(f_\mathcal{A} \circ T_{A_1}^i)'(z) < 0$ for all $i \geq 1$, by Lemma 17. Now $z \in X_{A_0}$, so (55) in Lemma 16(iii) gives $f_\mathcal{A}'(z) = -(z + \sigma_{A_0})^{-1}$ (which is positive), and formula (60) from Corollary 7 gives $\sum_{i=1}^{\infty} (f_\mathcal{A} \circ T_{A_1}^i)'(z) = (z + \varrho_{A_1})^{-1}$ (which is negative), so (103) implies that

$$\sum_{i=0}^{i_{k+1}-i_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) > \frac{-1}{z + \sigma_{A_0}} + \frac{1}{z + \varrho_{A_1}}.$$
(104)

However, $\mathcal{A} \in \mathfrak{D}$, so $\varrho_{A_1} < \sigma_{A_0}$, and therefore the right-hand side of (104) is positive, so we have shown that

$$\sum_{i=0}^{i_{k+1}-i_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) > 0.$$

It follows that, for all $k \geq 0$,

$$\sum_{n=i_k}^{i_{k+1}-1} (f_\mathcal{A} \circ \tau_\Gamma^n)'(x) = (\tau_\Gamma^{i_k})'(x) \sum_{i=0}^{i_{k+1}-i_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) > 0,$$

and hence

$$(f_\mathcal{A} + \varphi_\Gamma)'(x) = \sum_{n=0}^{\infty} (f_\mathcal{A} \circ \tau_\Gamma^n)'(x) = \sum_{k=0}^{\infty} \sum_{n=j_k}^{j_{k+1}-1} (f_\mathcal{A} \circ \tau_\Gamma^n)'(x) > 0,$$

so $f_\mathcal{A} + \varphi_\Gamma$ is strictly increasing on X_{A_0}.

Now suppose that $x \in X_{A_1}$. The proof proceeds analogously to the above. Let $0 = j_0 < j_1 < j_2 < \ldots$ be the sequence of all integers such that $\tau_\Gamma^{j_k}(x) \in X_{A_1}$.

For $k \geq 0$, writing $z = \tau_\Gamma^{j_k}(x)$, we see that if $1 \leq i < j_{k+1} - j_k$, then $\tau_\Gamma^i(z) \in X_{A_0}$, and thus $\tau_\Gamma^i(z) = T_{A_0}^i(z)$, so that

$$\sum_{i=0}^{j_{k+1}-j_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) = f_\mathcal{A}'(z) + \sum_{i=1}^{j_{k+1}-j_k-1} (f_\mathcal{A} \circ T_{A_0}^i)'(z) < f_\mathcal{A}'(z) + \sum_{i=1}^{\infty} (f_\mathcal{A} \circ T_{A_0}^i)'(z),$$
(105)

using the fact that $(f_\mathcal{A} \circ T_{A_0}^i)'(z) > 0$ for all $i \geq 1$, by Lemma 17. The right-hand side of (105) can be written as $-(z + \sigma_{A_1})^{-1} + (z + \varrho_{A_0})^{-1}$ using Lemma 16(iii) and Corollary 7, and this is strictly negative since $\sigma_{A_1} < \varrho_{A_0}$, because $\mathcal{A} \in \mathfrak{D}$, so we have shown that

$$\sum_{i=0}^{j_{k+1}-j_k-1} (f_\mathcal{A} \circ \tau_\Gamma^i)'(z) < \frac{-1}{z + \sigma_{A_1}} + \frac{1}{z + \varrho_{A_0}} < 0.$$

It follows that, for all $k \geq 0$,

$$\sum_{n=j_k}^{j_{k+1}-1} (f_{\mathcal{A}} \circ \tau_{\Gamma}^n)'(x) = (\tau_{\Gamma}^{j_k})'(x) \sum_{i=0}^{j_{k+1}-j_k-1} (f_{\mathcal{A}} \circ \tau_{\Gamma}^i)'(z) < 0,$$

and hence

$$(f_{\mathcal{A}} + \varphi_{\Gamma})'(x) = \sum_{n=0}^{\infty} (f_{\mathcal{A}} \circ \tau_{\Gamma}^n)'(x) = \sum_{k=0}^{\infty} \sum_{n=j_k}^{j_{k+1}-1} (f_{\mathcal{A}} \circ \tau_{\Gamma}^n)'(x) < 0,$$

so $f_{\mathcal{A}} + \varphi_{\Gamma}$ is strictly decreasing on X_{A_1}.   □

THEOREM 7. *Let $\mathcal{A} \in \mathfrak{D}$ and $t \in \mathcal{T}_{\mathcal{A}} = (t_0(\mathcal{A}), t_1(\mathcal{A}))$. The $\mathcal{A}$-Sturmian measure supported by the $\mathcal{A}$-Sturmian interval $\Gamma_{\mathcal{A}}(t)$ is the unique maximizing measure for $f_{\mathcal{A}(t)}$; and thus the corresponding Sturmian measure on $\Omega = \{0, 1\}^{\mathbb{N}}$ is the unique $\mathcal{A}(t)$-maximizing measure.*

Proof. Let us write $\varphi = \varphi_{\Gamma_{\mathcal{A}(t)}}$ and $T = T_{\mathcal{A}} = T_{\mathcal{A}(t)}$. We know that $f_{\mathcal{A}(t)} + \varphi - \varphi \circ T$ is a constant function when restricted to $\Gamma_{\mathcal{A}}(t) =: [\gamma_t^-, \gamma_t^+] \cap X_{\mathcal{A}}$, by Proposition 10. In particular,

$$(f_{\mathcal{A}(t)} + \varphi - \varphi \circ T)(\gamma_t^-) = (f_{\mathcal{A}(t)} + \varphi - \varphi \circ T)(\gamma_t^+),$$

and because $T(\gamma_t^-) = T(\gamma_t^+)$, we deduce that

$$(f_{\mathcal{A}(t)} + \varphi)(\gamma_t^-) = (f_{\mathcal{A}(t)} + \varphi)(\gamma_t^+). \tag{106}$$

But Proposition 11 implies that $f_{\mathcal{A}(t)} + \varphi$ is strictly increasing on X_{A_0} and strictly decreasing on X_{A_1}, so, together with (106), we deduce that

$$(f_{\mathcal{A}(t)} + \varphi)(x) > (f_{\mathcal{A}(t)} + \varphi)(y) \quad \text{for all } x \in \Gamma_{\mathcal{A}}(t), \ y \in X_{\mathcal{A}} \backslash \Gamma_{\mathcal{A}}(t). \tag{107}$$

Consequently, if z, z' are such that $T(z) = T(z')$, with $z \in \Gamma_{\mathcal{A}}(t)$ and $z' \notin \Gamma_{\mathcal{A}}(t)$, then

$$(f_{\mathcal{A}(t)} + \varphi)(z) > (f_{\mathcal{A}(t)} + \varphi)(z'),$$

and hence

$$(f_{\mathcal{A}(t)} + \varphi - \varphi \circ T)(z) > (f_{\mathcal{A}(t)} + \varphi - \varphi \circ T)(z').$$

In other words, the constant value of $f_{\mathcal{A}(t)} + \varphi - \varphi \circ T$ on $\Gamma_{\mathcal{A}}(t)$ is its global maximum, and this value is not attained at any point in $X_{\mathcal{A}} \backslash \Gamma_{\mathcal{A}}(t)$.

It follows that the Sturmian measure supported by $\Gamma_{\mathcal{A}}(t)$ is the unique maximizing measure for $f_{\mathcal{A}(t)} + \varphi - \varphi \circ T$, and hence the unique maximizing measure for $f_{\mathcal{A}(t)}$. Thus the corresponding Sturmian measure on $\Omega = \{0, 1\}^{\mathbb{N}}$ is the unique $\mathcal{A}(t)$-maximizing measure.   □

Recall from §1 that $\mathfrak{E} \subset M_2(\mathbb{R})^2$ denotes the set of matrix pairs which are equivalent to some pair in $\mathfrak{D}$, where equivalence of $\mathcal{A} = (A_0, A_1)$ and $\mathcal{A}' = (A'_0, A'_1)$ means that $A'_0 = u P^{-1} A_0 P$ and $A'_1 = v P^{-1} A_1 P$ for some invertible P and $u, v > 0$. We deduce the following theorem.

THEOREM 8. *If $\mathcal{A} \in \mathfrak{E}$ and $t \in \mathbb{R}^+$, then $\mathcal{A}(t)$ has a unique maximizing measure, and this maximizing measure is Sturmian.*

Proof. It suffices to prove the result for $\mathcal{A} \in \mathfrak{D}$, and this is immediate from Corollaries 13 and 14, Theorem 7 and the fact that $\mathfrak{D} \subset \mathfrak{E}$.   □

11. *The parameter map is a devil's staircase*

As noted in Remark 7, if $\mathcal{A} \in \mathfrak{C}$, then there is a topological conjugacy $h_{\mathcal{A}} : \Omega \to Y_{\mathcal{A}}$ between the shift map $\sigma : \Omega \to \Omega$ and the restriction of $T_{\mathcal{A}}$ to the Cantor set $Y_{\mathcal{A}} \subset X_{\mathcal{A}}$; the map $h_{\mathcal{A}}$ is strictly increasing with respect to the orders on Ω and $Y_{\mathcal{A}}$ (cf. Remark 10). If $d : \Omega \to [0, 1]$ is as in Proposition 1(c), associating to $\omega \in \Omega$ the Sturmian parameter of the measure supported by $[0\omega, 1\omega]$, then the map $d_{\mathcal{A}} : Y_{\mathcal{A}} \to [0, 1]$ given by $d_{\mathcal{A}} = d \circ h_{\mathcal{A}}^{-1}$ enjoys the same properties as d.

LEMMA 27. *The map $d_{\mathcal{A}} : Y_{\mathcal{A}} \to [0, 1]$ is continuous, non-decreasing and surjective. The preimage $d_{\mathcal{A}}^{-1}(\mathcal{P})$ is a singleton if $\mathcal{P}$ is irrational and a positive-length closed interval if $\mathcal{P}$ is rational.*

Proof. This is immediate from Proposition 1(c) and the fact that $h_{\mathcal{A}}$ is strictly increasing. □

Note that $d_{\mathcal{A}}$ associates to $y \in Y_{\mathcal{A}}$ the parameter of the $\mathcal{A}$-Sturmian measure supported by the $\mathcal{A}$-Sturmian interval $c_{\mathcal{A}}^{-1}(y)$, where we recall from Definition 12 that the identification map $c_{\mathcal{A}} : \mathcal{I}_{\mathcal{A}} \to [0, 1]$ is defined by $c_{\mathcal{A}}(\Gamma) = T_{\mathcal{A}}(\min \Gamma) = T_{\mathcal{A}}(\max \Gamma)$. Of the extensions of the function $d_{\mathcal{A}}$ from the Cantor set $Y_{\mathcal{A}}$ to the interval $X = [0, 1]$, there is a unique one giving a non-decreasing self-map $d_{\mathcal{A}} : X \to [0, 1]$. This extension, which we shall also denote by $d_{\mathcal{A}}$, is continuous, and $d_{\mathcal{A}}(c)$ is just the parameter of the $\mathcal{A}$-Sturmian measure $s_{c_{\mathcal{A}}^{-1}(c)}$ (i.e. of the $\mathcal{A}$-Sturmian measure supported by the $\mathcal{A}$-Sturmian interval $c_{\mathcal{A}}^{-1}(c)$) for each $c \in X$. We therefore have the following corollary.

COROLLARY 15. *The map $d_{\mathcal{A}} : X \to [0, 1]$ is continuous, non-decreasing and surjective. The preimage $d_{\mathcal{A}}^{-1}(\mathcal{P})$ is a singleton if $\mathcal{P}$ is irrational and a positive-length closed interval if $\mathcal{P}$ is rational.*

Definition 17. For $\mathcal{A} \in \mathfrak{D}$, let $\mathcal{P}_{\mathcal{A}}(t)$ denote the parameter of the Sturmian maximizing measure for $\mathcal{A}(t)$ or, equivalently, of the $\mathcal{A}$-Sturmian $f_{\mathcal{A}(t)}$-maximizing measure. This defines the *parameter map* $\mathcal{P}_{\mathcal{A}} : \mathbb{R}^+ \to [0, 1]$.

Recalling (see Proposition 10) the map $t \mapsto \Gamma_{\mathcal{A}}(t)$ associating $\mathcal{A}$-Sturmian interval to parameter $t \in \mathcal{T}_{\mathcal{A}} = (t_0(\mathcal{A}), t_1(\mathcal{A}))$, we see that, in fact, the map $\mathcal{P}_{\mathcal{A}} : \mathcal{T}_{\mathcal{A}} \to X$ can be written as

$$\mathcal{P}_{\mathcal{A}} = d_{\mathcal{A}} \circ c_{\mathcal{A}} \circ \Gamma_{\mathcal{A}}. \tag{108}$$

This means that $\mathcal{P}_{\mathcal{A}}$ will enjoy the same properties as established for $d_{\mathcal{A}}$ in Corollary 15, provided $c_{\mathcal{A}} \circ \Gamma_{\mathcal{A}}$ is strictly increasing.

LEMMA 28. *For $\mathcal{A} \in \mathfrak{C}$, the map $c_{\mathcal{A}} \circ \Gamma_{\mathcal{A}} : \mathcal{T}_{\mathcal{A}} \to X$ is strictly increasing and surjective.*

Proof. Recall, from (90), the function $G_{\mathcal{A}}$ given by

$$G_{\mathcal{A}}(t) = \log\left(\left(\frac{a_0 + c_0}{b_1 + d_1}\right) t^{-1}\right)$$

and that $\Gamma_{\mathcal{A}}(t) \in \mathcal{I}_{\mathcal{A}}$ is defined (see (92)) by the identity

$$\Delta_{\mathcal{A}} \circ \Gamma_{\mathcal{A}} = G_{\mathcal{A}}.$$

Now $G_{\mathcal{A}}$ is strictly decreasing, so is, in particular, injective, and therefore the map $\Gamma_{\mathcal{A}}$ is necessarily injective. Note that $\Gamma_{\mathcal{A}}$ clearly extends to a continuous injection on $\overline{\mathcal{T}_{\mathcal{A}}} = [t_0(\mathcal{A}), t_1(\mathcal{A})]$, with $\Gamma_{\mathcal{A}}(t_i(\mathcal{A})) = \Gamma_i$ for $i \in \{0, 1\}$.

Now $c_{\mathcal{A}} : \mathcal{I}_{\mathcal{A}} \to X$ is a bijection, so $c_{\mathcal{A}} \circ \Gamma_{\mathcal{A}} : \overline{\mathcal{T}_{\mathcal{A}}} \to X$ is injective, and its continuity means that it is strictly monotone. But $c_{\mathcal{A}}(\Gamma_{\mathcal{A}}(t_0(\mathcal{A}))) = 0$ and $c_{\mathcal{A}}(\Gamma_{\mathcal{A}}(t_1(\mathcal{A}))) = 1$, so the map $c_{\mathcal{A}} \circ \Gamma_{\mathcal{A}}$ must be strictly increasing and surjective, as required. $\square$

We can now prove that the parameter map $\mathcal{P}_{\mathcal{A}} : \mathbb{R}^+ \to [0, 1]$ is singular. More specifically, its properties described by the following Theorem 9 mean that it is a *devil's staircase*. These properties of the parameter map were noted by Bousch and Mairesse [**4**] in the context of the family (3) and were proved in detail by Morris and Sidorov [**16**] for the family (5). The following result can be viewed as a more detailed version of Theorem 3 from §1.

THEOREM 9. *If $\mathcal{A} \in \mathfrak{E}$ and $t \in \mathbb{R}^+$, then $\mathcal{A}(t)$ has a unique maximizing measure, and this maximizing measure is Sturmian. Let $\mathcal{P}_{\mathcal{A}}(t)$ denote the parameter of the Sturmian maximizing measure for $\mathcal{A}(t)$. The parameter map $\mathcal{P}_{\mathcal{A}} : \mathbb{R}^+ \to [0, 1]$ is continuous, nondecreasing and surjective. The preimage $\mathcal{P}_{\mathcal{A}}^{-1}(\mathcal{P})$ is a singleton if $\mathcal{P}$ is irrational and a positive-length closed interval if $\mathcal{P}$ is rational.*

Proof. The set $\mathfrak{E}$ consists of matrix pairs which are equivalent to a matrix pair in $\mathfrak{D}$, so it suffices to prove the result for $\mathcal{A} \in \mathfrak{D}$. Theorem 8 gives that $\mathcal{A}(t)$ has a unique maximizing measure and that this maximizing measure is Sturmian.

For $t \in \mathbb{R}^+ \setminus \mathcal{T}_{\mathcal{A}}$ we know that

$$\mathcal{P}_{\mathcal{A}}(t) = 0 \quad \text{for } t \in (0, t_0(\mathcal{A})) \tag{109}$$

by Theorem 13, and that

$$\mathcal{P}_{\mathcal{A}}(t) = 1 \quad \text{for } t \in (t_1(\mathcal{A}), \infty) \tag{110}$$

by Theorem 14, since the Dirac measures at the fixed points p_{A_0} and p_{A_1} are $\mathcal{A}$-Sturmian measures of parameters zero and one, respectively.

In view of (109) and (110), it suffices to establish the required properties of $\mathcal{P}_{\mathcal{A}}$ on the sub-interval $\mathcal{T}_{\mathcal{A}} = (t_0(\mathcal{A}), t_1(\mathcal{A}))$. Using the factorization (108), we see that this follows from Corollary 15 and Lemma 28. $\square$

REFERENCES

[1] V. D. Blondel. The birth of the joint spectral radius: an interview with Gilbert Strang. *Linear Algebra Appl.* **428** (2008), 2261–2264.

[2] V. D. Blondel, J. Theys and A. A. Vladimirov. An elementary counterexample to the finiteness conjecture. *SIAM J. Matrix Anal.* **24** (2003), 963–970.

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

[4] T. Bousch and J. Mairesse. Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture. *J. Amer. Math. Soc.* **15** (2002), 77–111.

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

[**6**] I. Daubechies and J. C. Lagarias. Sets of matrices all infinite products of which converge. *Linear Algebra Appl.* **162** (1992), 227–261.

[**7**] L. Gurvits. Stability of Linear Inclusions–Part 2. *NECI Technical Report*, 1996, pp. 96–173.

[**8**] K. G. Hare, I. D. Morris, N. Sidorov and J. Theys. An explicit counterexample to the Lagarias–Wang finiteness conjecture. *Adv. Math.* **226** (2011), 4667–4701.

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

[**10**] R. Jungers. *The Joint Spectral Radius (Lecture Notes in Control and Information Sciences, 385)*. Springer, Berlin, 2009.

[**11**] R. Kannan and C. K. Krueger. *Advanced Analysis on the Real Line*. Springer, New York, 1996.

[**12**] V. S. Kozyakin. A dynamical systems construction of a counterexample to the finiteness conjecture. *Proc. 44th IEEE Conf. Decision and Control and European Control Conf. (Seville, Spain, December 2005)*. IEEE, Piscataway, NJ, pp. 2338–2343.

[**13**] V. S. Kozyakin. An annotated bibliography on convergence of matrix products and the theory of joint/generalized spectral radius. *Preprint*, 2013, available online at https://pdfs.semanticscholar.org/f65 0/fcd2954934f3c9c9414c37f276cff4909b94.pdf.

[**14**] J. C. Lagarias and Y. Wang. The finiteness conjecture for the generalized spectral radius of a set of matrices. *Linear Algebra Appl.* **214** (1995), 17–42.

[**15**] M. Maesumi. Optimal norms and the computation of joint spectral radius of matrices. *Linear Algebra Appl.* **428** (2008), 2324–2338.

[**16**] I. D. Morris and N. Sidorov. On a devil's staircase associated to the joint spectral radii of a family of pairs of matrices. *J. Eur. Math. Soc. (JEMS)* **15** (2013), 1747–1782.

[**17**] M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62** (1940), 1–42.

[**18**] K. R. Parthasarathy. On the category of ergodic measures. *Illinois J. Math.* **5** (1961), 648–656.

[**19**] G.-C. Rota and G. Strang. A note on the joint spectral radius. *Indag. Math. (N.S.)* **22** (1960), 379–381.

[**20**] K. Sigmund. Generic properties of invariant measures for Axiom *A* diffeomorphisms. *Invent. Math.* **11** (1970), 99–109.

[**21**] G. Strang. The joint spectral radius, commentary by Gilbert Strang on paper number 5, in *Collected Works of Gian-Carlo Rota*, 2001; available online from http://www-math.mit.edu/~gs.

[**22**] P. Walters. *An Introduction to Ergodic Theory*. Springer, Berlin, 1981.