Symbol ratio minimax sequences in the lexicographic order

Philip Boyland, André de Carvalho, and Toby Hall

ABSTRACT

Consider the space of sequences of k letters ordered lexicographically. We study the set $\mathcal{M}(\boldsymbol{\alpha})$ of all maximal sequences for which the asymptotic proportions $\boldsymbol{\alpha}$ of the letters are prescribed, where a sequence is said to be maximal if it is at least as great as all of its tails. The infimum of $\mathcal{M}(\boldsymbol{\alpha})$ is called the $\boldsymbol{\alpha}$-infimax sequence, or the $\boldsymbol{\alpha}$-minimax sequence if the infimum is a minimum. We give an algorithm which yields all infimax sequences, and show that the infimax is *not* a minimax if and only if it is the $\boldsymbol{\alpha}$-infimax for every $\boldsymbol{\alpha}$ in a simplex of dimension 1 or greater. These results have applications to the theory of rotation sets of beta-shifts and torus homeomorphisms.

1. Introduction

Symbolic dynamics is a fundamental tool in dynamical systems theory, and the interaction between the dynamics of the shift map and an order structure is frequently important. For example, kneading theory [**11**] describes the dynamics of a unimodal map as the set of sequences which are less than or equal to the kneading sequence of the map in the unimodal order; while in Parry's work [**22**] on beta-shifts it is the relationship between the shift map and the lexicographic order which plays a central rôle. In such systems, a particular orbit is present if the maximum (or more generally supremum) of the orbit is less than or equal to a given sequence: hence, in order to decide whether or not a given dynamical feature is present, the key question is the size of the minimum, or infimum, of the set of maximal sequences which exhibit the feature. It is for this reason that such *minimax* and *infimax* sequences are important.

This paper provides a description of minimax and infimax sequences in the lexicographic order, where the relevant dynamical feature – closely related to rotation vectors – is the asymptotic proportions of the letters. In the remainder of the introduction we will first give an informal description of the main results, and then expand on their dynamical significance.

Given $k \geq 2$, let $\Sigma = \{1, 2, \ldots, k\}^{\mathbb{N}}$ be the space of sequences in the letters $1, 2, \ldots, k$, ordered lexicographically, and let $\sigma \colon \Sigma \to \Sigma$ be the shift map. A sequence $w \in \Sigma$ is said to be *maximal* if $\sigma^r(w) \leq w$ for all $r \geq 0$.

We are interested in maximal sequences for which the asymptotic proportions of the letters are given by some $\boldsymbol{\alpha} \in \Delta$, where Δ is the set of vectors in $\mathbb{R}^k$ with non-negative entries summing

PB: Department of Mathematics, University of Florida, 372 Little Hall, Gainesville, FL 32611-8105, USA

AdC: Departamento de Matemática Aplicada, IME-USP, Rua Do Matão 1010, Cidade Universitária, 05508-090 São Paulo SP, Brazil

TH: Department of Mathematical Sciences, University of Liverpool, Liverpool L69 7ZL, UK

to 1. Denote by $\mathcal{M}(\boldsymbol{\alpha})$ the subset of Σ consisting of maximal sequences w with the property that, for each i with $1 \leq i \leq k$, the asymptotic proportion of the letter i in w is given by α_i. Let $\mathcal{I}(\boldsymbol{\alpha})$ denote the infimum of the set $\mathcal{M}(\boldsymbol{\alpha})$, the $\boldsymbol{\alpha}$-*infimax* sequence. This infimum, while necessarily maximal, need not in general be an element of $\mathcal{M}(\boldsymbol{\alpha})$: when it is, it is called the $\boldsymbol{\alpha}$-*minimax* sequence.

The main results of the paper can be summarised as follows:

THEOREM 22. (DESCRIPTION OF INFIMAXES)   *There is an algorithm for computing $\mathcal{I}(\boldsymbol{\alpha})$ (to an arbitrary number of letters) in terms of a sequence of substitutions. This sequence of substitutions is determined by the itinerary of $\boldsymbol{\alpha}$ under a multi-dimensional continued fraction map $K \colon \Delta \to \Delta$.*

THEOREM 24. (INFIMAX OR MINIMAX)   *The infimum $\mathcal{I}(\boldsymbol{\alpha})$ of $\mathcal{M}(\boldsymbol{\alpha})$ is a minimum, i.e. is an element of $\mathcal{M}(\boldsymbol{\alpha})$, if and only if $\boldsymbol{\alpha}$ is the only point of Δ with its itinerary.*

We say that $\boldsymbol{\alpha}$ is *regular* if it is the only point of Δ with its itinerary, and that it is *exceptional* otherwise. Whether $\boldsymbol{\alpha}$ is regular or exceptional appears to depend on the growth rate of the itinerary of $\boldsymbol{\alpha}$ in a delicate way: our final result gives a flavour of this dependence.

THEOREM 27. (REGULAR OR EXCEPTIONAL)   *If the itinerary of $\boldsymbol{\alpha}$ grows at most quadratically then $\boldsymbol{\alpha}$ is regular; on the other hand, if it grows sufficiently fast then $\boldsymbol{\alpha}$ is exceptional.*

We now discuss the dynamical implications of these results in more detail. Let X be a shift-invariant subset of Σ. The vector $\boldsymbol{\rho}(w) \in \Delta$ of asymptotic proportions of the letters in an element w of X, if well-defined, is called the *rotation vector* of w, and the collection of all of the rotation vectors of elements of X is called the *rotation set* $\rho(X)$ of X. This terminology is by analogy with manifold dynamics: in fact, in the authors' forthcoming paper *"New rotation sets in a family of torus homeomorphisms"*, these symbolic rotation vectors are related directly to rotation vectors for torus homeomorphisms, and the techniques developed in this paper make it possible to provide a detailed description of all of the rotation sets which arise in a parameterised family of torus homeomorphisms.

When X is a subshift of finite type, a theorem of Ziemian [**26**] states that $\rho(X)$ is a convex set with finitely many extreme points, given by the rotation vectors of the minimal loops of the transition diagram. While this result is useful, subshifts of finite type are rather special, and are often ill-suited to understand dynamical behaviour in parameterised families, since Markov partitions can change dramatically under small changes in the map. Here we consider a broader class: in analogy with kneading theory and beta-shifts, we consider subshifts of the form

$$X(v) = \{w \in \Sigma \,:\, \sigma^r(w) \leq v \text{ for all } r \geq 0\},$$

where $v \in \Sigma$. In fact, since the supremum of any shift-invariant set is a maximal sequence, and since $X(v) = X(\sup X(v))$, it suffices to consider the case where v is maximal, which we henceforth assume.

Now if there is some $w \in \mathcal{M}(\boldsymbol{\alpha})$ with $w \leq v$ then it is clear that $\boldsymbol{\alpha} \in \rho(X(v))$, since if $w \leq v$ and w is maximal than $w \in X(v)$. Recalling that $\mathcal{I}(\boldsymbol{\alpha})$ denotes the infimum of all of the

$w \in \mathcal{M}(\boldsymbol{\alpha})$, it follows that

$$v > \mathcal{I}(\boldsymbol{\alpha}) \implies \boldsymbol{\alpha} \in \rho(X(v)).$$

Similarly, it can be shown (see Lemma 19 below) that if w is *any* (not necessarily maximal) element of Σ with $\boldsymbol{\rho}(w) = \boldsymbol{\alpha}$, then the supremum of the orbit of w is at least $\mathcal{I}(\boldsymbol{\alpha})$. Therefore

$$v < \mathcal{I}(\boldsymbol{\alpha}) \implies \boldsymbol{\alpha} \notin \rho(X(v)).$$

Whether or not $\boldsymbol{\alpha} \in \rho(X(v))$ when $v = \mathcal{I}(\boldsymbol{\alpha})$ depends on whether or not $\mathcal{I}(\boldsymbol{\alpha})$ has rotation vector $\boldsymbol{\alpha}$: that is, on whether it is an $\boldsymbol{\alpha}$-minimax, or only an $\boldsymbol{\alpha}$-infimax. Therefore the results of this paper make it possible to determine whether or not $\boldsymbol{\alpha} \in \rho(X(v))$ by comparing v with the single sequence $\mathcal{I}(\boldsymbol{\alpha})$. Moreover, since a consequence of the above discussion is that $\rho(X(v))$ can only change as v passes through an element of the set $\mathcal{I} = \{\mathcal{I}(\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in \Delta\}$ of infimaxes, understanding how the structure of $\rho(X(v))$ changes as v increases is closely related to understanding the structure of the set of initial segments of $\mathcal{I}$.

It is well known [14, 25] that when $k = 2$, all of the infimaxes are minimaxes and are the Sturmian sequences studied by Morse and Hedlund [21, 16]. Thus the infimax sequences with $k \geq 3$ letters can be seen as extensions of the two letter Sturmians. The construction of infimax sequences described here is reminiscent of the construction of Sturmian sequences through their relationship with continued fraction expansions. First there is a division-remainder procedure, similar to the standard Euclidean algorithm, which produces a sequence $\mathbf{n}$ of non-negative integers, analogous to the partial quotients of a continued fraction expansion (this sequence is the itinerary of the orbit of $\boldsymbol{\alpha}$ under $K \colon \Delta \to \Delta$ with respect to a certain partition of Δ, just as the sequence of partial quotients of the continued fraction expansion of $\alpha \in (0,1)$ is the itinerary of α under the Gauss map). Second, this itinerary is used to construct a sequence of substitutions which are applied successively to the single letter k, producing a sequence of words of increasing lengths, each of which is an initial subword of the infimax. If $\boldsymbol{\alpha}$ is a rational vector then the minimax sequence is periodic, and is determined after finitely many steps of the algorithm.

Since their introduction in 1940, Sturmian sequences have been studied extensively [1]. The wide range of fascinating and deep properties which they exhibit has inspired attempts to find classes of sequences in $k > 2$ symbols which share these properties. It was quickly discovered that no single class generalizes all of the key properties of Sturmians.

The infimax sequences defined here generalize the infimax property of Sturmians to an arbitrary number of symbols, and it is therefore natural to ask: what other properties do they share with Sturmians, and how do they relate to the other generalizations which have been defined?

Some properties are easy to determine. Infimax sequences in the case $k > 2$ are very far from being balanced and, in general, are not of Arnoux-Rauzy type [4], and do not provide good rational approximations. In addition, their rate of convergence to their digit frequency vector can be much slower than that of rigid rotation on a torus (this is shown in the authors' forthcoming paper *"New rotation sets in a family of torus homeomorphisms"*), and their orbit closures are not always uniquely ergodic [8].

Other more subtle properties merit further investigation. Jenkinson [18, 17] has shown that the measures associated with Sturmian sequences minimize the dispersion amongst ergodic

invariant measures of $x \mapsto 2x \mod 1$ having the same digit frequency. Thus, for example, treating binary words as base 2 expansions of points of $[0, 1]$, amongst binary words of length q which contain the symbol 1 exactly p times, the Sturmian word has the smallest geometric mean and the smallest standard deviation about its arithmetic mean p/q. The analogs of this beautiful result for infimaxes are currently being studied.

A fundamental property of Sturmians is that they code the orbits of a rigid rotation of the circle. This is sometimes expressed by saying that rigid circle rotation is their *geometric model*. Rigid rotation on a circle can be viewed as an interval exchange map on two intervals, and interval exchanges on more intervals occur as geometric models for known classes of sequences with $k > 2$ symbols [**23**, **2**, **12**, **19**]. Thus another natural question is whether or not k-symbol infimaxes are always codings of interval exchange maps on k intervals. Some of them are, but there are straightforward counter-examples (such as infimaxes with itinerary $1111\overline{0}$): the issue is therefore to distinguish the two cases. Using the terminology and results of Ferenczi and Zamboni [**13**], this is closely related to the problem of determining which rational frequency vectors give rise to clustering Burrows-Wheeler transforms. As we discuss at the end of Section 2, the infimax sequences code the attractors in the family of interval *translation* maps on k intervals studied by Bruin and Troubetzkoy [**7**, **8**].

Another natural question is the relationship of the multi-dimensional continued fraction algorithm used in this paper with existing algorithms. As with multi-symbol generalizations of Sturmians, no single multi-dimensional continued fraction algorithm yields all of the properties enjoyed in the classical one-dimensional case [**24**]. The purpose of the algorithm defined here is to generate infimax sequences and as such it contributes little to the understanding of rational approximation in higher dimensions. However, the existence of exceptional sets on which all points have the same expansion make it an interesting algorithm in its own right, and questions such as its ergodicity merit further study.

Section 2 contains basic definitions and precise statements of the theorems described above. Some preliminary results are presented in Section 3, and a finite version of the problem is then treated in Section 4: given non-negative integers $a_1, \dots, a_k$, what is the smallest maximal *word* which contains exactly a_i occurrences of each letter i? The solution of this problem is required later in the paper, and also introduces the main ideas in a more straightforward context.

In Section 5 we prove the validity of the algorithm for determining infimax sequences, before finishing, in Section 6, by considering the conditions under which infimax sequences are minimaxes.

2. *Definitions, notation, and statement of results*

Let $k \geq 2$ be the number of letters in our alphabet $\mathcal{A} = \{1, \dots, k\}$. We fix k throughout, and suppress the dependence of objects on it, except in Remark 9 and in the final part of the proof of Theorem 24.

Denote by Σ the space $\mathcal{A}^{\mathbb{N}}$ of sequences with entries in $\mathcal{A}$: we consider 0 to be a natural number, so that elements w of Σ are indexed as $w = (w_r)_{r \geq 0}$. Order Σ lexicographically, and endow it with the product topology (where $\mathcal{A}$ is discrete).

Similarly, denote by $\mathcal{A}^*$ the set of non-trivial finite words over the alphabet $\mathcal{A}$, ordered lexicographically with the convention that any proper initial subword of $W \in \mathcal{A}^*$ is greater than W (this convention is simply to ensure that $\mathcal{A}^*$ is totally ordered, and does not affect any of the results of the paper). Given $W \in \mathcal{A}^*$ and $i \in \mathcal{A}$, write $|W| \geq 1$ for the length of W, and $|W|_i \geq 0$ for the number of occurences of the letter i in W.

If $V, W \in \mathcal{A}^*$, denote by VW the concatenation of V and W, by $\overline{W} = WWWW\ldots$ the element of Σ given by infinite repetition of W, and by $V\overline{W}$ the element $VWWWW\ldots$ of Σ. An element of Σ of the form $\overline{W}$ is said to be *periodic*. Given $W \in \mathcal{A}^*$ and $n \geq 0$, denote $W^n = WW\ldots W$ the n-fold repetition of W, an element of $\mathcal{A}^*$ provided that $n > 0$: if $n = 0$ then W^n denotes the empty word, which will be used only when concatenated with elements of $\mathcal{A}^*$.

If $w \in \Sigma$ and $r \geq 1$ is an integer, write $w^{(r)} = w_0 w_1 \ldots w_{r-1}$, the element of $\mathcal{A}^*$ formed by the first r letters of w.

The shift map $\sigma \colon \Sigma \to \Sigma$ is defined by $\sigma(w)_r = w_{r+1}$. An element w of Σ is said to be *maximal* if it is the maximum element of its σ-orbit: that is, if $\sigma^r(w) \leq w$ for all $r \geq 0$. We write $\mathcal{M} \subset \Sigma$ for the set of maximal elements. Observe that $\mathcal{M}$ is a closed subset of Σ, for if $w \in \Sigma$ is not maximal then there is some $r \geq 0$ with $\sigma^r(w) > w$, and it follows that $\sigma^r(w') > w'$ for all $w' \in \Sigma$ sufficiently close to w.

Given $W \in \mathcal{A}^*$, write $\boldsymbol{\rho}(W) \in \mathbb{Q}^k$ for the vector whose i^{th} component is the proportion of the letter i in W: that is, $\boldsymbol{\rho}(W)_i = |W|_i/|W|$. Let

$$\Delta = \left\{ \boldsymbol{\alpha} \in \mathbb{R}_{\geq 0}^k \ : \ \alpha_k > 0, \ \sum \alpha_i = 1 \right\},$$

the simplex which contains these rational vectors, with the face $\alpha_k = 0$ removed, equipped with the maximum metric d_∞. Removing the face $\alpha_k = 0$ makes the statements of the results of the paper cleaner, and clearly if $\alpha_k = 0$ then the problem reduces to one with a smaller value of k.

Given $\boldsymbol{\alpha} \in \Delta$, denote by $\mathcal{R}(\boldsymbol{\alpha})$ the set of elements of Σ with asymptotic proportions of letters $\boldsymbol{\alpha}$:

$$\mathcal{R}(\boldsymbol{\alpha}) = \left\{ w \in \Sigma \ : \ \boldsymbol{\rho}\left(w^{(r)} \right) \to \boldsymbol{\alpha} \text{ as } r \to \infty \right\} \subset \Sigma.$$

REMARK 1. $\mathcal{R}(\boldsymbol{\alpha})$ is not closed in Σ. For example, when $k = 2$ the sequence $2^r \, \overline{21}$ is an element of $\mathcal{R}(1/2, 1/2)$ for all $r \geq 0$, but $2^r \, \overline{21} \to \overline{2} \notin \mathcal{R}(1/2, 1/2)$ as $r \to \infty$. This is a consequence of the more general observation that the asymptotic proportions of elements of Σ, which depend on their tails, do not interact well with the order and topology on Σ, which are defined using the heads of its elements.

We define also the set of maximal sequences with proportions $\boldsymbol{\alpha}$,

$$\mathcal{M}(\boldsymbol{\alpha}) = \mathcal{M} \cap \mathcal{R}(\boldsymbol{\alpha}).$$

Following on from Remark 1, observe that it is easy to construct elements of $\mathcal{M}(\boldsymbol{\alpha})$. Provided that $\boldsymbol{\alpha} \neq (0, 0, \ldots, 0, 1)$ then there are elements of $\mathcal{R}(\boldsymbol{\alpha})$ for which there is an upper bound N on the number of consecutive occurences of the letter k, and prepending k^{N+1} to such an element yields an element of $\mathcal{M}(\boldsymbol{\alpha})$. On the other hand, if $\boldsymbol{\alpha} = (0, 0, \ldots, 0, 1)$ then $\overline{k} \in \mathcal{M}(\boldsymbol{\alpha})$.

In particular, since every non-empty subset of Σ has an infimum, we can define the $\boldsymbol{\alpha}$-*infimax* sequence $\mathcal{I}(\boldsymbol{\alpha})$ by

$$\mathcal{I}(\boldsymbol{\alpha}) = \inf \mathcal{M}(\boldsymbol{\alpha}).$$

$\mathcal{I}(\boldsymbol{\alpha})$ is necessarily an element of $\mathcal{M}$, but need not be an element of $\mathcal{R}(\boldsymbol{\alpha})$, which is not closed in Σ. In the case that it is (and so is an element of $\mathcal{M}(\boldsymbol{\alpha})$), we call it the $\boldsymbol{\alpha}$-*minimax sequence*.

Having introduced the basic objects of study, we now turn to the algorithm for constructing $\mathcal{I}(\boldsymbol{\alpha})$, which is given in terms of the itinerary of $\boldsymbol{\alpha}$ under a certain dynamical system $K \colon \Delta \to \Delta$, defined piecewise on the subsets

$$\Delta_n = \left\{ \boldsymbol{\alpha} \in \Delta \; : \; \left\lfloor \frac{\alpha_1}{\alpha_k} \right\rfloor = n \right\} \subset \Delta \qquad (n \in \mathbb{N}),$$

where $\lfloor x \rfloor$ denotes the integer part of x. First, let $K_n \colon \Delta_n \to \Delta$ be given by

$$K_n(\boldsymbol{\alpha}) = \left(\frac{\alpha_2}{1 - \alpha_1}, \; \frac{\alpha_3}{1 - \alpha_1}, \; \ldots, \; \frac{\alpha_{k-1}}{1 - \alpha_1}, \; \frac{\alpha_1 - n\alpha_k}{1 - \alpha_1}, \; \frac{(n+1)\alpha_k - \alpha_1}{1 - \alpha_1} \right). \tag{2.1}$$

Each K_n is a *projectivity*: an embedding induced on a subset of $\mathbb{R}^k$ by the action of an element of $\mathrm{GL}_{k+1}(\mathbb{R})$ on projective coordinates in $\mathbb{R}\mathrm{P}^k$. As such, it sends convex sets to convex sets. Its inverse $K_n^{-1} \colon \Delta \to \Delta_n$ is given by

$$K_n^{-1}(\boldsymbol{\alpha}) = \left(\frac{(n+1)\alpha_{k-1} + n\alpha_k}{D}, \; \frac{\alpha_1}{D}, \; \frac{\alpha_2}{D}, \; \ldots, \; \frac{\alpha_{k-2}}{D}, \; \frac{\alpha_{k-1} + \alpha_k}{D} \right), \tag{2.2}$$

where $D = (n+1)\alpha_{k-1} + n\alpha_k + 1$.

Let $J \colon \Delta \to \mathbb{N}$ be given by $J(\boldsymbol{\alpha}) = \lfloor \alpha_1/\alpha_k \rfloor$, so that $\boldsymbol{\alpha} \in \Delta_{J(\boldsymbol{\alpha})}$. Then define $K \colon \Delta \to \Delta$ by

$$K(\boldsymbol{\alpha}) = K_{J(\boldsymbol{\alpha})}(\boldsymbol{\alpha}),$$

which is a multi-dimensional continued fraction map [**24**]. Associated to K is an itinerary map $\Phi \colon \Delta \to \mathbb{N}^{\mathbb{N}}$ defined by

$$\Phi(\boldsymbol{\alpha})_r = J(K^r(\boldsymbol{\alpha})) \qquad (r \in \mathbb{N}).$$

We shall see that the infimax sequence $\mathcal{I}(\boldsymbol{\alpha})$ is obtained from a sequence of substitutions associated with $\Phi(\boldsymbol{\alpha})$. Recall that a *substitution on* $\mathcal{A}$ is a map $\Lambda \colon \mathcal{A} \to \mathcal{A}^*$. Overloading notation, this induces maps $\Lambda \colon \mathcal{A}^* \to \mathcal{A}^*$ and $\Lambda \colon \Sigma \to \Sigma$ which replace each letter of the input sequence with its image: $\Lambda(w_0 w_1 w_2 \ldots) = \Lambda(w_0)\Lambda(w_1)\Lambda(w_2)\ldots$. Define substitutions Λ_n for each $n \in \mathbb{N}$ by

$$\Lambda_n \colon \quad \begin{cases} i & \mapsto & (i+1) & \text{if } 1 \le i \le k-2, \\ (k-1) & \mapsto & k\,1^{n+1}, & \text{and} \\ k & \mapsto & k\,1^n. \end{cases} \tag{2.3}$$

Observe that the expression (2.2) for $K_n^{-1}(\boldsymbol{\alpha})$ results precisely from translating (2.3) in such a way as to give the proportions of each letter in $\Lambda_n(w)$ in terms of the proportions in w, that is,

$$w \in \mathcal{R}(\boldsymbol{\alpha}) \iff \Lambda_n(w) \in \mathcal{R}(K_n^{-1}(\boldsymbol{\alpha})).$$

Given $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$, define substitutions $\Lambda_{\mathbf{n},r}$ for each $r \in \mathbb{N}$ by

$$\Lambda_{\mathbf{n},r} = \Lambda_{n_0} \circ \Lambda_{n_1} \circ \cdots \circ \Lambda_{n_r}.$$

Then define a map $S \colon \mathbb{N}^{\mathbb{N}} \to \Sigma$ by

$$S(\mathbf{n}) = \lim_{r \to \infty} \Lambda_{\mathbf{n},r}\left(\overline{k}\right) = \lim_{r \to \infty} \overline{\Lambda_{\mathbf{n},r}(k)},$$

where in the first definition $\Lambda_{\mathbf{n},r}$ is regarded as a map $\Sigma \to \Sigma$, and in the second as a map $\mathcal{A}^* \to \mathcal{A}^*$. The limit exists since $\Lambda_{n_{r+1}}(k)$ begins with the letter k, and hence $\Lambda_{\mathbf{n},r}(k)$ is an initial subword of $\Lambda_{\mathbf{n},r+1}(k)$ for all r.

The first main theorem of the paper states that, for every $\boldsymbol{\alpha} \in \Delta$, the corresponding infimax sequence is given by $S(\Phi(\boldsymbol{\alpha}))$.

THEOREM 22.   *Let $\boldsymbol{\alpha} \in \Delta$. Then $\mathcal{I}(\boldsymbol{\alpha}) = S(\Phi(\boldsymbol{\alpha}))$.*

The question of whether or not the infimax sequence is a minimax (that is, of whether or not an $\boldsymbol{\alpha}$-minimax exists) is answered by the following result:

THEOREM 24.   *Let $\boldsymbol{\alpha} \in \Delta$. Then*
a) $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) \subset \Delta$ is a d-dimensional simplex for some d with $0 \leq d \leq k - 2$.
b) $\mathcal{I}(\boldsymbol{\alpha})$ is the minimum of $\mathcal{M}(\boldsymbol{\alpha})$ if and only if $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a point.

There is therefore a fundamental distinction between *regular* elements $\boldsymbol{\alpha}$ of Δ, for which $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a point, and *exceptional* elements for which this is not the case. That both possibilities occur is the content of the following theorem.

THEOREM 27.   *Let $\boldsymbol{\alpha} \in \Delta$ and $\mathbf{n} = \Phi(\boldsymbol{\alpha})$.*
a) If there is some C such that $0 < n_r \leq Cr^2$ for all r, then $\boldsymbol{\alpha}$ is regular.
b) If $k \geq 3$ and $n_r \geq 2^{r+2} \prod_{i=0}^{r-1}(n_i + 2)$ for all $r \geq 1$, then $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a simplex of dimension $k - 2$, so that $\boldsymbol{\alpha}$ is exceptional.

Notice that if $k = 2$ then Theorem 24 a) gives that every $\boldsymbol{\alpha} \in \Delta$ is regular.

The growth condition of Theorem 27 b) is designed for ease of proof and can be improved without difficulty. Providing a precise characterisation of the set of regular $\boldsymbol{\alpha}$ when $k \geq 3$, by contrast, appears to be a challenging problem.

The substitutions Λ_n which play a central rôle here appear in a different context in papers of Bruin and Troubetzkoy [8] and Bruin [7] (dealing respectively with the case $k = 3$ and the case $k \geq 3$). These papers are concerned with a certain class of *interval translation mappings* (which are defined similarly to interval exchange mappings except that the images of the monotone pieces can overlap). The most interesting case is when the maps are of infinite type, which means that the attractor is a Cantor set. An interval translation mapping with k monotone pieces which is of infinite type can be renormalized infinitely often, with each renormalization being described by a substitution on the space of k-symbol itineraries. The

dynamics on the attractor is therefore given by the subshift generated by the sequence of substitutions corresponding to the sequence of renormalizations.

It turns out that the substitutions arising from renormalization of k-piece interval translation mappings in the class considered by Bruin and Troubetzkoy are exactly the substitutions Λ_n of (2.3). Their results provide extensions of some of the results of this paper, particularly in the case $k = 3$: see Remarks 25 b) and 29 a).

Cassaigne and Nicolas [10] use the family of limiting sequences obtained by Bruin and Troubetskoy to illustrate techniques for calculating the factor complexity of sequences associated with s-adic constructions. In the language of this paper, they show that the factor complexity function $p(n)$ of infimax sequences in the case $k = 3$ satisfies $p(n + 1) - p(n) \in \{2, 3\}$ for all n, excluding certain trivial cases.

3. Preliminaries

In this section we state some basic facts about the maps defined in Section 2. The proofs are routine, and could be omitted on first reading. The crucial result for what follows is Corollary 6, which asserts that the map $S \circ \Phi$ is lower semi-continuous.

LEMMA 2.   Let $n \in \mathbb{N}$. Then the substitution $\Lambda_n \colon \mathcal{A}^* \to \mathcal{A}^*$ is strictly order-preserving. Similarly $\Lambda_n \colon \Sigma \to \Sigma$ is strictly order-preserving, with $\Lambda_n(\mathcal{M}) \subseteq \mathcal{M}$.

Proof.   To show that $\Lambda_n \colon \mathcal{A}^* \to \mathcal{A}^*$ is strictly order-preserving, suppose that $V, W \in \mathcal{A}^*$ with $V < W$. Then either W is a proper initial subword of V, in which case $\Lambda_n(W)$ is a proper initial subword of $\Lambda_n(V)$, so that $\Lambda_n(V) < \Lambda_n(W)$ as required; or there is some $R \geq 0$ with $V_r = W_r$ for $0 \leq r < R$ and $V_R < W_R$. If $V_R \leq k - 2$ then it is obvious that $\Lambda_n(V) < \Lambda_n(W)$. On the other hand, if $V_R = k - 1$ and $W_R = k$, then $\Lambda_n(V) = \Lambda_n(V_0 \ldots V_{R-1})k1^{n+1} \ldots$, and $\Lambda_n(W) = \Lambda_n(V_0 \ldots V_{R-1})k1^n \ldots$. If W has length $R + 1$ then $\Lambda_n(W)$ is a proper initial subword of $\Lambda_n(V)$, so that $\Lambda_n(V) < \Lambda_n(W)$; while if W has length greater than $R + 1$, then the letter following $\Lambda_n(V_0 \ldots V_{R-1})k1^n$ in $\Lambda_n(W)$, being the first letter in the Λ_n-image of a letter, is not 1, so again $\Lambda_n(V) < \Lambda_n(W)$ as required.

The proof that $\Lambda_n \colon \Sigma \to \Sigma$ is strictly order-preserving is similar but simpler, since there is no longer any need to worry about the ends of the words.

To show that $\Lambda_n(\mathcal{M}) \subseteq \mathcal{M}$, let $w \in \mathcal{M}$. Consider w_0, the first, and hence largest, letter in w. If $w_0 < k - 1$ then $\Lambda_n(w_r) = w_r + 1$ for all r, and it is clear that $\Lambda_n(w) \in \mathcal{M}$. Assume therefore that $w_0 \geq k - 1$, so that $\Lambda_n(w)$ begins with the letter k. Suppose for a contradiction that $\Lambda_n(w)$ is not maximal, so that $\Lambda_n(w) = Vv$ for some $V \in \mathcal{A}^*$ and $v \in \Sigma$ with $v > Vv$. Since $V_0 = k$ we must have $v_0 = k$. Since k can only occur as the first letter in the Λ_n-image of a letter, it follows that $w = Uu$ with $\Lambda_n(U) = V$ and $\Lambda_n(u) = v$. Since w is maximal we have $u \leq Uu$, and since Λ_n is order-preserving we have $v = \Lambda_n(u) \leq \Lambda_n(Uu) = Vv$, which is the required contradiction. $\square$

The following lemma is an immediate consequence of the definition (2.3) of the substitutions Λ_n.

LEMMA 3.   *Let $n_0, n_1, \ldots, n_{k-2}$ be any natural numbers. Then $\Lambda_{n_0} \circ \Lambda_{n_1} \circ \cdots \circ \Lambda_{n_{k-2}}(W)$ has initial letter k for all $W \in \mathcal{A}^*$.*

Proof.   If $W_0 = i < k$, then $(\Lambda_n(W))_0 = i + 1$ for all $n \in \mathbb{N}$; while if $W_0 = k$ then $(\Lambda_n(W))_0 = k$ also.   □

Endow $\mathbb{N}^{\mathbb{N}}$ with the product topology, and order it *reverse lexicographically*: that is, lexicographically with the convention that $0 > 1 > 2 > 3 > \cdots$. This convention is to ensure that $S : \mathbb{N}^{\mathbb{N}} \to \Sigma$ is order-preserving.

LEMMA 4.   *$S : \mathbb{N}^{\mathbb{N}} \to \Sigma$ is continuous and strictly order-preserving, with image contained in $\mathcal{M}$.*

Proof.   Let $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$.

To show that $S(\mathbf{n}) \in \mathcal{M}$, observe that $\Lambda_{\mathbf{n},r}\left(\overline{k}\right) \in \mathcal{M}$ for each r by Lemma 2. The result follows since $\mathcal{M}$ is closed in Σ.

To show that S is continuous at $\mathbf{n}$, observe that since $\Lambda_n(k) = k1^n$, the word $L_{\mathbf{n},r} := \Lambda_{\mathbf{n},r}(k)$ has length at least $1 + \sum_{s=0}^{r} n_s$. Therefore if $\mathbf{m} \in \mathbb{N}^{\mathbb{N}}$ satisfies $\mathbf{m}^{(r+1)} = \mathbf{n}^{(r+1)}$, then $S(\mathbf{m})$ and $S(\mathbf{n})$ agree to at least $1 + \sum_{s=0}^{r} n_s$ letters. This establishes that S is continuous at $\mathbf{n}$ provided that $n_s \neq 0$ for arbitrarily large s.

To show continuity at $\mathbf{n}$ in the case where $\mathbf{n} = n_0 \ldots n_{r-1}\overline{0}$ for some $r \geq 0$, observe that, for $R \geq k - 1$,

$$\Lambda_0^R(k1) = \Lambda_0^{R-1}(k2) = \cdots = \Lambda_0^{R-k+2}(k\,(k-1)) = \Lambda_0^{R-k+1}(k\,k\,1),$$

and, repeating the argument, $\Lambda_0^R(k1)$ has initial subword $k^{1+\lfloor R/(k-1) \rfloor}$. Now if $\mathbf{m} \neq \mathbf{n}$ is very close to $\mathbf{n}$, then $\mathbf{m} = n_0 \ldots n_{r-1} 0^R m_{r+R} \ldots$, where R is very large and $m_{r+R} > 0$. It follows that

$$L_{\mathbf{m},r+R} = \Lambda_{\mathbf{n},r-1}(\Lambda_0^R(k1\ldots)) = \Lambda_{\mathbf{n},r-1}(k^{1+\lfloor R/(k-1) \rfloor} \ldots)$$

agrees with $S(\mathbf{n}) = \Lambda_{\mathbf{n},r-1}\left(\overline{k}\right)$ to at least $1 + \lfloor R/(k-1) \rfloor$ letters, establishing continuity at $\mathbf{n}$ as required.

To show that S is strictly order-preserving, let $\mathbf{m} \in \mathbb{N}^{\mathbb{N}}$ with $\mathbf{m} < \mathbf{n}$, so that there is some $r \in \mathbb{N}$ with $\mathbf{m}^{(r)} = \mathbf{n}^{(r)}$ but $m_r > n_r$ (since $\mathbb{N}^{\mathbb{N}}$ is ordered reverse lexicographically). Then $\Lambda_{\mathbf{n},r-1}(\Lambda_{n_r}(k\ell))$ is an initial subword of $S(\mathbf{n})$ for some letter $\ell \in \mathcal{A}$, while $\Lambda_{\mathbf{n},r-1}(\Lambda_{m_r}(k))$ is an initial subword of $S(\mathbf{m})$. Now $\Lambda_{m_r}(k) = k1^{m_r} < k1^{n_r}\Lambda_{n_r}(\ell) = \Lambda_{n_r}(k\ell)$ since $m_r > n_r$, so that $S(\mathbf{m}) < S(\mathbf{n})$ by Lemma 2 as required.   □

Using the definitions of the product topology on Σ and the lexicographical order on $\mathcal{A}^*$, the standard definition of lower semi-continuity for functions from a metric space X into Σ can be phrased as follows: $f : X \to \Sigma$ is lower semi-continuous at $x \in X$ if

$$\forall R \in \mathbb{N}, \exists \epsilon > 0, \quad d(x,y) < \epsilon \implies f(y)^{(R)} \geq f(x)^{(R)}.$$

Similarly, $f\colon X \to \mathbb{N}^{\mathbb{N}}$ is lower semi-continuous at x if the same condition holds, bearing in mind that the $\geq$ should be interpreted reverse lexicographically.

Although the itinerary map $\Phi\colon \Delta \to \mathbb{N}^{\mathbb{N}}$ is discontinuous at all preimages under K of the discontinuity set of K, it is everywhere lower semi-continuous:

LEMMA 5. $\Phi\colon \Delta \to \mathbb{N}^{\mathbb{N}}$ is lower semi-continuous.

Proof.   We need to show that for all $R \in \mathbb{N}$ and all $\boldsymbol{\alpha} \in \Delta$, there is an $\epsilon > 0$ such that if $d_{\infty}(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ then $\Phi(\boldsymbol{\beta})^{(R)} \geq \Phi(\boldsymbol{\alpha})^{(R)}$. The proof is by induction on R.

For the case $R = 0$, observe that for all $\boldsymbol{\alpha} \in \Delta$ there is some $\epsilon > 0$ such that if $d_{\infty}(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ then $\boldsymbol{\beta} \in \Delta_{J(\boldsymbol{\alpha})} \cup \Delta_{J(\boldsymbol{\alpha})-1}$, so that $J(\boldsymbol{\beta}) \leq J(\boldsymbol{\alpha})$ and hence $\Phi(\boldsymbol{\beta})^{(0)} \geq \Phi(\boldsymbol{\alpha})^{(0)}$ as required.

If $R > 0$, then for each $\boldsymbol{\alpha} \in \Delta$ there is, by the inductive hypothesis, some $\delta > 0$ such that if $d_{\infty}(K(\boldsymbol{\alpha}), \boldsymbol{\gamma}) < \delta$ then $\Phi(\boldsymbol{\gamma})^{(R-1)} \geq \Phi(K(\boldsymbol{\alpha}))^{(R-1)}$. Then, by continuity of $K_{J(\boldsymbol{\alpha})}$, there is some $\epsilon > 0$ such that if $d_{\infty}(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ then either $J(\boldsymbol{\beta}) = J(\boldsymbol{\alpha})$ and $d_{\infty}(K(\boldsymbol{\alpha}), K(\boldsymbol{\beta})) < \delta$; or $J(\boldsymbol{\beta}) < J(\boldsymbol{\alpha})$. In either case, $\Phi(\boldsymbol{\beta})^{(R)} \geq \Phi(\boldsymbol{\alpha})^{(R)}$ as required.   □

Combining Lemma 4 and Lemma 5 gives

COROLLARY 6. $S \circ \Phi\colon \Delta \to \mathcal{M}$ is lower semi-continuous.

The next lemma and remark describe the case in which one of the components of $\boldsymbol{\alpha}$ is zero, so that the problem can be reduced to one over a smaller alphabet.

LEMMA 7.   Let $\boldsymbol{\alpha} \in \Delta$ have itinerary $\mathbf{n} = \Phi(\boldsymbol{\alpha})$, and let $1 \leq i \leq k - 1$. Then $\alpha_i = 0$ if and only if $n_r = 0$ for all $r \equiv i - 1 \bmod k - 1$.

Proof.   By (2.1), if $\alpha_1 = 0$ then $K(\boldsymbol{\alpha})_{k-1} = K_0(\boldsymbol{\alpha})_{k-1} = 0$, while if $\alpha_i = 0$ for some i with $2 \leq i \leq k - 1$ then $K(\boldsymbol{\alpha})_{i-1} = 0$. Since $\Phi(\boldsymbol{\alpha})_0 = 0$ whenever $\alpha_1 = 0$, it follows immediately that if $\alpha_i = 0$ then $n_r = 0$ for all $r \equiv i - 1 \bmod k - 1$.

For the converse observe first, by a straightforward induction on i, that if $0 \leq i \leq k - 2$ and if $n_0 = 0$, then

$$K_{n_i} \circ K_{n_{i-1}} \circ \cdots \circ K_{n_0}(\boldsymbol{\alpha})_j = \begin{cases} \alpha_{j+i+1} / \left(1 - \sum_{\ell=1}^{i+1} \alpha_\ell\right) & \text{if } 1 \leq j \leq k - i - 2, \\ \alpha_1 / \left(1 - \sum_{\ell=1}^{i+1} \alpha_\ell\right) & \text{if } j = k - i - 1, \end{cases}$$

independently of $n_1, \ldots, n_i$. The case $i = k - 2$ gives

$$K_{n_{k-2}} \circ K_{n_{k-1}} \circ \cdots \circ K_{n_0}(\boldsymbol{\alpha})_1 = \frac{\alpha_1}{\alpha_k} \geq \frac{\alpha_1}{1 - \alpha_1},$$

provided only that $n_0 = 0$. Now if $n_r = 0$ for all $r \equiv 0 \bmod k - 1$ then repeated application of this inequality gives $K^{s(k-1)}(\boldsymbol{\alpha})_1 \geq \alpha_1/(1 - \alpha_1)^s$ for all $s \geq 0$, and since $K^{s(k-1)}(\boldsymbol{\alpha})_1 < 1$ for all s it follows that $\alpha_1 = 0$, establishing the converse in the case $i = 1$.

The statement for arbitrary $i \leq k - 1$ follows. For if $n_r = 0$ for all $r \equiv i - 1 \bmod k - 1$, then $\boldsymbol{\beta} = K^{i-1}(\boldsymbol{\alpha})$ has itinerary $\mathbf{m} = \Phi(\boldsymbol{\beta})$ satisfying $m_r = 0$ for all $r \equiv 0 \bmod k - 1$. Therefore $\beta_1 = 0$, and hence $\alpha_i = 0$ by $i - 1$ applications of (2.2). $\qquad\square$

COROLLARY 8. *Let $\boldsymbol{\alpha} \in \Delta$ and $\mathbf{n} = \Phi(\boldsymbol{\alpha})$. Then the following are equivalent:*
a) *For all $r \geq 0$, no component of $K^r(\boldsymbol{\alpha})$ is zero; and*
b) *For all $r \geq 0$, there is some $s \geq 0$ with $n_{r+s(k-1)} \neq 0$.*

REMARK 9. This remark relates the K-orbit of $\boldsymbol{\alpha}$ when some $\alpha_i = 0$ to the K-orbit of the point $\boldsymbol{\beta}$ obtained by deleting the ith component of $\boldsymbol{\alpha}$. We therefore include the value of k in our notation, writing Δ^k instead of Δ. For each i with $1 \leq i \leq k - 1$, write $\Delta^{k,i} = \{\boldsymbol{\alpha} \in \Delta^k : \alpha_i = 0\}$, the ith face of Δ, and let $\pi_i \colon \Delta^{k,i} \to \Delta^{k-1}$ be the bijection which forgets α_i: that is, $\pi_i(\boldsymbol{\alpha}) = (\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_k)$. The bijection π_i clearly also depends on k, as do the maps K, but no confusion will arise from continuing to suppress this dependence.

Now if $k \geq 3$ and $\boldsymbol{\alpha} \in \Delta^{k,i}$, then it follows directly from (2.1) that

$$K(\boldsymbol{\alpha}) = \begin{cases} \pi_{k-1}^{-1} \circ \pi_1(\boldsymbol{\alpha}) & \text{if } i = 1, \\ \pi_{i-1}^{-1} \circ K \circ \pi_i(\boldsymbol{\alpha}) & \text{if } 2 \leq i \leq k - 1. \end{cases}$$

In particular, the itinerary $\Phi(\pi_i(\boldsymbol{\alpha}))$ is obtained from $\Phi(\boldsymbol{\alpha})$ by deleting the zeroes which occur at each position $r \equiv i - 1 \bmod k - 1$.

As stated in Section 2, if $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ it is not in general the case that there is only a single point of Δ with itinerary $\mathbf{n}$. However, it is a straightforward consequence of Lemma 7 that $\Phi^{-1}(\mathbf{n})$ is a single point for itineraries of the form $\mathbf{n} = W\,\overline{0}$.

LEMMA 10. *Let $\mathbf{n} = n_0 n_1 \ldots n_{r-1}\overline{0} \in \mathbb{N}^{\mathbb{N}}$. Then there is a unique $\boldsymbol{\alpha} \in \Delta$ with $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$, namely $\boldsymbol{\alpha} = K_{n_0}^{-1} \circ K_{n_1}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1}(0, 0, \ldots, 0, 1)$.*

Proof. $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$ if and only if $\boldsymbol{\alpha} = K_{n_0}^{-1} \circ K_{n_1}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1}(\boldsymbol{\beta})$ for some $\boldsymbol{\beta}$ with itinerary $\overline{0}$. But $\Phi(\boldsymbol{\beta}) = \overline{0}$ if and only if $\boldsymbol{\beta} = (0, 0, \ldots, 0, 1)$ by Lemma 7. $\qquad\square$

In particular, if $\boldsymbol{\alpha}$ has an itinerary of this form then $\boldsymbol{\alpha} \in \mathbb{Q}^k$. Theorem 16 states that, conversely, every element $\boldsymbol{\alpha}$ of $\Delta \cap \mathbb{Q}^k$ has such an itinerary: that is, that $K^r(\boldsymbol{\alpha}) = (0, 0, \ldots, 0, 1)$ for some r.

4. *The finite version*

In this section we solve a finite version of the minimax problem, which is a necessary precursor to our later results. The simplicity of the solution makes it straightforward to understand the origin of the maps K_n and the substitutions Λ_n.

A word $W \in \mathcal{A}^*$ is said to be maximal if $\overline{W}$ is a maximal element of Σ or, equivalently, if $W = UV \implies W \geq VU$, i.e. W is at least as large as all of its cyclic permutations.

Let $\widehat{\Delta} = \{\boldsymbol{a} = (a_1, \ldots, a_k) \in \mathbb{N}^k : a_k > 0\}$, the discrete analogue of the space Δ. Continuing the analogy, we write for each $\boldsymbol{a} \in \widehat{\Delta}$

$$\widehat{\mathcal{R}}(\boldsymbol{a}) = \{W \in \mathcal{A}^* : |W|_i = a_i \text{ for } 1 \leq i \leq k\} \qquad \text{(a finite set)},$$
$$\widehat{\mathcal{M}}(\boldsymbol{a}) = \{W \in \widehat{\mathcal{R}}(\boldsymbol{a}) : W \text{ is maximal}\}, \qquad \text{and}$$
$$\widehat{\mathcal{I}}(\boldsymbol{a}) = \min \widehat{\mathcal{M}}(\boldsymbol{a}).$$

REMARK 11.   An obvious comment, which is nevertheless important for the proof of Theorem 13 below, is that every $W \in \widehat{\mathcal{R}}(\boldsymbol{a})$ has a cyclic permutation which belongs to $\widehat{\mathcal{M}}(\boldsymbol{a})$.

For each $n \in \mathbb{N}$ write $\widehat{\Delta}_n = \{\boldsymbol{a} \in \widehat{\Delta} : n\, a_k \leq a_1 < (n+1)\, a_k\}$, and define a bijection $\widehat{K}_n \colon \widehat{\Delta}_n \to \widehat{\Delta}$ by

$$\widehat{K}_n(\boldsymbol{a}) = (a_2, a_3, \ldots, a_{k-1}, a_1 - na_k, (n+1)a_k - a_1),$$

whose inverse $\widehat{K}_n^{-1} \colon \widehat{\Delta} \to \widehat{\Delta}_n$, the Abelianization of the substitution Λ_n, is given by

$$\widehat{K}_n^{-1}(\boldsymbol{a}) = ((n+1)a_{k-1} + na_k, a_1, a_2, \ldots, a_{k-2}, a_{k-1} + a_k).$$

LEMMA 12.   Let $\boldsymbol{a} \in \widehat{\Delta}$. Then the set $\Lambda_n^{-1}(\widehat{\mathcal{M}}(\boldsymbol{a}))$ of words whose image under Λ_n lies in $\widehat{\mathcal{M}}(\boldsymbol{a})$ is exactly $\widehat{\mathcal{M}}(\widehat{K}_n(\boldsymbol{a}))$.

Proof.   To show that $\Lambda_n^{-1}(\widehat{\mathcal{M}}(\boldsymbol{a})) \subset \widehat{\mathcal{M}}(\widehat{K}_n(\boldsymbol{a}))$, let $W \in \mathcal{A}^*$ with $\Lambda_n(W) \in \widehat{\mathcal{M}}(\boldsymbol{a})$. Then $W \in \widehat{\mathcal{R}}(\widehat{K}_n(\boldsymbol{a}))$ by comparison of the right-hand side of (2.3) with the formula for $\widehat{K}_n^{-1}$. Moreover, W is maximal: for if $W = UV < VU$ then $\Lambda_n(W) = \Lambda_n(U)\Lambda_n(V) < \Lambda_n(V)\Lambda_n(U)$ by Lemma 2, contradicting the maximality of $\Lambda_n(W)$.

To show that $\Lambda_n(W) \in \widehat{\mathcal{M}}(\boldsymbol{a})$ for all $W \in \widehat{\mathcal{M}}(\widehat{K}_n(\boldsymbol{a}))$, it follows as above that $\Lambda_n(W) \in \widehat{\mathcal{R}}(\boldsymbol{a})$. That it is maximal follows from translating the statement $\Lambda_n(\mathcal{M}) \subseteq \mathcal{M}$ (Lemma 2) into the finite setting:

$$W \text{ maximal} \implies \overline{W} \in \mathcal{M} \implies \Lambda_n(\overline{W}) = \overline{\Lambda_n(W)} \in \mathcal{M} \implies \Lambda_n(W) \text{ maximal}.$$

$\square$

The following theorem gives the fundamental relationship between the substitutions, the linear maps associated to the division-remainder algorithm, and the minimax: the substitution Λ_n sends the minimax for $\boldsymbol{a}$ to the minimax for $\widehat{K}_n^{-1}(\boldsymbol{a})$.

THEOREM 13.   Let $\boldsymbol{a} \in \widehat{\Delta}$.
(a) If $a_i = 0$ for all $i < k$ then $\widehat{\mathcal{I}}(\boldsymbol{a}) = k^{a_k}$.
(b) Otherwise, $\widehat{\mathcal{I}}(\boldsymbol{a}) = \Lambda_n(\widehat{\mathcal{I}}(\widehat{K}_n(\boldsymbol{a})))$, where $n = \lfloor a_1/a_k \rfloor$.

REMARKS 14.
a) The theorem gives rise to a straightforward algorithm for calculating $\widehat{\mathcal{I}}(\boldsymbol{a})$: the key point
   is that the sum of the entries of $\widehat{K}_n(\boldsymbol{a})$ is a_1 less than the sum of the entries of $\boldsymbol{a}$; and if
   $a_i > 0$ for any $i < k$, then a_1 will be positive after applying $i - 1$ terms of the appropriate
   sequence of $\widehat{K}_n$'s. Therefore repeatedly applying $\widehat{K}_{\lfloor a_1/a_k \rfloor}$ eventually yields an $\boldsymbol{a}$ with $a_i = 0$
   for all $i < k$. The hand implementation of this algorithm is illustrated in Examples 15
   below, while the statement of Theorem 13 translates directly into a recursive algorithm for
   computer implementation.
b) By linearity of the $\widehat{K}_n$ we have $\widehat{\mathcal{I}}(N\boldsymbol{a}) = \widehat{\mathcal{I}}(\boldsymbol{a})^N$ for each integer $N \geq 1$.

EXAMPLES 15.   Let $k = 3$ and $\boldsymbol{a} = (24, 3, 14)$. We have

$$(24, 3, 14) \xrightarrow{\widehat{K}_1} (3, 10, 4) \xrightarrow{\widehat{K}_0} (10, 3, 1) \xrightarrow{\widehat{K}_{10}} (3, 0, 1) \xrightarrow{\widehat{K}_3} (0, 0, 1),$$

so that

$$\widehat{\mathcal{I}}(24, 3, 14) = \Lambda_1 \Lambda_0 \Lambda_{10} \Lambda_3(3) = \Lambda_1 \Lambda_0 \Lambda_{10}(31^3) = \Lambda_1 \Lambda_0(31^{10}2^3) = \Lambda_1(3\,2^{10}\,(31)^3) = 31\,(311)^{10}\,(312)^3.$$

Notice that the intermediate words 31^3, $31^{10}2^3$, and $3\,2^{10}\,(31)^3$ are $\widehat{\mathcal{I}}(3, 0, 1)$, $\widehat{\mathcal{I}}(10, 3, 1)$, and
$\widehat{\mathcal{I}}(3, 10, 4)$ respectively.
   Similarly, if $k = 4$ and $\boldsymbol{a} = (2, 3, 1, 3)$ then

$$(2, 3, 1, 3) \xrightarrow{\widehat{K}_0} (3, 1, 2, 1) \xrightarrow{\widehat{K}_3} (1, 2, 0, 1) \xrightarrow{\widehat{K}_1} (2, 0, 0, 1) \xrightarrow{\widehat{K}_2} (0, 0, 0, 1),$$

so that

$$\widehat{\mathcal{I}}(2, 3, 1, 3) = \Lambda_0 \Lambda_3 \Lambda_1 \Lambda_2(4) = \Lambda_0 \Lambda_3 \Lambda_1(411) = \Lambda_0 \Lambda_3(4122) = \Lambda_0(4111233) = 422234141.$$

Proof of Theorem 13.   Statement (a) is obvious, since k^{a_k} is the unique element of
$\widehat{\mathcal{R}}(0, 0, \ldots, 0, a_k)$.
   For (b), it suffices to show that $\widehat{\mathcal{I}}(\boldsymbol{a})$ is in the image of Λ_n, where $n = \lfloor a_1/a_k \rfloor$: the result
then follows immediately from Lemmas 12 and 2.
   Since $a_1 \geq na_k$ there are elements of $\widehat{\mathcal{R}}(\boldsymbol{a})$, and hence, by Remark 11, of $\widehat{\mathcal{M}}(\boldsymbol{a})$, in which
every occurence of the letter k is followed by the word 1^n, and such elements of $\widehat{\mathcal{M}}(\boldsymbol{a})$ are
smaller than any element of $\widehat{\mathcal{M}}(\boldsymbol{a})$ which does not have this property. Therefore

$$\widehat{\mathcal{I}}(\boldsymbol{a}) = k1^n W_1 \, k1^n W_2 \, \cdots \, k1^n W_{a_k}$$

for some words W_r which do not contain the letter k. Moreover, the letters must be arranged in ascending order in each W_r: that is,

$$W_r = 1^{n_{r,1}} 2^{n_{r,2}} \cdots (k-1)^{n_{r,k-1}}$$

for each r, where the $n_{r,s}$ are non-negative integers. For if this were not the case, then replacing each W_r with a word in which the same letters are arranged in ascending order would decrease every cyclic permutation of $\widehat{\mathcal{I}}(\boldsymbol{a})$ starting with k, so that there would be an element of $\widehat{\mathcal{M}}(\boldsymbol{a})$ smaller than $\widehat{\mathcal{I}}(\boldsymbol{a})$.

To show that $\widehat{\mathcal{I}}(\boldsymbol{a})$ is in the image of Λ_n, it therefore suffices to show that $n_{r,1} \leq 1$ for all r. Observe first that $\sum_{r=1}^{a_k} n_{r,1} = a_1 - na_k < a_k$, so that at least one $n_{r,1}$ is zero, and in particular $n_{1,1} = 0$ by maximality of $\widehat{\mathcal{I}}(\boldsymbol{a})$.

Suppose for a contradiction that $n_{s,1} \geq 2$ for some least s. Define words W_r' for $1 \leq r \leq a_k$ by $W_{s-1}' = 1 W_{s-1}$, $W_s = 1 W_s'$, and $W_r' = W_r$ for $r \neq s-1, s$: that is, push one of the 1s from W_s to W_{s-1}. Then taking an appropriate cyclic permutation yields an element W' of $\widehat{\mathcal{M}}(\boldsymbol{a})$ given by

$$W' = k1^n W_t' \, k1^n W_{t+1}' \, \cdots \, k1^n W_{a_k}' \, k1^n W_1' \, \cdots k1^n W_{t-1}',$$

where t is not equal to s since W_s' starts with the letter 1 by choice of s, but W_t' does not start with the letter 1 by maximality of W'. Now

$$W' < \ k1^n W_t \, k1^n W_{t+1} \, \cdots \, k1^n W_{a_k} \, k1^n W_1 \, \cdots k1^n W_{t-1} \ \leq \ \widehat{\mathcal{I}}(\boldsymbol{a}),$$

where the first inequality is by definition of the words W_r' together with $t \neq s$, and the second is by maximality of $\widehat{\mathcal{I}}(\boldsymbol{a})$. This contradicts that $\widehat{\mathcal{I}}(\boldsymbol{a})$ is the minimum element of $\widehat{\mathcal{M}}(\boldsymbol{a})$, establishing that $\widehat{\mathcal{I}}(\boldsymbol{a})$ is in the image of Λ_n as required. $\qquad\square$

To connect this result with the formalism used in the general case, observe that

$$
\begin{array}{ccc}
\widehat{\Delta}_n & \xrightarrow{\ \widehat{K}_n\ } & \widehat{\Delta} \\
\downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\
\Delta_n & \xrightarrow{\ K_n\ } & \Delta
\end{array}
\tag{4.1}
$$

commutes, where $\pi \colon \widehat{\Delta} \to \Delta$ is defined by $\pi(\boldsymbol{a}) = \boldsymbol{a}/\sum a_i$. Moreover, the functions $\widehat{K}_n$ can be gathered into a single function $\widehat{K} \colon \widehat{\Delta} \to \widehat{\Delta}$ defined by $\widehat{K}(\boldsymbol{a}) = \widehat{K}_{\lfloor a_1/a_k \rfloor}(\boldsymbol{a})$, giving rise to an itinerary map $\widehat{\Phi} \colon \widehat{\Delta} \to \mathbb{N}^{\mathbb{N}}$ defined by

$$\widehat{\Phi}(\boldsymbol{a})_r = n \ \Longleftrightarrow \ \widehat{K}^r(\boldsymbol{a}) \in \widehat{\Delta}_n.$$

Since $\widehat{K}^r(\boldsymbol{a}) = (0,0,\ldots,0,1)$, a fixed point of $\widehat{K}$, for some r, the itinerary $\widehat{\Phi}(\boldsymbol{a})$ has only finitely many non-zero entries.

The following is then a restatement of Theorem 13. Note that it does not claim to give the minimum element of $\mathcal{M}(\boldsymbol{\alpha})$ for rational $\boldsymbol{\alpha}$, but only the minimum *periodic* element: that this is in fact the minimum of $\mathcal{M}(\boldsymbol{\alpha})$ will follow from Theorem 22 below.

THEOREM 16.   *Let $\boldsymbol{\alpha} \in \Delta \cap \mathbb{Q}^k$. Then the itinerary of $\boldsymbol{\alpha}$ is of the form $\Phi(\boldsymbol{\alpha}) = n_0\, n_1\, \ldots\, n_{r-1}\,\overline{0}$, and the minimum periodic element $\mathcal{P}(\boldsymbol{\alpha})$ of $\mathcal{M}(\boldsymbol{\alpha})$ is equal to $S(\Phi(\boldsymbol{\alpha}))$.*

Proof.   Let $\boldsymbol{a} \in \widehat{\Delta}$ be the smallest integer vector which is a positive multiple of $\boldsymbol{\alpha}$. Then any periodic element of $\mathcal{M}(\boldsymbol{\alpha})$ is of the form $\overline{W}$, where $W \in \widehat{\mathcal{M}}(N\boldsymbol{a})$ for some $N \geq 1$. However $\widehat{\mathcal{I}}(N\boldsymbol{a}) = \widehat{\mathcal{I}}(\boldsymbol{a})^N$ by Remark 14b), so that the smallest periodic element of $\mathcal{M}(\boldsymbol{\alpha})$ is $\overline{\widehat{\mathcal{I}}(\boldsymbol{a})}$.

It is immediate from (4.1) that $\Phi(\boldsymbol{\alpha}) = \widehat{\Phi}(\boldsymbol{a})$, so that in particular $\mathbf{n} = \Phi(\boldsymbol{\alpha})$ is of the given form. Then

$$\mathcal{P}(\boldsymbol{\alpha}) = \overline{\widehat{\mathcal{I}}(\boldsymbol{a})} = \overline{\Lambda_{\mathbf{n},r-1}(k)} = S(\Phi(\boldsymbol{\alpha}))$$

as required.   □

REMARK 17.   In the computer science and combinatorics of words literature, the term *Lyndon words* is used for words that are minimal amongst their cyclic permutations with respect to the lexicographic order [**3**, **20**]. Therefore maximal words are the same as Lyndon words when the ordering of $\mathcal{A}$ is reversed, and the results of this section can be rephrased as determining the largest Lyndon word with a given number of each of the letters.

5.   Proof of Theorem 22: $\mathcal{I}(\boldsymbol{\alpha}) = S(\Phi(\boldsymbol{\alpha}))$

In this section we prove that the infimum $\mathcal{I}(\boldsymbol{\alpha})$ of $\mathcal{M}(\boldsymbol{\alpha})$ is given by $S(\Phi(\boldsymbol{\alpha}))$. We show first (Lemma 19) that $S(\Phi(\boldsymbol{\alpha}))$ is a lower bound of $\mathcal{M}(\boldsymbol{\alpha})$, and then (Lemma 21) that it lies in the closure of $\mathcal{M}(\boldsymbol{\alpha})$.

That $S(\Phi(\boldsymbol{\alpha}))$ is a lower bound of $\mathcal{M}(\boldsymbol{\alpha})$ is a special case of a more general result. Given any $w \in \Sigma$, define $\sup w \in \mathcal{M}$ by

$$\sup w = \sup_{r \geq 0} \sigma^r(w),$$

so that $w = \sup w$ if and only if $w \in \mathcal{M}$. Lemma 19 below states that if $w \in \mathcal{R}(\boldsymbol{\alpha})$ then $S(\Phi(\boldsymbol{\alpha})) \leq \sup w$: in particular, if $w \in \mathcal{M}(\boldsymbol{\alpha})$ then $S(\Phi(\boldsymbol{\alpha})) \leq w$ as required.

The proof uses the finite version of the result as expressed by Theorem 16, and we start with a lemma which provides appropriate rational approximations to $\boldsymbol{\alpha}$ together with corresponding periodic approximations to the supremum of an element of $\mathcal{R}(\boldsymbol{\alpha})$.

LEMMA 18.   *Let $\boldsymbol{\alpha} \in \Delta$, $w \in \mathcal{R}(\boldsymbol{\alpha})$, $R \in \mathbb{N}$ and $\epsilon > 0$. Then there is some $\boldsymbol{\beta} \in \Delta \cap \mathbb{Q}^k$ and a periodic $v \in \mathcal{M}(\boldsymbol{\beta})$ such that $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ and $(\sup w)^{(R)} = v^{(R)}$.*

Proof.   Write $s = \sup w$. By definition of the supremum, there is some $r \geq 0$ such that $(\sigma^r(w))^{(R)} = s^{(R)}$. Since $\sigma^r(w) \in \mathcal{R}(\boldsymbol{\alpha})$, there is an initial subword $s^{(R)}\, W$ of $\sigma^r(w)$ long enough that

$$d_\infty(\boldsymbol{\alpha}, \boldsymbol{\rho}(s^{(R)}\, W\, 1^R)) < \epsilon.$$

Let U be the length R word with the property that $(\sigma^r(w))^{(2R+|W|)} = s^{(R)} \, W \, U$.

Let $v \in \mathcal{M}$ be the maximal shift of the periodic sequence $u = \overline{s^{(R)} \, W \, 1^R}$. We shall show that $v^{(R)} = s^{(R)}$ which will establish the result, with $\boldsymbol{\beta} = \boldsymbol{\rho}(s^{(R)} \, W \, 1^R)$.

Since s, and hence u, begins with the letter k, $v^{(R)}$ is a subword of $s^{(R)} \, W \, 1^R \le s^{(R)} \, W \, U$: but every length R subword of $s^{(R)} \, W \, U$ is a subword of w, and hence is less than or equal to $s^{(R)}$ by the definition of the supremum. Therefore $v^{(R)} \le s^{(R)}$. On the other hand, however, $v^{(R)} \ge u^{(R)} = s^{(R)}$, since $v \ge u$. This establishes the result. $\qquad\square$

LEMMA 19.   *Let $\boldsymbol{\alpha} \in \Delta$ and $w \in \mathcal{R}(\boldsymbol{\alpha})$. Then $S(\Phi(\boldsymbol{\alpha})) \le \sup w$. In particular, $S(\Phi(\boldsymbol{\alpha}))$ is a lower bound of $\mathcal{M}(\boldsymbol{\alpha})$.*

Proof.   Write $s = \sup w$. To show that $S(\Phi(\boldsymbol{\alpha})) \le s$, it suffices to show that $S(\Phi(\boldsymbol{\alpha}))^{(R)} \le s^{(R)}$ for every $R \in \mathbb{N}$. Fix such an R.

By the lower semi-continuity of $S \circ \Phi$ (Corollary 6), there is some $\epsilon > 0$ such that if $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ then $S(\Phi(\boldsymbol{\alpha}))^{(R)} \le S(\Phi(\boldsymbol{\beta}))^{(R)}$.

By Lemma 18 there is some $\boldsymbol{\beta}$ with $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ and some periodic $v \in \mathcal{M}(\boldsymbol{\beta})$ with $s^{(R)} = v^{(R)}$. Theorem 16 gives $v \ge S(\Phi(\boldsymbol{\beta}))$. Then

$$S(\Phi(\boldsymbol{\alpha}))^{(R)} \le S(\Phi(\boldsymbol{\beta}))^{(R)} \le v^{(R)} = s^{(R)}$$

as required. $\qquad\square$

We now turn to proving that $S(\Phi(\boldsymbol{\alpha})) \in \overline{\mathcal{M}(\boldsymbol{\alpha})}$. To do this we need to construct elements of $\mathcal{M}(\boldsymbol{\alpha})$ which agree with $S(\Phi(\boldsymbol{\alpha}))$ on arbitrarily long initial subwords, and the following straightforward lemma will be used for this purpose.

LEMMA 20.   *Let $\boldsymbol{\alpha} \in \Delta$, $R \in \mathbb{N}$, and $\epsilon > 0$. Then there is some $\boldsymbol{\beta} \in \Delta \cap \mathbb{Q}^k$ such that $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$ and $\Phi(\boldsymbol{\beta})^{(R)} = \Phi(\boldsymbol{\alpha})^{(R)}$.*

Proof.   The proof is by induction on R, with the base case $R = 0$ being the statement that rational elements are dense in Δ.

Suppose then that $R > 0$. Let $n = J(\boldsymbol{\alpha})$, so that $\boldsymbol{\alpha} \in \Delta_n$. Recall that $K|_{\Delta_n} = K_n : \Delta_n \to \Delta$ is a homeomorphism. By the inductive hypothesis, there is a sequence $(\boldsymbol{\gamma}_i)$ in $\Delta \cap \mathbb{Q}^k$ converging to $K(\boldsymbol{\alpha})$ with $\Phi(\boldsymbol{\gamma}_i)^{(R-1)} = \Phi(K(\boldsymbol{\alpha}))^{(R-1)}$ for all i. Let $\boldsymbol{\beta} = K_n^{-1}(\boldsymbol{\gamma}_i)$ for some i large enough that $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}) < \epsilon$. $\qquad\square$

LEMMA 21.   *Let $\boldsymbol{\alpha} \in \Delta$. Then $S(\Phi(\boldsymbol{\alpha})) \in \overline{\mathcal{M}(\boldsymbol{\alpha})}$.*

Proof.   If $\boldsymbol{\alpha} \in \Delta \cap \mathbb{Q}^k$ then $S(\Phi(\boldsymbol{\alpha})) \in \mathcal{M}(\boldsymbol{\alpha})$ by Theorem 16, so we assume that $\boldsymbol{\alpha} \notin \mathbb{Q}^k$, and in particular, by Lemma 10, that $\mathbf{n} = \Phi(\boldsymbol{\alpha})$ has infinitely many non-zero entries.

It suffices to find, for each R, an element w of $\mathcal{M}(\boldsymbol{\alpha})$ with initial subword $\Lambda_{\mathbf{n},R}(k)$. We can assume that $n_{R+1} > 0$, since otherwise we increase R until this is the case.

Using Lemma 20, find for each $r \geq 0$ an element $\boldsymbol{\beta}_r$ of $\Delta \cap \mathbb{Q}^k$ with $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\beta}_r) < 1/2^r$, whose itinerary

$$\Phi(\boldsymbol{\beta}_r) = \mathbf{n}_r = n_{r,0} n_{r,1} \ldots n_{r,L_r} \overline{0},$$

satisfies $n_{r,s} = n_s$ for $0 \leq s \leq R+1$.

Set $U = \Lambda_{n_0} \Lambda_{n_1} \cdots \Lambda_{n_R} \Lambda_{n_{R+1}-1}(k)$, and $W_r = \Lambda_{n_{r,0}} \Lambda_{n_{r,1}} \cdots \Lambda_{n_{r,L_r}}(k)$ for each r, so that $\boldsymbol{\rho}(W_r) = \boldsymbol{\beta}_r$. Write $L = |U|$ and $L_r = |U_r|$ for $r \geq 0$. Choose integers $p_r \geq 1$ for $r \geq 0$ inductively to satisfy $\sum_{s=0}^{r} p_s L_s > 2^r L_{r+1}$. Finally, set

$$w = U\, W_0^{p_0} W_1^{p_1} W_2^{p_2} \ldots . \tag{5.1}$$

We will show that $w \in \mathcal{M}(\boldsymbol{\alpha})$, which will establish the result since it has initial subword $\Lambda_{\mathbf{n},R}(k)$. To show that $w \in \mathcal{R}(\boldsymbol{\alpha})$, let $I = \{(r,s) : r \in \mathbb{N}, 0 \leq s < p_r\}$ ordered lexicographically, and define an increasing function $\ell \colon I \to \mathbb{N}$ by $\ell(r,s) = L + sL_r + \sum_{t=0}^{r-1} p_t L_t$, the index of the beginning of the $(s+1)^{\text{th}}$ subword W_r in (5.1). Now since $\boldsymbol{\rho}(W_r) \to \boldsymbol{\alpha}$ as $r \to \infty$ we have that for all $\epsilon > 0$ there is some J such that $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\rho}(w^{(\ell(r,s))})) < \epsilon$ for all $(r,s) > (J,0)$. On the other hand, given any $t \geq \ell(1,0)$, we have $d_\infty(\boldsymbol{\rho}(w^{(t)}), \boldsymbol{\rho}(w^{(\ell(r,s))})) < 1/2^r$, where (r,s) is greatest with $\ell(r,s) \leq t$, by choice of the p_r. Therefore $d_\infty(\boldsymbol{\alpha}, \boldsymbol{\rho}(w^{(t)})) \to 0$ as $t \to \infty$ as required.

It remains to show that w is maximal. Now we can write

$$w = \Lambda_{n_0} \Lambda_{n_1} \cdots \Lambda_{n_R} (\Lambda_{n_{R+1}-1}(k) \Lambda_{n_{R+1}}(u))$$

for some $u \in \Sigma$. However $\Lambda_{n_{R+1}-1}(k) \Lambda_{n_{R+1}}(u) = k\, 1^{n_{R+1}-1} \Lambda_{n_{R+1}}(u)$ is maximal, since it has initial subword $k\, 1^{n_{R+1}-1}$ followed by a letter other than 1, whereas every letter k in $\Lambda_{n_{R+1}}(u)$ is followed by at least n_{R+1} consecutive 1s. Therefore w is also maximal by Lemma 2. $\qquad\square$

Combining Lemmas 19 and 21 gives the result we have been working towards.

THEOREM 22. *Let $\boldsymbol{\alpha} \in \Delta$. Then $\mathcal{I}(\boldsymbol{\alpha}) = S(\Phi(\boldsymbol{\alpha}))$.* $\qquad\square$

REMARKS 23.
a) The proofs of Lemmas 18 and 19 only depend on being able to find arbitrarily long initial subwords W of $w \in \mathcal{R}(\boldsymbol{\alpha})$ with $\boldsymbol{\rho}(W)$ arbitrarily close to $\boldsymbol{\alpha}$. It follows that the results of this section remain true if elements of $\mathcal{R}(\boldsymbol{\alpha})$ are only required to have subsequential limits $\boldsymbol{\alpha}$, which is a common approach in the definition of rotation sets. To be precise, for each $\boldsymbol{\alpha} \in \Delta$ write

$$\mathcal{R}'(\boldsymbol{\alpha}) = \left\{ w \in \Sigma \, : \, \boldsymbol{\rho}\left(w^{(r_i)}\right) \to \boldsymbol{\alpha} \text{ for some } r_i \to \infty \right\} \subset \Sigma,$$

and $\mathcal{M}'(\boldsymbol{\alpha}) = \mathcal{M} \cap \mathcal{R}'(\boldsymbol{\alpha})$. Then $S(\Phi(\boldsymbol{\alpha}))$ is the infimum of $\mathcal{M}'(\boldsymbol{\alpha})$, and $S(\Phi(\boldsymbol{\alpha})) \leq \sup w$ for all $w \in \mathcal{R}'(\boldsymbol{\alpha})$.

b) The infimax sequences $S(\Phi(\boldsymbol{\alpha}))$ are *almost periodic*: for every initial subword W of $S(\Phi(\boldsymbol{\alpha}))$, there is some N with the property that every length N subword of $S(\Phi(\boldsymbol{\alpha}))$ contains W. As a consequence, the orbit closure

$$\Sigma_{\boldsymbol{\alpha}} = \overline{\{\sigma^r(S(\Phi(\boldsymbol{\alpha}))) \, : \, r \geq 0\}}$$

is a minimal σ-invariant set.

To show almost periodicity, assume that $\boldsymbol{\alpha} \notin \mathbb{Q}^k$ (since otherwise $S(\Phi(\boldsymbol{\alpha}))$ is periodic and therefore almost periodic), and write $\mathbf{n} = \Phi(\boldsymbol{\alpha})$. Pick r large enough that $\Lambda_{\mathbf{n},r}(k)$ has initial subword W. Now $\Lambda_{n_{r+1}} \circ \Lambda_{n_{r+2}} \circ \cdots \circ \Lambda_{n_{r+k-1}}(i)$ has initial letter k for all i with $1 \leq i \leq k$ by Lemma 3, so that $U_i := \Lambda_{\mathbf{n},r+k-1}(i)$ has initial subword W for each i. However $S(\Phi(\boldsymbol{\alpha})) = \Lambda_{\mathbf{n},r+k-1}(u)$ for some $u \in \Sigma$, and is therefore a concatenation of the words U_i. This establishes the result, with $N = 2 \max_{1 \leq i \leq k} |U_i|$.

6. Minimax sequences

In this section we address the question of when the infimum $\mathcal{I}(\boldsymbol{\alpha})$ of $\mathcal{M}(\boldsymbol{\alpha})$ is a minimum. Since the set of maximal elements is a closed subset of Σ, $\mathcal{I}(\boldsymbol{\alpha})$ is necessarily maximal, and the issue is whether or not it belongs to $\mathcal{R}(\boldsymbol{\alpha})$. We will show that this happens exactly when $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$. We shall also show that this condition holds for some values of $\boldsymbol{\alpha}$ (in fact we already know by Lemma 10 and Theorem 16 that it holds for $\boldsymbol{\alpha}$ rational), but fails when the itinerary $\Phi(\boldsymbol{\alpha})$ grows too rapidly.

THEOREM 24.   Let $\boldsymbol{\alpha} \in \Delta$. Then
a) $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a d-dimensional simplex for some d with $0 \leq d \leq k - 2$.
b) $\mathcal{I}(\boldsymbol{\alpha})$ is the minimum of $\mathcal{M}(\boldsymbol{\alpha})$ if and only if $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a point.

Proof.   Write $\mathbf{n} = \Phi(\boldsymbol{\alpha})$.
a) The homeomorphisms $K_n^{-1} \colon \Delta \to \Delta_n$ of (2.2) extend by the same formulae to homeomorphisms $K_n^{-1} \colon \overline{\Delta} \to \overline{\Delta}_n \subseteq \overline{\Delta}$ of compact simplices. Define, for each $r \in \mathbb{N}$, an embedding

$$\Upsilon_{\mathbf{n},r} = K_{n_0}^{-1} \circ K_{n_1}^{-1} \circ \cdots \circ K_{n_r}^{-1} \colon \overline{\Delta} \to \overline{\Delta}.$$

The images $A_{\mathbf{n},r} = \Upsilon_{\mathbf{n},r}(\overline{\Delta})$ of these embeddings form a decreasing sequence of non-empty compact subsets of $\overline{\Delta}$, which are $(k-1)$-dimensional simplices since each K_n^{-1} is a projectivity. Moreover $A_{\mathbf{n},r} \subset \Delta$ for all $r \geq k - 1$, since if $\alpha_i > 0$ for some $1 \leq i < k$ then $\Upsilon_{\mathbf{n},k-1-i}(\boldsymbol{\alpha})_{k-1} > 0$, and therefore $\Upsilon_{\mathbf{n},k-i}(\boldsymbol{\alpha})_k > 0$: it follows that

$$\Phi^{-1}(\mathbf{n}) = \bigcap_{r \geq 0} A_{\mathbf{n},r}$$

is a non-empty compact convex subset of Δ, consisting of all those points which have itinerary $\mathbf{n}$: this set is a simplex by a theorem of Borovikov [6], which states that the intersection of a decreasing sequence of simplices is a simplex. Since rational elements of Δ

do not share their itineraries with any other points by Lemma 10 and Theorem 16, $\Phi^{-1}(\mathbf{n})$ cannot contain more than one rational point, and hence has dimension at most $k-2$.

b) If $\boldsymbol{\alpha} \in \mathbb{Q}^k$ then the result follows by Lemma 10 and Theorem 16, so suppose that $\boldsymbol{\alpha} \notin \mathbb{Q}^k$. In particular $n_r > 0$ for arbitrarily large r, and hence $|\Lambda_{\mathbf{n},r}(k)| \to \infty$ as $r \to \infty$. Set

$$\boldsymbol{\alpha}_r^{(i)} = \Upsilon_{\mathbf{n},r}(e^{(i)}) = \boldsymbol{\rho}(\Lambda_{\mathbf{n},r}(i))$$

for each $r \in \mathbb{N}$ and $1 \le i \le k$, where $e^{(i)} = (0,\ldots,0,1,0,\ldots,0)$ is the ith vertex of $\overline{\Delta}$. By compactness and convexity, $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$ if and only if $\boldsymbol{\alpha}_r^{(i)} \to \boldsymbol{\alpha}$ as $r \to \infty$ for each i.

Suppose first then that $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$: we need to show that $S(\mathbf{n}) \in \mathcal{R}(\boldsymbol{\alpha})$.

Let $\epsilon > 0$: we will show that $d_\infty(\boldsymbol{\rho}(S(\mathbf{n})^{(m)}), \boldsymbol{\alpha}) < \epsilon$ for all sufficiently large m. To do this, let $R \ge 0$ be such that $d_\infty(\boldsymbol{\alpha}_R^{(i)}, \boldsymbol{\alpha}) < \epsilon/2$ for all $1 \le i \le k$, and write W_i for the word $\Lambda_{\mathbf{n},R}(i)$: thus $d_\infty(\boldsymbol{\rho}(W_i), \boldsymbol{\alpha}) < \epsilon/2$ for all i. Let $L = \max_{1 \le i \le k} |W_i|$. Now

$$S(\mathbf{n}) = \lim_{r \to \infty} \Lambda_{\mathbf{n},r}(\overline{k}) = \lim_{r \to \infty, r > R} \Lambda_{\mathbf{n},R}(\Lambda_{n_{R+1}} \circ \cdots \circ \Lambda_{n_r}(\overline{k}))$$

is a concatenation of the words W_i. Therefore $d_\infty(\boldsymbol{\rho}(S(\mathbf{n})^{(m)}), \boldsymbol{\alpha}) < \epsilon$ whenever $m > 2L/\epsilon$, as required.

Conversely, suppose that $S(\mathbf{n}) \in \mathcal{R}(\boldsymbol{\alpha})$, so that $\boldsymbol{\alpha}_r^{(k)} \to \boldsymbol{\alpha}$ as $r \to \infty$. We need to show that $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$, or equivalently that $\boldsymbol{\alpha}_r^{(i)} \to \boldsymbol{\alpha}$ as $r \to \infty$ for each i. The proof is by induction on $k \ge 2$, with the case $k = 2$ immediate since then $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$ for all $\boldsymbol{\alpha}$ by a).

We distinguish two cases.

(i) Suppose first that for every i with $1 \le i \le k-1$, there are arbitrarily large integers $r \equiv i - 1 \bmod k - 1$ with the property that $n_r > 0$.

Write $L_r^{(i)} = |\Lambda_{\mathbf{n},r}(i)|$ for each $r \in \mathbb{N}$ and $1 \le i \le k$, so that $\boldsymbol{\alpha}_r^{(i)} L_r^{(i)}$ is an integer vector whose entries give the number of occurences of each letter in $\Lambda_{\mathbf{n},r}(i)$. Comparing the expressions $\Lambda_{\mathbf{n},r}(k-1) = \Lambda_{\mathbf{n},r-1}(k\,1^{n_r+1})$ and $\Lambda_{\mathbf{n},r}(k) = \Lambda_{\mathbf{n},r-1}(k\,1^{n_r})$ gives

$$\boldsymbol{\alpha}_r^{(k-1)} = \frac{\boldsymbol{\alpha}_r^{(k)} L_r^{(k)} + \boldsymbol{\alpha}_{r-1}^{(1)} L_{r-1}^{(1)}}{L_r^{(k)} + L_{r-1}^{(1)}}. \tag{6.1}$$

On the other hand, the first of these two expressions alone gives

$$\boldsymbol{\alpha}_r^{(k-1)} = \frac{\boldsymbol{\alpha}_{r-1}^{(k)} L_{r-1}^{(k)} + (n_r + 1)\boldsymbol{\alpha}_{r-1}^{(1)} L_{r-1}^{(1)}}{L_{r-1}^{(k)} + (n_r + 1)L_{r-1}^{(1)}}. \tag{6.2}$$

Solving (6.1) and (6.2) for $\boldsymbol{\alpha}_r^{(k-1)}$ in terms of $\boldsymbol{\alpha}_r^{(k)}$ and $\boldsymbol{\alpha}_{r-1}^{(k)}$ under the assumption $n_r > 0$ gives

$$\boldsymbol{\alpha}_r^{(k-1)} = \boldsymbol{\alpha}_r^{(k)} + \frac{L_{r-1}^{(k)}}{(n_r + 1)L_r^{(k)} - L_{r-1}^{(k)}}\left(\boldsymbol{\alpha}_r^{(k)} - \boldsymbol{\alpha}_{r-1}^{(k)}\right).$$

Since $\boldsymbol{\alpha}_r^{(k)} \to \boldsymbol{\alpha}$ as $r \to \infty$, it follows that for any $\epsilon > 0$ there is some R such that $d_\infty(\boldsymbol{\alpha}_r^{(k)}, \boldsymbol{\alpha}) < \epsilon$ for all $r \geq R$, and $d_\infty(\boldsymbol{\alpha}_r^{(k-1)}, \boldsymbol{\alpha}) < \epsilon$ for all $r \geq R$ with $n_r > 0$: for $L_r^{(k)} \geq L_{r-1}^{(k)}$ for all r, so that $L_{r-1}^{(k)} \leq (n_r + 1)L_r^{(k)} - L_{r-1}^{(k)}$ provided that $n_r > 0$.

Now suppose that $r \geq R$ with $n_r > 0$. Then the expressions $\Lambda_{\mathbf{n},r}(i) = \Lambda_{\mathbf{n},r-1}(i+1)$ for $1 \leq i \leq k-2$ give $\boldsymbol{\alpha}_{r+k-2}^{(1)} = \boldsymbol{\alpha}_r^{(k-1)}$ so that, by (6.1),

$$\boldsymbol{\alpha}_{r+k-1}^{(k-1)} = \frac{\boldsymbol{\alpha}_{r+k-1}^{(k)}L_{r+k-1}^{(k)} + \boldsymbol{\alpha}_r^{(k-1)}L_r^{(k-1)}}{L_{r+k-1}^{(k)} + L_r^{(k-1)}}$$

and hence $\boldsymbol{\alpha}_{r+k-1}^{(k-1)}$ is a convex combination of points within ϵ of $\boldsymbol{\alpha}$ and so is itself within ϵ of $\boldsymbol{\alpha}$. Inductively it follows that $d_\infty(\boldsymbol{\alpha}_{r+s(k-1)}^{(k-1)}, \boldsymbol{\alpha}) < \epsilon$ for all $s \in \mathbb{N}$. Since, by the defining assumption of this case, there are $r \geq R$ with $n_r > 0$ in every congruence class modulo $k-1$, we have $d_\infty(\boldsymbol{\alpha}_r^{(k-1)}, \boldsymbol{\alpha}) < \epsilon$ for all sufficiently large r. Therefore $\boldsymbol{\alpha}_r^{(k-1)} \to \boldsymbol{\alpha}$ as $r \to \infty$.

Since $\boldsymbol{\alpha}_r^{(i)} = \boldsymbol{\alpha}_{r+i-(k-1)}^{(k-1)}$ for all $1 \leq i \leq k-2$ and $r \geq (k-1)-i$, it follows that $\boldsymbol{\alpha}_r^{(i)} \to \boldsymbol{\alpha}$ as $r \to \infty$ for all i as required.

(ii) Suppose then that there is some i with $1 \leq i \leq k-1$ such that $n_r = 0$ for all sufficiently large $r \equiv i - 1 \bmod k - 1$. We shall show that if $\Phi(\boldsymbol{\beta}) = \mathbf{n}$ then $\boldsymbol{\beta} = \boldsymbol{\alpha}$. Now for each $r \in \mathbb{N}$ we have that $\Phi(\boldsymbol{\beta}) = \mathbf{n}$ if and only if $\boldsymbol{\beta} = K_{n_0}^{-1} \circ K_{n_1}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1}(\boldsymbol{\beta}')$ for some $\boldsymbol{\beta}'$ with $\Phi(\boldsymbol{\beta}') = n_r n_{r+1} \ldots$, so we can suppose without loss of generality that $n_r = 0$ for every $r \equiv 0 \bmod k-1$, and hence by Lemma 7 that $\alpha_1 = \beta_1 = 0$. By Remark 9 (and using the notation introduced there), $\mathbf{m} := \Phi(\pi_1(\boldsymbol{\alpha})) = \Phi(\pi_1(\boldsymbol{\beta}))$ is obtained from $\mathbf{n}$ by deleting the zero entries in positions which are multiples of $k-1$. Write $\boldsymbol{\alpha}_r = \boldsymbol{\alpha}_r^{(k)}$ and $\boldsymbol{\alpha}_r' = \Upsilon_{\mathbf{m},r}(0,0,\ldots,0,1) \in \Delta^{k-1}$ for each $r \in \mathbb{N}$. We shall show that

$$\boldsymbol{\alpha}_{s(k-2)+i}' = \pi_1(\boldsymbol{\alpha}_{s(k-1)+i+1}) \quad \text{whenever } 0 \leq i \leq k-3 \text{ and } s \in \mathbb{N}. \tag{6.3}$$

That is, the proportions of letters in each $\Lambda_{\mathbf{m},r}(k-1)$ (a sequence over $k-1$ letters) is obtained from the proportions of letters in $\Lambda_{\mathbf{n},r'}(k)$ by deleting an initial zero, where r' is an appropriate index which increases with r. This will establish the result. For then $\boldsymbol{\alpha}_r \to \boldsymbol{\alpha}$ implies $\boldsymbol{\alpha}_r' \to \pi_1(\boldsymbol{\alpha})$, or in other words $S(\mathbf{m}) \in \mathcal{R}(\pi_1(\boldsymbol{\alpha}))$. Hence $\pi_1(\boldsymbol{\alpha}) = \pi_1(\boldsymbol{\beta})$ by the inductive hypothesis, so that $\boldsymbol{\alpha} = \boldsymbol{\beta}$ as required.

Observe first that when $s = 0$, equation (6.3) reads

$$K_{m_0}^{-1} \circ K_{m_1}^{-1} \circ \cdots \circ K_{m_i}^{-1}(0,\ldots,0,1) = \pi_1(K_0^{-1} \circ K_{m_0}^{-1} \circ K_{m_1}^{-1} \circ \cdots \circ K_{m_i}^{-1}(0,0,\ldots,0,1))$$

$$\tag{6.4}$$

for $0 \leq i \leq k-3$, where on the left hand side $(0,\ldots,0,1) \in \Delta^{k-1}$, and on the right hand side $(0,0,\ldots,0,1) \in \Delta^k$. This is a straightforward consequence of (2.2): since $i \leq k-3$, the $(k-1)^{\text{th}}$ component of $K_{m_0}^{-1} \circ K_{m_1}^{-1} \circ \cdots \circ K_{m_i}^{-1}(0,0,\ldots,0,1) \in \Delta^k$ is zero, so that applying K_0^{-1} cyclically permutes the first $k-1$ components.

Now it follows from (2.2) that

$$\pi_{i+1}^{-1} \circ K_m^{-1} = K_m^{-1} \circ \pi_i^{-1} : \Delta^{k-1} \to \Delta^{k,i+1}$$

for all $1 \leq i \leq k - 2$ and all $m \in \mathbb{N}$. Applying this for $i = 1, 2, \ldots, k - 2$ in succession gives

$$\pi_{k-1}^{-1} \circ K_{m_{k-2}}^{-1} \circ \cdots \circ K_{m_1}^{-1} = K_{m_{k-2}}^{-1} \circ \cdots \circ K_{m_1}^{-1} \circ \pi_1^{-1} : \Delta^{k-1} \to \Delta^{k,k-1}$$

for all $m_1, \ldots, m_{k-2}$. Then, using again the observation that if $\boldsymbol{\alpha} \in \Delta^{k,k-1}$ then K_0^{-1} cyclically permutes its components,

$$K_{m_{k-2}}^{-1} \circ \cdots \circ K_{m_1}^{-1} = \pi_1 \circ K_0^{-1} \circ K_{m_{k-2}}^{-1} \circ \cdots \circ K_{m_1}^{-1} \circ \pi_1^{-1} : \Delta^{k-1} \to \Delta^{k-1} \qquad (6.5)$$

for all $m_1, \ldots, m_{k-2}$.

Applying (6.4) followed by s applications of (6.5) establishes (6.3) as required.

$\square$

REMARKS 25.

a) It is clear that if $\Phi^{-1}(\mathbf{n})$ is more than just one point, then $S(\mathbf{n})$ can only be the minimum of $\mathcal{M}(\boldsymbol{\alpha})$ for at most one $\boldsymbol{\alpha} \in \Phi^{-1}(\mathbf{n})$. The content of the final part of the proof is that in fact it is not the minimum of any of the sets $\mathcal{M}(\boldsymbol{\alpha})$, and indeed does not belong to $\mathcal{R}(\boldsymbol{\alpha})$ for any $\boldsymbol{\alpha} \in \Delta$.

b) Combining Theorem 24 b) with Lemma 17 of [8] and Lemma 4.2 of [7] yields the following result: the action of the shift map on the orbit closure $\Sigma_{\boldsymbol{\alpha}}$ of Remarks 23 b) is uniquely ergodic if and only if $\boldsymbol{\alpha}$ is regular.

In view of this result, we make the following definitions:

DEFINITIONS 26. $\boldsymbol{\alpha} \in \Delta$ is *regular* if $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$, and *exceptional* otherwise.

When $k = 2$, every $\boldsymbol{\alpha} \in \Delta$ is regular by Theorem 24 a). Therefore, in the two letter case, there is an $\boldsymbol{\alpha}$-minimax sequence for all $\boldsymbol{\alpha}$: these are the well-known Sturmian sequences [14, 25]. When $k \geq 3$, we have already seen that $\boldsymbol{\alpha}$ is regular if it is rational (i.e. if $\Phi(\boldsymbol{\alpha})_r = 0$ for all sufficiently large r). The following theorem states that the same is true when $\Phi(\boldsymbol{\alpha})_r > 0$ grows at most quadratically with r, and, on the other hand, that if $\Phi(\boldsymbol{\alpha})_r$ grows too fast then $\boldsymbol{\alpha}$ is exceptional.

THEOREM 27.   Let $\boldsymbol{\alpha} \in \Delta$ and $\mathbf{n} = \Phi(\boldsymbol{\alpha})$.

a) If there is some C such that $0 < n_r \leq Cr^2$ for all r, then $\boldsymbol{\alpha}$ is regular.

b) If $k \geq 3$ and $n_r \geq 2^{r+2} \prod_{i=0}^{r-1}(n_i + 2)$ for all $r \geq 1$, then $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a simplex of dimension $k - 2$, so that $\boldsymbol{\alpha}$ is exceptional.

Proof.   We use the notation of the proof of Theorem 24.

a) We will use a theorem of Birkhoff [**5**, **9**] to show that Δ is contracted by the embeddings $\Upsilon_{\mathbf{n},r}$, and we start by giving some necessary definitions and stating this theorem. Let $A = (a_{ij})$ be a k by k matrix with strictly positive entries, and f_A be its projective action on $\overline{\Delta}$: that is, $f_A \colon \overline{\Delta} \to \overline{\Delta}$ is defined by

$$f_A(\boldsymbol{\alpha}) = \frac{A\,\boldsymbol{\alpha}}{\|A\,\boldsymbol{\alpha}\|_1}.$$

Define also

$$d(A) = \max_{1 \leq i,j,l,m \leq k} \frac{a_{il}a_{jm}}{a_{im}a_{jl}} \geq 1 \tag{6.6}$$

(that is, $d(A)$ is the largest number that can be obtained by choosing four elements of A arranged in a rectangle, and dividing the product of the two elements on one diagonal by the product of the two elements on the other). $d(A)$ is stricly greater than one unless A has rank 1.

Let $\tau \colon [1, \infty) \to [0, 1)$ be the strictly increasing function $\tau(d) = (\sqrt{d} - 1)/(\sqrt{d} + 1)$. Write $\mathring{\Delta}$ for the simplex Δ less its faces, and let $\delta \colon \mathring{\Delta} \times \mathring{\Delta} \to \mathbb{R}_{\geq 0}$ be Hilbert's projective metric (which generates the Euclidean topology),

$$\delta(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \log \max_{1 \leq i,j \leq k} \frac{\alpha_i \, \beta_j}{\alpha_j \, \beta_i}.$$

Birkhoff's theorem states that, provided $d(A) > 1$, the restriction of f_A to $\mathring{\Delta}$ contracts the metric δ by $\tau(d(A))$: that is, $\delta(f_A(\boldsymbol{\alpha}), f_A(\boldsymbol{\beta})) \leq \tau(d(A))\, \delta(\boldsymbol{\alpha}, \boldsymbol{\beta})$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathring{\Delta}$.

Now let $A(n)$ be the k by k matrix with $A(n)_{1,k-1} = n + 1$, $A(n)_{1,k} = n$, $A(n)_{i,i-1} = 1$ for $2 \leq i \leq k$, $A(n)_{k,k} = 1$, and all other entries zero: as an example, when $k = 5$,

$$A(n) = \begin{pmatrix} 0 & 0 & 0 & n+1 & n \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

By (2.2), we have $f_{A(n)} = K_n^{-1} \colon \overline{\Delta} \to \overline{\Delta}$. Although $A(n)$ has some zero entries, we shall see that any product of $2k - 3$ such matrices $A(n_r)$ with each $n_r > 0$ is strictly positive. Write $A(n_0, \ldots, n_r) = \prod_{s=0}^r A(n_s)$. By considering the action of $\Lambda_{n_0} \circ \cdots \circ \Lambda_{n_{k-3}}$ on each of the letters $1, \ldots, k$, it can be seen that $A(n_0, \ldots, n_{k-3})$ has row i, for $1 \leq i \leq k - 2$, consisting of i zeros followed by $n_{i-1} + 1$ and then n_{i-1} in the other columns; row $k - 1$ has a 1 in column 1 and zeros in the other columns; and row k has a zero in column 1 and 1s in the other columns. Similarly $A(n_{k-2}, \ldots, n_{2k-4})$ has row i, for $1 \leq i \leq k - 1$, consisting of $i - 1$ zeros followed by $n_{k-3+i} + 1$ on the diagonal and n_{k-3+i} in the other columns; while

row k has 1 in every column. As an example, when $k = 5$, these two matrices are given by

$$\begin{pmatrix} 0 & n_0+1 & n_0 & n_0 & n_0 \\ 0 & 0 & n_1+1 & n_1 & n_1 \\ 0 & 0 & 0 & n_2+1 & n_2 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} n_3+1 & n_3 & n_3 & n_3 & n_3 \\ 0 & n_4+1 & n_4 & n_4 & n_4 \\ 0 & 0 & n_5+1 & n_5 & n_5 \\ 0 & 0 & 0 & n_6+1 & n_6 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

The product $A(n_0, \ldots, n_{2k-4}) = (a_{ij})_{1 \leq i,j \leq k}$ of these matrices is therefore strictly positive when each $n_r > 0$, with each a_{ij} a polynomial of degree at most 2 in $n_0, \ldots, n_{2k-4}$. We shall show that, for each $1 \leq i \leq k$ and each $1 \leq l < m \leq k$, the quotient a_{il}/a_{im} is bounded above by 2, while the quotient a_{im}/a_{il} is bounded above by a linear function of $n_{k-1}, \ldots, n_{2k-4}$. As a consequence, since (6.6) says that $d(A)$ is the product of one quotient of the first type and one of the second, there is some R, depending only on k, such that

$$d(A(n_0, \ldots, n_{2k-4})) \leq R(n_{k-1} + \cdots + n_{2k-4}) \tag{6.7}$$

provided that each $n_r > 0$.

The claim is straightforward when $i = k-1$, in which case a_{il} is either n_{k-2} or $n_{k-2}+1$; and when $i = k$, in which case $a_{i1} = 1$, $a_{il} = 2 + \sum_{j=k-1}^{k-3+l} n_j$ for $2 \leq l \leq k-1$, and $a_{ik} = a_{i,k-1} - 1$. When $1 \leq i \leq k-2$, the explicit descriptions of the elements of $A(n_0, \ldots, n_{k-3})$ and $A(n_{k-2}, \ldots, n_{2k-4})$ give

$$a_{il} = \begin{cases} n_{i-1} & \text{if } 1 \leq l \leq i, \\ n_{i-1}(n_{k+i-2}+2) + (n_{k+i-2}+1) & \text{if } l = i+1, \\ n_{i-1}\left(2 + \sum_{j=k+i-2}^{k+l-3} n_j\right) + n_{k+i-2} & \text{if } i+2 \leq l < k, \\ n_{i-1}\left(1 + \sum_{j=k+i-2}^{2k-4} n_j\right) + n_{k+i-2} & \text{if } l = k, \end{cases}$$

from which the claim follows.

Now let $\boldsymbol{\alpha} \in \Delta$, and suppose that there is some C such that $\mathbf{n} = \Phi(\boldsymbol{\alpha})$ satisfies $0 < n_r \leq Cr^2$ for all r. For each $r \geq 0$ we have

$$\Upsilon_{\mathbf{n},(r+1)(2k-3)-1}(\overline{\Delta}) = (K_{n_0}^{-1} \circ \cdots \circ K_{n_{2k-4}}^{-1}) \circ (K_{n_{2k-3}}^{-1} \circ \cdots \circ K_{n_{4k-7}}^{-1}) \circ \cdots \circ$$
$$(K_{n_{r(2k-3)}}^{-1} \circ \cdots \circ K_{n_{(r+1)(2k-3)-1}}^{-1})(\overline{\Delta})$$

Since $(K_{n_{r(2k-3)}}^{-1} \circ \cdots \circ K_{n_{(r+1)(2k-3)-1}}^{-1})(\overline{\Delta}) \subset \mathring{\Delta}$ (because the product of $2k-3$ matrices $A(n)$ is strictly positive), it is enough to show that

$$\prod_{r=0}^{\infty} \tau(d(A(n_{r(2k-3)}, \ldots, n_{(r+1)(2k-3)-1}))) = 0.$$

By (6.7) and $n_r \leq Cr^2$, there is some Q depending only on C and k such that $d_r := d(A(n_{r(2k-3)}, \ldots, n_{(r+1)(2k-3)-1})) \leq (Qr)^2$ for all $r \geq 1$, so that $\tau(d_r) \leq (Qr-1)/(Qr+1)$. Recall that if $0 < a_r \leq 1$ for all r then $\prod_{r=0}^{\infty} a_r = 0$ if and only if $\sum_{r=0}^{\infty} \left(\frac{1}{a_r} - 1\right)$ diverges.

Since

$$\frac{1}{\tau(d_r)} - 1 \geq \frac{2}{Qr - 1},$$

the result follows.

b) Set

$$\delta_r = \min_{1 \leq i < j \leq k-1} d_\infty(\boldsymbol{\alpha}_r^{(i)}, \boldsymbol{\alpha}_r^{(j)})$$

for each $r \geq 0$, the smallest distance between a pair of vertices in the simplex $A_{\mathbf{n},r}$ excluding the vertex $\Upsilon_{\mathbf{n},r}(0,0,\ldots,0,1)$. We shall show that $\delta_0 = 1$ and $\delta_r \geq \delta_{r-1} - 1/2^{r+2}$ for each $r \geq 1$, so that $\delta_r > 3/4$ for all r. It is therefore not possible for all of the $\boldsymbol{\alpha}_r^{(j)}$ to converge to the same point.

That $\delta_0 = 1$ is straightforward, since $\boldsymbol{\alpha}_0^{(i)} = K_{n_0}^{-1}(e^{(i)})$ is equal to $e^{(i+1)}$ if $1 \leq i \leq k - 2$, and to $((n_0 + 1)e^{(1)} + e^{(k)})/(n_0 + 2)$ if $i = k - 1$.

Now let $r \geq 1$. If $1 \leq i \leq k - 2$ then we have $\Lambda_{\mathbf{n},r}(i) = \Lambda_{\mathbf{n},r-1}(\Lambda_{n_r}(i)) = \Lambda_{\mathbf{n},r-1}(i+1)$, so that

$$\boldsymbol{\alpha}_r^{(i)} = \boldsymbol{\alpha}_{r-1}^{(i+1)} \qquad \text{for } 1 \leq i \leq k - 2.$$

Consider then the case $i = k - 1$. By (6.2)

$$\boldsymbol{\alpha}_r^{(k-1)} = \frac{\boldsymbol{\alpha}_{r-1}^{(k)} L_{r-1}^{(k)} + (n_r + 1)\boldsymbol{\alpha}_{r-1}^{(1)} L_{r-1}^{(1)}}{L_{r-1}^{(k)} + (n_r + 1)L_{r-1}^{(1)}},$$

so that

$$\boldsymbol{\alpha}_r^{(k-1)} - \boldsymbol{\alpha}_{r-1}^{(1)} = \frac{L_{r-1}^{(k)} \left(\boldsymbol{\alpha}_{r-1}^{(k)} - \boldsymbol{\alpha}_{r-1}^{(1)} \right)}{(n_r + 1)L_{r-1}^{(1)} + L_{r-1}^{(k)}},$$

in which each component has absolute value bounded above by $\prod_{i=0}^{r-1}(n_i + 2)/n_r \leq 1/2^{r+2}$, using $L_{r-1}^{(k)} \leq \prod_{i=0}^{r-1}(n_i + 2)$ in the numerator and $L_{r-1}^{(i)} \geq 1$ in the denominator.

Therefore $d_\infty(\boldsymbol{\alpha}_r^{(k-1)}, \boldsymbol{\alpha}_{r-1}^{(1)}) \leq 1/2^{r+2}$, and we saw in the first part of the proof that $d_\infty(\boldsymbol{\alpha}_r^{(i)}, \boldsymbol{\alpha}_{r-1}^{(i+1)}) = 0$ for $1 \leq i \leq k - 2$. This gives $\delta_r \geq \delta_{r-1} - 1/2^{r+2}$ as required.

To show that $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$ is a simplex of dimension $k - 2$, let $\pi \colon \mathbb{R}^k \to \mathbb{R}^{k-1}$ be projection onto the first $k - 1$ coordinates. Then

$$V_0 := \{\pi(\boldsymbol{\alpha}_0^{(i)}) : 1 \leq i \leq k - 1\} = \{(n_0 + 1)\pi(e^{(1)})/(n_0 + 2), \pi(e^{(2)}), \pi(e^{(3)}), \ldots, \pi(e^{(k-1)})\},$$

and $(n_0 + 1)/(n_0 + 2) \geq 1/2$. Now for each $r \geq 1$, the set $V_r := \{\pi(\boldsymbol{\alpha}_r^{(i)}) : 1 \leq i \leq k - 1\}$ is within d_∞-Hausdorff distance $1/4$ of V_0, and hence the same is true for the limit V_∞. The $k - 1$ points of V_∞ therefore span a simplex of dimension $k - 2$, which is the π-image of a simplex of dimension $k - 2$ contained in $\Phi^{-1}(\Phi(\boldsymbol{\alpha}))$.

$\square$

EXAMPLE 28.   The conditions of Theorem 27a) are obviously satisfied when $\mathbf{n} = \Phi(\boldsymbol{\alpha}) = \overline{n_0 \dots n_{r-1}}$ is periodic without any zero entries: by the theorem, such a sequence is the itinerary of a unique periodic point of K. The corresponding minimax sequence $\mathcal{I}(\boldsymbol{\alpha})$ is the fixed point of the substitution $\Lambda_{n_0} \circ \cdots \circ \Lambda_{n_{r-1}}$, and therefore generates a substitution minimal set [**15**].

The simplest such example is when $k = 3$ and $\Phi(\boldsymbol{\alpha}) = \overline{1}$. The minimum of $\mathcal{M}(\boldsymbol{\alpha})$ is then given by

$$\lim_{r \to \infty} \Lambda_1^r(3) = 312311312231231131131231131223122312311312231231131131231131312\dots,$$

the unique fixed point of Λ_1. In this example $\boldsymbol{\alpha}$ is the unique fixed point of K_1 or, equivalently, the (suitably normalized) strictly positive eigenvector of the matrix

$$A(1) = \begin{pmatrix} 0 & 2 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

from the proof of Theorem 27a). Notice that the minimum of $\mathcal{M}(\boldsymbol{\alpha})$ is not of Arnoux-Rauzy type [**4**]: for example, it has six factors of length two, and the substitution Λ_1 is not Pisot.

REMARKS 29.
a) Theorems 11 and 12 of [**8**] improve substantially on Theorem 27 in the case $k = 3$ (only). After translation to the notation used here, they read:
THEOREM (BRUIN AND TROUBETZKOY)   Let $k = 3$, and let $\boldsymbol{\alpha} \in \Delta$ and $\mathbf{n} = \Phi(\boldsymbol{\alpha})$.
  – For each $r \geq 0$, let $L_{2r} = \min\{s \geq 1 : n_{2r+s} \neq 0\}$. If either

$$\sum_r \frac{n_{2r}}{n_{2r}+1} \sqrt{\frac{1}{(n_{2r-1}+1)L_{2r}}} = \infty, \quad \text{or}$$

$$\prod_r \frac{n_{2r}+1}{n_{2r-1}+1+\frac{1}{L_{2r}}} = 0,$$

   or if either condition holds for the shift $\sigma(\mathbf{n}) = \Phi(K(\boldsymbol{\alpha}))$ of $\mathbf{n}$, then $\boldsymbol{\alpha}$ is regular.
  – If there is some $\lambda > 1$ such that $n_{r+1} \geq \lambda n_r$ for all sufficiently large r, then $\boldsymbol{\alpha}$ is exceptional.
   This result gives rise to a striking pair of examples: on the one hand, if $\Phi(\boldsymbol{\alpha})_r = 2^r$ for all r then $\boldsymbol{\alpha}$ is exceptional by the second statement; while on the other hand, if $\Phi(\boldsymbol{\alpha})_r = 2^r$ when r is even and $\Phi(\boldsymbol{\alpha})_r = 3^r$ when r is odd, then $\boldsymbol{\alpha}$ is regular by the second condition in the former statement.
b) The result of Theorem 27a) clearly extends to the case where finitely many of the n_r are zero. When $n_r = 0$ for arbitrarily large r the situation is more complicated, as the product of $2k - 3$ successive matrices need not be strictly positive. This can not always be remedied by grouping the sequence of matrices more judiciously: in the case where $n_{r(k-1)} = 0$ for all r, no product $A(n_s, n_{s+1}, \dots, n_{s+t})$ is strictly positive. This case arises when considering the itinerary of an element $\boldsymbol{\alpha}$ of Δ which has some zero coordinates (Lemma 7), and can be treated by induction on k.
c) The fact that the bound of (6.7) depends only on $k - 2$ of the $2k - 3$ variables means that it is sufficient for regularity to have control over the n_r along an appropriate subsequence.

d) The growth condition in Theorem 27b) — which, for example, is satisfied by $n_r = 2^{2^{3r}}$ — could easily be improved by improving the bounds on $L_{r-1}^{(k)}$ and $L_{r-1}^{(i)}$ in the penultimate paragraph of the proof: the point here is simply to show that exceptional $\boldsymbol{\alpha}$ exist. In fact, numerical experiments suggest that, when $k = 3$, $\Phi^{-1}(\mathbf{n})$ is a non-trivial interval when $n_r = r^3$, so that even the $k = 3$ results of Bruin and Troubetzkoy are far from optimal.

We finish by showing – closely following the proof of Corollary 13 of [**8**] – that a generic element $\mathbf{n}$ of $\mathbb{N}^{\mathbb{N}}$ is the itinerary of only one point. We use the following lemma.

LEMMA 30.   *For all $n \geq 0$, the map $K_n^{-1} \colon \mathring{\Delta} \to \mathring{\Delta}$ does not expand the Hilbert metric: that is, $\delta(K_n^{-1}(\boldsymbol{\alpha}), K_n^{-1}(\boldsymbol{\beta})) \leq \delta(\boldsymbol{\alpha}, \boldsymbol{\beta})$ for all $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathring{\Delta}$.*

Proof.   Let $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathring{\Delta}$, and write

$$\boldsymbol{\alpha}' := K_n^{-1}(\boldsymbol{\alpha}) = C\left((n+1)\alpha_{k-1} + n\alpha_k,\ \alpha_1,\ \alpha_2,\ \ldots,\ \alpha_{k-2},\ \alpha_{k-1} + \alpha_k\right),$$

where $C = C(\boldsymbol{\alpha})$ is a constant, and similarly $\boldsymbol{\beta}' := K_n^{-1}(\boldsymbol{\beta})$. To prove the lemma, we need to show that whenever $1 \leq i < j \leq k$, there exist I and J between 1 and k with

$$\frac{\alpha_i' \, \beta_j'}{\beta_i' \, \alpha_j'} \leq \frac{\alpha_I \, \beta_J}{\beta_I \, \alpha_J}.$$

This can be established straightforwardly by cases, using the elementary fact that if a, b, c, d are positive reals then $(a+b)/(c+d)$ lies between a/c and b/d.

– If i and j are both between 2 and $k-1$ then $\dfrac{\alpha_i'\beta_j'}{\beta_i'\alpha_j'} = \dfrac{\alpha_{i-1}\beta_{j-1}}{\beta_{i-1}\alpha_{j-1}}$.

– If $i = 1$ and $j < k$ then $\dfrac{((n+1)\alpha_{k-1} + n\alpha_k)\beta_{j-1}}{((n+1)\beta_{k-1} + n\beta_k)\alpha_{j-1}}$ lies between $\dfrac{(n+1)\alpha_{k-1}\beta_{j-1}}{(n+1)\beta_{k-1}\alpha_{j-1}} = \dfrac{\alpha_{k-1}\beta_{j-1}}{\beta_{k-1}\alpha_{j-1}}$

and $\dfrac{n\alpha_k\beta_{j-1}}{n\beta_k\alpha_{j-1}} = \dfrac{\alpha_k\beta_{j-1}}{\beta_k\alpha_{j-1}}$.

– If $i > 1$ and $j = k$ then the argument is identical, except that the factors $n+1$ and n are omitted.

– If $i = 1$ and $j = k$ then $\dfrac{((n+1)\alpha_{k-1} + n\alpha_k)(\beta_{k-1} + \beta_k)}{((n+1)\beta_{k-1} + n\beta_k)(\alpha_{k-1} + \alpha_k)}$ lies between $\dfrac{((n+1)\alpha_{k-1} + n\alpha_k)\beta_{k-1}}{((n+1)\beta_{k-1} + n\beta_k)\alpha_{k-1}}$

and $\dfrac{((n+1)\alpha_{k-1} + n\alpha_k)\beta_k}{((n+1)\beta_{k-1} + n\beta_k)\alpha_k}$, each of which is between two terms of the required type by the argument above.

$\square$

Let $\mathcal{O} \subset \mathbb{N}^{\mathbb{N}}$ be the set of itineraries $\mathbf{n}$ which contain infinitely many disjoint subwords 1^{2k-3}, and let Reg $\subset \mathbb{N}^{\mathbb{N}}$ be the set of regular itineraries $\mathbf{n}$, i.e. those for which $\Phi^{-1}(\mathbf{n})$ is a point.

THEOREM 31. $\mathcal{O} \subset \mathrm{Reg}$, and $\mathcal{O}$ is a dense G_δ subset of $\mathbb{N}^{\mathbb{N}}$.

Proof.   Let $\mathbf{n} \in \mathcal{O}$. As shown in the proof of Theorem 27, the map $K_1^{-(2k-3)} \colon \mathring{\Delta} \to \mathring{\Delta}$ is represented by the strictly positive matrix $A(1)^{2k-3}$, and hence by Birkhoff's theorem contracts the Hilbert metric by a factor $\lambda \in (0, 1)$. It follows, using Lemma 30, that if r_i is the index of the start of the ith disjoint subword 1^{2k-3} in $\mathbf{n}$, then we have

$$A_{\mathbf{n},r_i+(2k-3)} = \Upsilon_{\mathbf{n},r_i+(2k-3)}(\overline{\Delta}) = \Upsilon_{\mathbf{n},r_i}\left(K_1^{-(2k-3)}(\overline{\Delta})\right)$$

has diameter bounded above by $\lambda^{i-1}D$, where D is the Hilbert diameter of $K_1^{-(2k-3)}(\overline{\Delta}) \subset \mathring{\Delta}$. Therefore $\Phi^{-1}(\mathbf{n})$ is a single point, so that $\mathbf{n} \in \mathrm{Reg}$.

For each $N \geq 0$ let $\mathcal{O}_N \subset \mathbb{N}^{\mathbb{N}}$ be the set of itineraries which contain a word 1^{2k-3} starting after the Nth symbol. Then $\mathcal{O}_N$ is open and dense in the Baire space $\mathbb{N}^{\mathbb{N}}$, so that $\mathcal{O} = \bigcap_{N \geq 0} \mathcal{O}_N$ is a dense G_δ subset of $\mathbb{N}^{\mathbb{N}}$ as required. $\square$

References

1. P. Arnoux. Sturmian sequences. In *Substitutions in dynamics, arithmetics and combinatorics*, volume 1794 of *Lecture Notes in Math.*, pages 143–198. Springer, Berlin, 2002.

2. P. Arnoux and G. Rauzy. Représentation géométrique de suites de complexité $2n + 1$. *Bull. Soc. Math. France*, 119(2):199–215, 1991.

3. J. Berstel and D. Perrin. The origins of combinatorics on words. *European J. Combin.*, 28(3):996–1022, 2007.

4. V. Berthé, S. Ferenczi, and L. Zamboni. Interactions between dynamics, arithmetics and combinatorics: the good, the bad, and the ugly. In *Algebraic and topological dynamics*, volume 385 of *Contemp. Math.*, pages 333–364. Amer. Math. Soc., Providence, RI, 2005.

5. G. Birkhoff. Extensions of Jentzsch's theorem. *Trans. Amer. Math. Soc.*, 85:219–227, 1957.

6. V. Borovikov. On the intersection of a sequence of simplexes. *Uspehi Matem. Nauk (N.S.)*, 7(6(52)):179–180, 1952.

7. H. Bruin. Renormalization in a class of interval translation maps of d branches. *Dyn. Syst.*, 22(1):11–24, 2007.

8. H. Bruin and S. Troubetzkoy. The Gauss map on a class of interval translation mappings. *Israel J. Math.*, 137:125–148, 2003.

9. J. E. Carroll. Birkhoff's contraction coefficient. *Linear Algebra Appl.*, 389:227–234, 2004.

10. J. Cassaigne and F. Nicolas. Factor complexity. In *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.*, pages 163–247. Cambridge Univ. Press, Cambridge, 2010.

11. W. de Melo and S. van Strien. *One-dimensional dynamics*, volume 25 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1993.

12. S. Ferenczi, C. Holton, and L. Zamboni. Structure of three-interval exchange transformations. II. A combinatorial description of the trajectories. *J. Anal. Math.*, 89:239–276, 2003.

13. S. Ferenczi and L. Zamboni. Clustering words and interval exchanges. *J. Integer Seq.*, 16(2):Article 13.2.1, 9, 2013.

14. J-M. Gambaudo, O. Lanford, III, and C. Tresser. Dynamique symbolique des rotations. *C. R. Acad. Sci. Paris Sér. I Math.*, 299(16):823–826, 1984.

15. W. Gottschalk. Substitution minimal sets. *Trans. Amer. Math. Soc.*, 109:467–491, 1963.

16. G. A. Hedlund. Sturmian minimal sets. *Amer. J. Math.*, 66:605–620, 1944.

17. O. Jenkinson. A partial order on $\times 2$-invariant measures. *Math. Res. Lett.*, 15(5):893–900, 2008.

18. O. Jenkinson. Balanced words and majorization. *Discrete Math. Algorithms Appl.*, 1(4):463–483, 2009.

19. L-M. Lopez and P. Narbel. Substitutions and interval exchange transformations of rotation class. *Theoret. Comput. Sci.*, 255(1-2):323–344, 2001.

20. R. C. Lyndon. On Burnside's problem. *Trans. Amer. Math. Soc.*, 77:202–215, 1954.

21. M. Morse and G. A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.*, 62:1–42, 1940.

22. W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, 11:401–416, 1960.

23. G. Rauzy. Échanges d'intervalles et transformations induites. *Acta Arith.*, 34(4):315–328, 1979.

24. F. Schweiger. *Multidimensional continued fractions*. Oxford Science Publications. Oxford University Press, Oxford, 2000.

25. P. Veerman. Symbolic dynamics of order-preserving orbits. *Phys. D*, 29(1-2):191–201, 1987.

26. K. Ziemian. Rotation sets for subshifts of finite type. *Fund. Math.*, 146(2):189–201, 1995.