# ON DIGIT FREQUENCIES IN $\beta$-EXPANSIONS

PHILIP BOYLAND, ANDRÉ DE CARVALHO, AND TOBY HALL

ABSTRACT. We study the sets $\mathrm{DF}(\beta)$ of digit frequencies of $\beta$-expansions of numbers in $[0,1]$. We show that $\mathrm{DF}(\beta)$ is a compact convex set with countably many extreme points which varies continuously with $\beta$; that there is a full measure collection of non-trivial closed intervals on each of which $\mathrm{DF}(\beta)$ mode locks to a constant polytope with rational vertices; and that the generic digit frequency set has infinitely many extreme points, accumulating on a single non-rational extreme point whose components are rationally independent.

## 1. Introduction

1.1. **$\beta$-expansions.** Let $\beta > 1$ be a real number and write $k = \lceil \beta \rceil$, the smallest integer which is not less than $\beta$. A *$\beta$-expansion* [16] of a number $x \in [0,1]$ is any representation of $x$ of the form

$$x = \sum_{r=0}^{\infty} w_r \beta^{-(r+1)},$$

in which the sequence $w = (w_r)_{r\geq 0}$ of digits belongs to $\Sigma_k = \{0, 1, \dots, k-1\}^{\mathbb{N}}$.

In general, a number $x$ can have many distinct $\beta$-expansions [10, 18]. However there is a canonical choice, known as the *greedy $\beta$-expansion*, or simply as *the* $\beta$-expansion, for which the sequence $w \in \Sigma_k$ of digits is lexicographically greatest. By analogy with the usual algorithm for determining expansions to integer bases, it is found by choosing each digit in turn to be as large as possible. To be precise, except in the trivial case where $\beta$ is an integer and $x = 1$, if $f_\beta \colon [0,1] \to [0,1)$ is defined by $f_\beta(x) = \beta x \bmod 1$, then the sequence $d_\beta(x) \in \Sigma_k$ of digits of the greedy $\beta$-expansion of $x$ is given by $d_\beta(x)_r = \lfloor \beta f_\beta^r(x) \rfloor$, the integer part of $\beta f_\beta^r(x)$. Equivalently, $d_\beta(x)$ is the itinerary of $x$ under $f_\beta$ with respect to the intervals $I_j$ $(0 \leq j \leq k-1)$ defined by $I_j = [j/\beta, (j+1)/\beta)$ for $0 \leq j < k-1$, and $I_{k-1} = [(k-1)/\beta, 1]$: that is, $d_\beta(x)_r = j$ if and only if $f_\beta^r(x) \in I_j$.

1.2. **The digit frequency set.** Let

$$\Delta = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^k_{\geq 0} \; : \; \sum_{i=0}^{k-1} \alpha_i = 1 \right\}$$

be the standard $(k-1)$-simplex. Given $\boldsymbol{\alpha} \in \Delta$, we say that a number $x \in [0,1]$ *has $\beta$-digit frequency $\boldsymbol{\alpha}$*, and write $\delta_\beta(x) = \boldsymbol{\alpha}$, if $\lim_{r\to\infty} N_{i,r}(d_\beta(x))/r = \alpha_i$ for each $i$, where $N_{i,r}(d_\beta(x))$ denotes the number of $i^{\mathrm{s}}$ in the first $r$ entries of $d_\beta(x)$. Parry [14] observed that since the digit frequency can be written as a Birkhoff sum, it exists and is constant for almost every $x \in [0,1]$ with respect to any ergodic invariant measure, such as the measure of maximal entropy which he himself defined.

---

*Date*: October 2014.

2010 *Mathematics Subject Classification.* Primary 11A63, Secondary 37B10, 68R15.

The authors would like to thank the referee, whose careful reading of the paper led to many significant improvements. We are grateful for the support of FAPESP grants 2010/09667-0 and 2011/17581-0. This research has also been supported in part by EU Marie-Curie IRSES Brazilian-European partnership in Dynamical Systems (FP7-PEOPLE-2012-IRSES 318999 BREUDS).

In this paper we study the sets

$$\mathrm{DF}(\beta) = \{\delta_\beta(x) \,:\, x \in [0,1] \text{ has well-defined } \beta\text{-digit frequency}\}$$

of *all* digit frequencies of $\beta$-expansions of numbers in $[0,1]$. For example, Figure 1 depicts $\mathrm{DF}(\beta)$ for $\beta = 2.1901$ (and indeed for all $\beta$ in a neighbourhood of this value), projected into the $(\alpha_0, \alpha_2)$-plane, and shown within the 2-simplex $\Delta$. It is a pentagon, with vertices $(0,1,0)$, $(1,0,0)$, $(3/4,0,1/4)$, $(5/8,1/8,2/8)$, and $(4/9,3/9,2/9)$ (see Examples 36a), 48, and 60). So, for example, a $\beta$-expansion with this $\beta$ can have at most $1/4$ of its digits equal to 2; and if it has this many $2^{\mathrm{s}}$, then at most $1/8$ of its digits can be equal to 1.

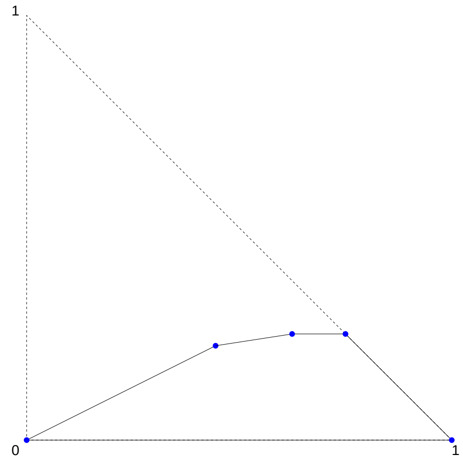


FIGURE 1. $\mathrm{DF}(\beta)$ for $\beta$ in a neighbourhood of 2.1901 (so $k = 3$), projected into the $(\alpha_0, \alpha_2)$-plane

1.3. **Bifurcations of the digit frequency set.** The case where $1 < \beta \le 2$ (i.e., when $k = 2$) was studied by Chi and Kwon [8] and by Kwon [12]. In this case, the digit frequency set is a compact interval contained in the line $\alpha_0 + \alpha_1 = 1$,

$$\mathrm{DF}(\beta) = \{(1-\alpha_1, \alpha_1) \,:\, \alpha_1 \in [0, \mathrm{rhe}(\beta)]\}.$$

As $\beta$ increases through $(1,2]$, the right hand endpoint $\mathrm{rhe}(\beta)$ varies as a devil's staircase (Figure 8). That is, for each rational $m/n \in (0,1)$ there is a non-trivial interval $I_{m/n} \subset (1,2)$ with $\mathrm{rhe}(\beta) = m/n$ for all $\beta \in I_{m/n}$ ("mode-locking"). On the other hand, each irrational in $(0,1)$ is equal to $\mathrm{rhe}(\beta)$ for only one $\beta$; such "irrational" values of $\beta$ are buried points of a Cantor set whose complementary gaps are the interiors of the mode-locking intervals. Bifurcations in the digit frequency set take place at those values of $\beta$ for which the itinerary of 1 under $f_\beta$ is a Sturmian sequence $s_\alpha$ for some $\alpha \in (0,1)$: this is the smallest value of $\beta$ for which the frequency $(1-\alpha, \alpha)$ belongs to $\mathrm{DF}(\beta)$. These results can also be derived from the theory of rotation intervals of bimodal degree one circle maps.

When $\beta \in (k-1, k]$ for $k \ge 3$ the situation, while somewhat analogous, is more complicated. Now $\mathrm{DF}(\beta)$ behaves as a convex set-valued devil's staircase, mode-locking to a polytope with rational vertices on each of an infinite collection of closed subintervals of $(k-1, k]$. "Irrational" behaviour again occurs when $\beta$ is a buried point of the Cantor set whose complementary gaps are the interiors of the mode-locking intervals. For such values of $\beta$, the digit frequency set $\mathrm{DF}(\beta)$ has a countably infinite set of rational vertices, which limits on a finite number of non-rational extreme points, of which there are at least 1 and at most $k-1$ (see Figures 4 and 5). $\beta$ is said to be *regular* when there is only one non-rational extreme point (or when there are none, i.e. when $\mathrm{DF}(\beta)$ is a polytope); and is said to be *exceptional* otherwise.

As in the case $k = 2$, bifurcations of the digit frequency set occur when the itinerary of 1 under $f_\beta$ passes through elements of a particular set of sequences: these are the *lexicographic infimax sequences* of [6], which are described in Section 2. In the regular case the infimax sequence has a well-defined digit frequency, which becomes an element of the digit frequency set. In the exceptional case, on the other hand, the convex hull of the set of non-rational extreme points is a "prime" collection of frequencies which is either contained in or disjoint from $\mathrm{DF}(\beta)$ for every $\beta \in (k-1, k]$ (see Remark 46).

The topologically generic behaviour, with respect to the parameter $\beta$, is therefore that $\mathrm{DF}(\beta)$ is a polytope with rational vertices. We shall also show (Theorem 49) that $\mathrm{DF}(\beta)$ is a polytope with rational vertices for Lebesgue almost every $\beta$ in $(k-1, k]$. This contrasts, however, with the generic behaviour in the collection of all digit frequency sets with the Hausdorff topology. This space is homeomorphic to an interval, and we show (Theorem 51) that in this interval, the generic digit frequency set is of non-rational regular type, with its single limiting extreme point having components that are independent over the rationals.

1.4. **Summary of results.** The following list summarises the main results of the paper, which are contained in Theorems 45, 49, and 51. The statements are for $\beta \in (k-1, k]$, where $k \geq 3$: the simpler situation when $\beta \in (1, 2]$ is discussed in Example 47.

- $\mathrm{DF}(\beta)$ is a compact convex set of dimension $k-1$.
- The function $\beta \mapsto \mathrm{DF}(\beta)$ is increasing, and is continuous with respect to the Hausdorff topology on the set of non-empty compact subsets of $\Delta$.
- $\mathrm{DF}(\beta)$ has countably many extreme points, and there is an algorithm which lists them.
- All but at most $k-1$ extreme points of $\mathrm{DF}(\beta)$ are rational. There exist $\beta$ for which the set of extreme points accumulates on $k-1$ non-rational points.
- There are infinitely many disjoint closed intervals, whose union has full Lebesgue measure in $(k-1, k]$, on each of which $\mathrm{DF}(\beta)$ mode locks to a constant polytope with rational vertices.
- Digit frequency sets which are not polytopes are realised by only one value of $\beta$.
- The set $\mathcal{D} := \{\mathrm{DF}(\beta) : \beta \in [k-1, k]\}$ with the Hausdorff topology is homeomorphic to an interval. There is a dense $G_\delta$ subset of $\mathcal{D}$ consisting of digit frequency sets having a single non-rational extreme point, whose components are rationally independent.

1.5. **Outline of the paper.** Let $Z_\beta = \{d_\beta(x) : x \in [0, 1]\}$ be the set of all digit sequences of greedy $\beta$-expansions, or equivalently the set of all itineraries of orbits of $f_\beta$. The set $Z_\beta$ is determined by the "kneading sequence" of $f_\beta$, the itinerary of the rightmost point 1, once a minor correction has been made to account for the ambiguity of coding at the endpoints of the intervals $I_j$ – this is the same ambiguity which arises in expansions to integer bases. Let $w_\beta = \lim_{x \nearrow 1} d_\beta(x) \in \Sigma_k$, the *quasi-greedy* expansion of 1. Then (see for example [13] Theorem 7.2.9)

$$Z_\beta = \{v \in \Sigma_k : \sigma^r(v) < w_\beta \text{ for all } r \in \mathbb{N}\} \cup \{d_\beta(1)\},$$

where $<$ is the lexicographic order on $\Sigma_k$ and $\sigma \colon \Sigma_k \to \Sigma_k$ is the shift map. The quasi-greedy expansion $w_\beta$ is equal to $d_\beta(1)$ except for those values of $\beta$ for which $f_\beta^r(1) = 0$ for some $r \geq 1$, when $w_\beta$ is periodic.

The problem of determining $\mathrm{DF}(\beta)$ can therefore be rephrased as follows: for which $\boldsymbol{\alpha} \in \Delta$ is there some $v \in \Sigma_k$ with digit frequency $\boldsymbol{\alpha}$, whose entire $\sigma$-orbit is less than $w_\beta$? Techniques for answering this type of question were developed in [6]. The results of that paper will be summarised and extended in Section 2, and applied in Section 3 to the closely related problem of describing digit frequency sets of *symbolic $\beta$-shifts*

$$X(w) = \{v \in \Sigma_k \,:\, \sigma^r(v) \le w \text{ for all } r \in \mathbb{N}\},$$

where $w \in \Sigma_k$. In Section 4 we interpret these results in terms of digit frequency sets of $\beta$-expansions; describe typical digit frequency sets from both the measure-theoretic and topological points of view; investigate the smoothness of digit frequency sets at non-rational extreme points; discuss the set of accumulation points of the sequences $(N_{i,r}(w_\beta))_{r\ge 0}$; and present some examples illustrating how the digit frequency set $\mathrm{DF}(\beta)$ can be calculated (or approximated in the non-polytope case) in practice for a specific value of $\beta$.

1.6. **The subsequential approach.** Digit frequencies could alternatively be defined subsequentially. Let

$$\delta'_\beta(x) = \{\boldsymbol{\alpha} \in \Delta \,:\, \lim_{s\to\infty} N_{i,r_s}(d_\beta(x))/r_s = \alpha_i \text{ for some } r_s \to \infty \text{ and each } i\}$$

for each $x \in [0,1]$, and set

$$\mathrm{DF}'(\beta) = \bigcup_{x\in[0,1]} \delta'_\beta(x).$$

A priori $\mathrm{DF}'(\beta)$ is a bigger set than $\mathrm{DF}(\beta)$, but it turns out (Remark 18) that the two are equal for all $\beta > 1$.

1.7. **An example.** When $d_\beta(1)$ is preperiodic, a bare hands calculation of $\mathrm{DF}(\beta)$ can be carried out using Markov partition techniques (compare [11, 20]). For example, Figure 2 shows a Markov partition $(J_1, J_2, J_3, J_4, J_5)$ for $f_\beta$ when $d_\beta(1) = 2\,1\,2\,1\,\overline{0}$, so that $\beta^4 - 2\beta^3 - \beta^2 - 2\beta - 1 = 0$ ($\beta \simeq 2.7$). We see the digit 0 every time we visit $J_1$, the digit 1 every time we visit $J_2$ or $J_3$, and the digit 2 every time we visit $J_4$ or $J_5$.

The associated Markov transition graph has 10 *minimal loops* (loops which visit each interval at most once). These are $\overline{1}$, $\overline{1\,2}$, $\overline{1\,2\,4}$, $\overline{1\,3\,5\,2}$, $\overline{1\,3\,5\,2\,4}$, $\overline{1\,4}$, $\overline{1\,5\,2}$, $\overline{1\,5\,2\,4}$, $\overline{2}$, and $\overline{2\,3\,5}$, with corresponding digit frequencies $(1,0,0)$, $(1/2,1/2,0)$, $(1/3,1/3,1/3)$, $(1/4,1/2,1/4)$, $(1/5,2/5,2/5)$, $(1/2,0,1/2)$, $(1/3,1/3,1/3)$, $(1/4,1/4,1/2)$, $(0,1,0)$, and $(0,2/3,1/3)$ respectively. The digit frequency set is obtained by taking the convex hull of these frequencies: it is a pentagon with vertices $(1,0,0)$, $(1/2,0,1/2)$, $(1/4,1/4,1/2)$, $(0,1,0)$, and $(0,2/3,1/3)$. See Example 61.

This observation by itself is sufficient, using Theorem 3.4 of [20], to establish that $\mathrm{DF}(\beta)$ is a convex polytope with rational vertices whenever the $f_\beta$-orbit of 1 is finite. (A straightforward concatenation argument shows that, in the case where $w_\beta \ne d_\beta(1)$, the "missing" digit frequency of $w_\beta$ is realised as $d_\beta(x)$ for some $x$: this digit frequency corresponds to a loop $L$ in the Markov graph, which can be concatenated with any other intersecting loop $M$ in the pattern $LMLLMLLLM\ldots$ to provide the itinerary of a suitable point $x$.) However, quite different techniques are needed to address the general case.

This calculation, and several of the results presented in this paper, are reminiscent of rotation sets of torus homeomorphisms. The connection between the two problems is made explicit in [5].

We note also the connection of this work with the *digit distribution problem* first studied by Besicovitch [2] and Eggleston [9]. For expansions to *integer* bases $N$ and elements $\boldsymbol{\alpha}$ of the $(N-1)$-simplex $\Delta$, they considered the properties (for example, the Hausdorff dimension) of the sets

$$H_{\boldsymbol{\alpha}} = \{x \in [0,1] \,:\, \delta_N(x) = \boldsymbol{\alpha}\}.$$

The results presented here enable one to determine, for arbitrary bases $\beta$, the values of $\boldsymbol{\alpha}$ for which $H_{\boldsymbol{\alpha}}$ is non-empty.

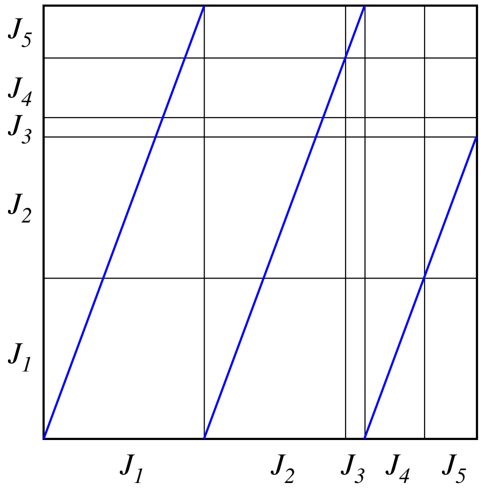


FIGURE 2. A Markov partition for $f_\beta$ when $d_\beta(1) = 2\,1\,2\,1\,\overline{0}$

## 2. Infimax sequences

2.1. **Notation and summary of results from** [6]**.** Let $k \geq 2$ be an integer, and $\Sigma_k = \{0, 1, \dots, k-1\}^{\mathbb{N}}$, the set of sequences $w = (w_r)_{r\geq 0}$ with entries in $\{0, 1, \dots, k-1\}$, endowed with the lexicographic order and the product topology (we consider the natural numbers $\mathbb{N}$ to include 0). The suffix $k$ will generally be suppressed, both on $\Sigma_k$ and on other $k$-dependent objects. We refer to elements of the alphabet $\{0, 1, \dots, k-1\}$ as *digits*, since elements of $\Sigma$ will be interpreted as digit sequences of $\beta$-expansions.

Denote by $\mathcal{W}$ the set of non-empty finite words $W$ over the alphabet $\{0, 1, \dots, k-1\}$, ordered lexicographically with the convention that any proper initial subword of $W$ is greater than $W$. We write $|W| \geq 1$ for the length of a word $W$.

If $V, W \in \mathcal{W}$, we write $VW$ for the concatenation of $V$ and $W$; $W^n$ for the $n$-fold repetition of $W$ (where $n \geq 1$); $\overline{W} = WWWW\dots$ for the element of $\Sigma$ given by infinite repetition of $W$; and $V\overline{W}$ for the element $VWWWW\dots$ of $\Sigma$. An element of $\Sigma$ of the form $\overline{W}$ is said to be *periodic*.

If $w \in \Sigma$ and $r \geq 1$, we write $w\,[\![r]\!] = w_0 w_1 \dots w_{r-1}$ for the word formed by the first $r$ digits of $w$.

The shift map $\sigma\colon \Sigma \to \Sigma$ is defined by $\sigma(w)_r = w_{r+1}$. An element $w$ of $\Sigma$ is said to be *maximal* if it is the maximum element of its $\sigma$-orbit: that is, if $\sigma^r(w) \leq w$ for all $r \geq 0$. We write $\mathcal{M}$ for the set of maximal elements of $\Sigma$.

As in the introduction, let $\Delta$ denote the standard $(k-1)$-simplex

$$\Delta = \left\{ \boldsymbol{\alpha} \in \mathbb{R}^k_{\geq 0} \,:\, \sum_{i=0}^{k-1} \alpha_i = 1 \right\}$$

with the Euclidean metric $d$. (In [6], $\Delta$ was endowed with the maximum metric to ease some of the calculations, but this is not necessary here.) $B_\epsilon(\boldsymbol{\alpha})$ denotes the open $\epsilon$-ball about $\boldsymbol{\alpha} \in \Delta$. Write $\mathcal{C}(\Delta)$ for the space of non-empty compact subsets of $\Delta$, with the Hausdorff metric $d_H$.

Given $\boldsymbol{\alpha} \in \Delta$, let $\mathcal{R}(\boldsymbol{\alpha}) \subset \Sigma$ be the set of sequences with digit frequency $\boldsymbol{\alpha}$,

$$\mathcal{R}(\boldsymbol{\alpha}) = \{ w \in \Sigma \,:\, \lim_{r\to\infty} N_{i,r}(w)/r = \alpha_i \text{ for each } i \},$$

and $\mathcal{M}(\boldsymbol{\alpha}) = \mathcal{M} \cap \mathcal{R}(\boldsymbol{\alpha})$, the set of maximal sequences with digit frequency $\boldsymbol{\alpha}$.

We begin with a brief summary of necessary results from [6]. We work here over the alphabet $\{0, 1, \dots, k-1\}$, as is appropriate for digit sequences of $\beta$-expansions, rather than over the alphabet

$\{1, 2, \ldots, k\}$ used in [6]. A second notational change is that we write

$$\Delta' = \{\boldsymbol{\alpha} \in \Delta \,:\, \alpha_{k-1} \neq 0\}$$

for the set of elements of the standard simplex whose final coordinate is non-zero: in [6] this set, which was the main object of study, was denoted $\Delta$, and the set here called $\Delta$ was denoted $\overline{\Delta}$.

Let $\Delta_n \subset \Delta$ be defined for $n \geq 0$ by

$$\Delta_n = \{\boldsymbol{\alpha} \in \Delta' \,:\, \lfloor \alpha_0/\alpha_{k-1} \rfloor = n\},$$

so that the $\Delta_n$ partition $\Delta'$. Define $J\colon \Delta' \to \mathbb{N}$ by $J(\boldsymbol{\alpha}) = \lfloor \alpha_0/\alpha_{k-1} \rfloor$, so that $\boldsymbol{\alpha} \in \Delta_{J(\boldsymbol{\alpha})}$ for each $\boldsymbol{\alpha} \in \Delta'$.

We define a multi-dimensional continued fraction map $K\colon \Delta' \to \Delta'$ by setting $K(\boldsymbol{\alpha}) = K_{J(\boldsymbol{\alpha})}(\boldsymbol{\alpha})$, where $K_n\colon \Delta_n \to \Delta'$ is the projective homeomorphism given by

$$K_n(\boldsymbol{\alpha}) = \left( \frac{\alpha_1}{1-\alpha_0}, \; \frac{\alpha_2}{1-\alpha_0}, \; \ldots, \; \frac{\alpha_{k-2}}{1-\alpha_0}, \; \frac{\alpha_0 - n\alpha_{k-1}}{1-\alpha_0}, \; \frac{(n+1)\alpha_{k-1} - \alpha_0}{1-\alpha_0} \right).$$

Let $\Phi\colon \Delta' \to \mathbb{N}^{\mathbb{N}}$ be the itinerary map of $K$ with respect to the partition $\{\Delta_n\}$. That is, for each $\boldsymbol{\alpha} \in \Delta'$, the sequence $\Phi(\boldsymbol{\alpha}) \in \mathbb{N}^{\mathbb{N}}$ is defined by

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

We order $\mathbb{N}^{\mathbb{N}}$ *reverse lexicographically*: if $\mathbf{m}$ and $\mathbf{n}$ are distinct elements of $\mathbb{N}^{\mathbb{N}}$, then $\mathbf{m} < \mathbf{n}$ if and only if $m_r > n_r$, where $r$ is the smallest index with $m_r \neq n_r$.

For each $n \in \mathbb{N}$, let $\Lambda_n\colon \Sigma \to \Sigma$ and $\Lambda_n\colon \mathcal{W} \to \mathcal{W}$ be the substitutions defined by

$$\Lambda_n\colon \quad \begin{cases} i & \mapsto \quad (i+1) \qquad\qquad \text{if } 0 \leq i \leq k-3 \\ (k-2) & \mapsto \quad (k-1)\, 0^{n+1} \\ (k-1) & \mapsto \quad (k-1)\, 0^n. \end{cases} \tag{1}$$

These substitutions are strictly order preserving and satisfy $\Lambda_n(\mathcal{M}) \subset \mathcal{M}$.

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

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

and let $S\colon \mathbb{N}^{\mathbb{N}} \to \mathcal{M} \subset \Sigma$ be given by

$$S(\mathbf{n}) = \lim_{r \to \infty} \Lambda_{\mathbf{n},r}(\overline{k-1}),$$

the limit existing since $\Lambda_{n_{r+1}}(k-1)$ begins with the digit $k-1$, so that $\Lambda_{\mathbf{n},r}(k-1)$ is an initial subword of $\Lambda_{\mathbf{n},r+1}(k-1)$ for all $r$. Finally, let $\mathcal{I} = S \circ \Phi\colon \Delta' \to \mathcal{M}$.

This family of substitutions is an example of an S-adic system. Such systems have been the subject of considerable work recently: see [1] for a survey.

The following results from [6] will be used here:

**Facts 1.**

*a) Let $\boldsymbol{\alpha} \in \Delta'$. Then $\mathcal{I}(\boldsymbol{\alpha})$ is the infimum of $\mathcal{M}(\boldsymbol{\alpha})$, the so-called $\boldsymbol{\alpha}$-*infimax *sequence.*

*b) Let $\boldsymbol{\alpha} \in \Delta'$ and $w \in \mathcal{R}(\boldsymbol{\alpha})$. Then $\mathcal{I}(\boldsymbol{\alpha}) \leq \sup_{r \geq 0} \sigma^r(w)$.*

*c) The itinerary $\Phi(\boldsymbol{\alpha})$ of $\boldsymbol{\alpha}$ is of the form $n_0\, n_1\, \ldots\, n_r\, \overline{0}$ if and only if $\boldsymbol{\alpha} \in \mathbb{Q}^k$. In this case $\mathcal{I}(\boldsymbol{\alpha}) = \overline{\Lambda_{\mathbf{n},r}(k-1)}$, and $\Lambda_{\mathbf{n},r}(k-1)$ is a primitive period of $\mathcal{I}(\boldsymbol{\alpha})$.*

*d) $\Phi\colon \Delta' \to \mathbb{N}^{\mathbb{N}}$ is lower semi-continuous and surjective. It is not injective except when $k = 2$: the preimage $\Phi^{-1}(\mathbf{n})$ of a point $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ is a $d$-simplex, where $0 \leq d = d(\mathbf{n}) \leq k-2$.*

*e) $S\colon \mathbb{N}^{\mathbb{N}} \to \mathcal{M}$ is a continuous order-preserving bijection onto its image $\mathcal{I}(\Delta') \subset \mathcal{M}$, the set of infimax sequences.*

*f)* $\mathcal{I}(\boldsymbol{\alpha}) \in \mathcal{M}(\boldsymbol{\alpha})$ *(and hence* $\mathcal{I}(\boldsymbol{\alpha}) = \min \mathcal{M}(\boldsymbol{\alpha})$*) if and only if* $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) = \{\boldsymbol{\alpha}\}$. *This is always the case when* $\boldsymbol{\alpha} \in \mathbb{Q}^k$. *If* $\Phi^{-1}(\Phi(\boldsymbol{\alpha})) \neq \{\boldsymbol{\alpha}\}$ *then* $\mathcal{I}(\boldsymbol{\alpha})$ *does not have well-defined digit frequency.*

*g) If* $\mathcal{R}'(\boldsymbol{\alpha})$ *denotes the set of sequences with subsequential digit frequency* $\boldsymbol{\alpha}$*, i.e.*

$$\mathcal{R}'(\boldsymbol{\alpha}) = \{w \in \Sigma \,:\, \lim_{s\to\infty} N_{i,r_s}(w)/r_s = \alpha_i \text{ for some } r_s \to \infty \text{ and each } i\},$$

*and* $\mathcal{M}'(\boldsymbol{\alpha}) = \mathcal{M} \cap \mathcal{R}'(\boldsymbol{\alpha})$*, then a) and b) hold in the primed versions: that is,* $\mathcal{I}(\boldsymbol{\alpha}) = \inf \mathcal{M}'(\boldsymbol{\alpha})$*, and* $\mathcal{I}(\boldsymbol{\alpha}) \le \sup_{r\ge 0} \sigma^r(w)$ *for all* $w \in \mathcal{R}'(\boldsymbol{\alpha})$.

*h) The set of itineraries* $\mathbf{n}$ *for which* $\Phi^{-1}(\mathbf{n})$ *is a single point contains the dense* $G_\delta$ *subset* $\mathcal{O}$ *of* $\mathbb{N}^{\mathbb{N}}$ *consisting of those sequences which contain infinitely many distinct subwords* $1^{2k-3}$. *(For* $k = 3$ *this is a result of Bruin and Troubetzkoy* [7]*.)*

*i) An element* $\boldsymbol{\alpha}$ *of* $\Delta'$ *has the property that its orbit* $(K^r(\boldsymbol{\alpha}))_{r\ge 0}$ *is disjoint from the faces of* $\Delta$ *if and only if its itinerary* $\mathbf{n} = \Phi(\boldsymbol{\alpha})$ *has the following property: for every* $r \ge 0$ *there is some* $s \ge 0$ *with* $n_{r+s(k-1)} \neq 0$ *(that is, there is no congruence class of indices modulo* $k-1$ *on which* $\mathbf{n}$ *is eventually zero).*

**Definitions 2** (Rational type, regular, exceptional)**.** Let $\boldsymbol{\alpha} \in \Delta'$ with itinerary $\mathbf{n} = \Phi(\boldsymbol{\alpha})$. We say that $\boldsymbol{\alpha}$ and $\mathbf{n}$ are of *rational type* if $\boldsymbol{\alpha} \in \mathbb{Q}^k$: that is, if $\mathbf{n} = n_0\, n_1\, \dots\, n_r\, \overline{0}$ for some $n_0, n_1, \dots, n_r$. We say that $\boldsymbol{\alpha}$ and $\mathbf{n}$ are *regular* if $\Phi^{-1}(\mathbf{n}) = \{\boldsymbol{\alpha}\}$, and that they are *exceptional* otherwise. In the latter case, we refer to the non-trivial simplex $\Phi^{-1}(\mathbf{n}) \subset \Delta'$ as an *exceptional set.*

It can be shown [6] that, for $k \ge 3$, an element $\mathbf{n}$ of $\mathbb{N}^{\mathbb{N}}$ is regular when it grows slowly enough and has only finitely many zero entries; and that it is exceptional when it grows rapidly enough.

**Example 3.** Let $k = 3$ and $\boldsymbol{\alpha} = (7/16, 5/16, 4/16) \in \Delta'$. We have

$$(7/16, 5/16, 4/16) \xrightarrow{K_1} (5/9, 3/9, 1/9) \xrightarrow{K_5} (3/4, 0, 1/4) \xrightarrow{K_3} (0, 0, 1),$$

and $K_0(0,0,1) = (0,0,1)$. Therefore $\boldsymbol{\alpha}$ has itinerary $\Phi(\boldsymbol{\alpha}) = 1\,5\,3\,\overline{0}$. The $\boldsymbol{\alpha}$-infimax sequence is

$$\mathcal{I}(\boldsymbol{\alpha}) = S(1\,5\,3\,\overline{0}) = \Lambda_1(\Lambda_5(\Lambda_3(\overline{2}))) = \overline{2\,0\,1\,1\,1\,1\,1\,2\,0\,0\,2\,0\,0\,2\,0\,0}.$$

This is the smallest maximal sequence with digit frequency $\boldsymbol{\alpha}$.

2.2. **Convergence to the exceptional set.** In this section we establish information about exceptional sets which goes beyond that contained in [6]. The results are technical, and their proofs could be omitted on first reading.

We fix an element $\mathbf{n}$ of $\mathbb{N}^{\mathbb{N}}$, and begin by describing, as in [6], a decreasing sequence $(A_{\mathbf{n},r})_{r\ge 0}$ of simplices whose intersection is $\Phi^{-1}(\mathbf{n})$.

The homeomorphism $K_n\colon \Delta_n \to \Delta'$ has inverse given by

$$K_n^{-1}(\boldsymbol{\alpha}) = \left( \frac{(n+1)\,\alpha_{k-2} + n\,\alpha_{k-1}}{D}, \frac{\alpha_0}{D}, \frac{\alpha_1}{D}, \dots, \frac{\alpha_{k-3}}{D}, \frac{\alpha_{k-2} + \alpha_{k-1}}{D} \right), \tag{2}$$

where $D = (n+1)\,\alpha_{k-2} + n\,\alpha_{k-1} + 1$. The homeomorphism $K_n^{-1}\colon \Delta' \to \Delta_n$ extends by the same formula to a homeomorphism $K_n^{-1}\colon \Delta \to \mathrm{Cl}\,(\Delta_n) \subset \Delta$.

**Definitions 4** ($\Upsilon_{\mathbf{n},r}$, $A_{\mathbf{n},r}$, $\mathcal{F}$, $\mathcal{F}_{\mathbf{n},r}$)**.** For each $r \in \mathbb{N}$, we define an embedding

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

Let $\mathcal{F} = \Delta \setminus \Delta'$ denote the face $\alpha_{k-1} = 0$ of $\Delta$, and write

$$A_{\mathbf{n},r} = \Upsilon_{\mathbf{n},r}(\Delta) \qquad \text{and} \qquad \mathcal{F}_{\mathbf{n},r} = \Upsilon_{\mathbf{n},r}(\mathcal{F}).$$

Since each $K_n^{-1}$ is projective, $A_{\mathbf{n},r}$ is a $(k-1)$-simplex and $\mathcal{F}_{\mathbf{n},r}$ is a $(k-2)$-simplex for all $r$. Moreover, since $K_{n_r}^{-1}(\Delta) \subset \Delta$, the sequence $(A_{\mathbf{n},r})_{r\geq 0}$ is decreasing.

Now

$$\Upsilon_{\mathbf{n},r}(\Delta') = A_{\mathbf{n},r} \setminus \mathcal{F}_{\mathbf{n},r} = \{\boldsymbol{\alpha} \in \Delta' \,:\, \Phi(\boldsymbol{\alpha})\,[\![r+1]\!] = \mathbf{n}\,[\![r+1]\!]\},$$

the set of frequencies whose itineraries agree with $\mathbf{n}$ on their first $r+1$ entries. On the other hand,

$$\boldsymbol{\alpha} \in \mathcal{F}_{\mathbf{n},r} \cap \Delta' \implies \Phi(\boldsymbol{\alpha})\,[\![r-i+1]\!] = n_0 n_1 \ldots n_{r-i-1}\,(n_{r-i}+1) \text{ for some } i \text{ with } 0 \leq i \leq k-2. \tag{3}$$

In particular, $\Phi(\boldsymbol{\alpha}) < \mathbf{n}$ for all $\boldsymbol{\alpha} \in \mathcal{F}_{\mathbf{n},r}$ when $r \geq k-2$, since $\mathcal{F}_{\mathbf{n},r} \subset \Delta'$ for $r \geq k-2$.

Let $\boldsymbol{\alpha} \in \Delta'$. If $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$ then $\boldsymbol{\alpha} \in A_{\mathbf{n},r}$ for all $r$. On the other hand, if $\Phi(\boldsymbol{\alpha}) \neq \mathbf{n}$, let $r \in \mathbb{N}$ be such that $\Phi(\boldsymbol{\alpha})_r \neq n_r$: then $\boldsymbol{\alpha} \notin \mathcal{F}_{\mathbf{n},r+k}$ by (3), and hence $\boldsymbol{\alpha} \notin A_{\mathbf{n},r+k}$. We therefore have

- every element of $\Delta'$ whose itinerary starts $n_0 \ldots n_r$ lies in $A_{\mathbf{n},r}$, and
- if $r \geq k$, then every element of $A_{\mathbf{n},r}$ has itinerary starting $n_0 \ldots n_{r-k}$.

In particular,

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

and the decreasing sequence $(A_{\mathbf{n},r})$ converges Hausdorff to $\Phi^{-1}(\mathbf{n})$. Moreover, since $(A_{\mathbf{n},r})$ is a decreasing sequence of simplices, it follows by a theorem of Borovikov [4] that $\Phi^{-1}(\mathbf{n})$ is also a simplex. By Facts 1c), this simplex cannot have interior in $\Delta$ (since then there would be both rational and non-rational points having itinerary $\mathbf{n}$), and so it has dimension at most $k-2$.

The following lemma plays a key rôle in the proofs of Lemma 16 and Theorem 37, two of the central results of the paper.

**Lemma 5.** *Let $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$. Then for every $\epsilon > 0$, there are infinitely many $r$ with $d_H(\mathcal{F}_{\mathbf{n},r}, \Phi^{-1}(\mathbf{n})) < \epsilon$.*

*Proof.* If $\mathbf{n}$ is regular then $\Phi^{-1}(\mathbf{n})$ is a point and $\mathcal{F}_{\mathbf{n},r} \subset A_{\mathbf{n},r} \to \Phi^{-1}(\mathbf{n})$, so the result is immediate. We can therefore assume that $\mathbf{n}$ is exceptional. In particular, $\boldsymbol{\alpha}$ is not of rational type, and hence $n_r \neq 0$ for infinitely many $r$.

Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ $(2 \leq n \leq k-1)$ be the vertices of the simplex $\Phi^{-1}(\mathbf{n})$. Write $\mathbf{e}_0, \ldots \mathbf{e}_{k-1}$ for the vertices of $\Delta$ (so that the $i^{\text{th}}$ component of $\mathbf{e}_i$ is 1), and let $\boldsymbol{\alpha}_r^{(i)} = \Upsilon_{\mathbf{n},r}(\mathbf{e}_i)$. Therefore the vertices of $\mathcal{F}_{\mathbf{n},r}$ are $\boldsymbol{\alpha}_r^{(i)}$ for $0 \leq i \leq k-2$, and $A_{\mathbf{n},r}$ has the additional vertex $\boldsymbol{\alpha}_r^{(k-1)}$. Notice, by comparison of (1) and (2), that $\boldsymbol{\alpha}_r^{(i)}$ is the digit frequency of the word $\Lambda_{\mathbf{n},r}(i)$. We write $L_r^{(i)}$ for the length of $\Lambda_{\mathbf{n},r}(i)$, so that $L_r^{(i)}\boldsymbol{\alpha}_r^{(i)}$ is an integer vector whose entries give the number of each digit in $\Lambda_{\mathbf{n},r}(i)$.

Since $A_{\mathbf{n},r} \to \Phi^{-1}(\mathbf{n})$ as $r \to \infty$ we have $\mathcal{F}_{\mathbf{n},r} \subset A_{\mathbf{n},r} \subset B_\epsilon(\Phi^{-1}(\mathbf{n}))$ for all sufficiently large $r$. It remains to prove that $\Phi^{-1}(\mathbf{n}) \subset B_\epsilon(\mathcal{F}_{\mathbf{n},r})$ for infinitely many $r$, and for this it is enough to show that, for infinitely many $r$, every vertex of $\Phi^{-1}(\mathbf{n})$ is approximated by a vertex of $\mathcal{F}_{\mathbf{n},r}$: that is,

$$\forall \epsilon > 0,\ \forall R,\ \exists r \geq R,\ \forall m \leq n,\ \exists i \leq k-2,\ \boldsymbol{\alpha}_r^{(i)} \in B_\epsilon(\mathbf{v}_m).$$

Suppose for a contradiction that there exist $\epsilon > 0$ and $R$ such that for all $r \geq R$ there is some $m$ for which $B_\epsilon(\mathbf{v}_m)$ doesn't contain $\boldsymbol{\alpha}_r^{(i)}$ for any $i \leq k-2$. Decrease $\epsilon$ if necessary so that the distance between any two vertices of $\Phi^{-1}(\mathbf{n})$ is at least $2\epsilon$; and increase $R$ if necessary so that $d_H(A_{\mathbf{n},r}, \Phi^{-1}(\mathbf{n})) < \epsilon/5$ for all $r \geq R$. In particular this means that, for all $r \geq R$, every $B_{\epsilon/5}(\mathbf{v}_\ell)$ contains some vertex $\boldsymbol{\alpha}_r^{(i)}$ of $A_{\mathbf{n},r}$.

Pick $r \geq R$ with $n_{r+1} \neq 0$. Since there is some $m$ for which none of the $\boldsymbol{\alpha}_r^{(i)}$ with $i \leq k-2$ lie in $B_\epsilon(\mathbf{v}_m)$, we must have $\boldsymbol{\alpha}_r^{(k-1)} \in B_{\epsilon/5}(\mathbf{v}_m)$.

Now (1) gives $\boldsymbol{\alpha}_{r+1}^{(i)} = \boldsymbol{\alpha}_r^{(i+1)}$ for $0 \le i \le k-3$, and

$$\boldsymbol{\alpha}_{r+1}^{(k-2)} = \frac{(n_{r+1}+1)L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(k-1)}\boldsymbol{\alpha}_r^{(k-1)}}{(n_{r+1}+1)L_r^{(0)} + L_r^{(k-1)}}, \qquad \text{and} \tag{4}$$

$$\boldsymbol{\alpha}_{r+1}^{(k-1)} = \frac{n_{r+1}L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(k-1)}\boldsymbol{\alpha}_r^{(k-1)}}{n_{r+1}L_r^{(0)} + L_r^{(k-1)}}. \tag{5}$$

Since the only vertices of $A_{\mathbf{n},r+1}$ which are not also vertices of $A_{\mathbf{n},r}$ are $\boldsymbol{\alpha}_{r+1}^{(k-2)}$ and $\boldsymbol{\alpha}_{r+1}^{(k-1)}$, one of these must lie in $B_{\epsilon/5}(\mathbf{v}_m)$. However, both lie along the line segment joining $\boldsymbol{\alpha}_r^{(0)}$ to $\boldsymbol{\alpha}_r^{(k-1)}$, and $\boldsymbol{\alpha}_{r+1}^{(k-1)}$ is the closer of the two to $\boldsymbol{\alpha}_r^{(k-1)}$. Therefore $\boldsymbol{\alpha}_{r+1}^{(k-1)} \in B_{\epsilon/5}(\mathbf{v}_m)$.

Let $d_1 < d_2$ be the distances from $\boldsymbol{\alpha}_r^{(k-1)}$ to $\boldsymbol{\alpha}_{r+1}^{(k-1)}$ and $\boldsymbol{\alpha}_{r+1}^{(k-2)}$ respectively. Then

$$\frac{d_2}{d_1} = \frac{n_{r+1}+1}{n_{r+1}} \, \frac{n_{r+1}L_r^{(0)} + L_r^{(k-1)}}{(n_{r+1}+1)L_r^{(0)} + L_r^{(k-1)}} < 2$$

since $n_{r+1} \ge 1$ by choice of $r$, and so $(n_{r+1}+1)/n_{r+1} \le 2$. However, since both $\boldsymbol{\alpha}_r^{(k-1)}$ and $\boldsymbol{\alpha}_{r+1}^{(k-1)}$ lie in $B_{\epsilon/5}(\mathbf{v}_m)$ we have $d_1 < 2\epsilon/5$, and hence $d_2 < 4\epsilon/5$. Therefore

$$d(\boldsymbol{\alpha}_{r+1}^{(k-2)}, \mathbf{v}_m) \le d_2 + d(\boldsymbol{\alpha}_r^{(k-1)}, \mathbf{v}_m) < \frac{4\epsilon}{5} + \frac{\epsilon}{5} = \epsilon.$$

This is the required contradiction. For since both $\boldsymbol{\alpha}_{r+1}^{(k-2)}$ and $\boldsymbol{\alpha}_{r+1}^{(k-1)}$ are within $\epsilon$ of $\mathbf{v}_m$, every other vertex $\mathbf{v}_\ell$ of $\Phi^{-1}(\mathbf{n})$ must have $d(\mathbf{v}_\ell, \boldsymbol{\alpha}_{r+1}^{(i)}) < \epsilon/5$ for some $i < k-2$. $\square$

As a consequence, the itinerary map $\Phi$ has no local minima:

**Corollary 6.** *Let $\boldsymbol{\alpha} \in \Delta'$. Then for all $\epsilon > 0$ there is some $\boldsymbol{\beta} \in B_\epsilon(\boldsymbol{\alpha})$ with $\Phi(\boldsymbol{\beta}) < \Phi(\boldsymbol{\alpha})$.*

*Proof.* Write $\mathbf{n} = \Phi(\boldsymbol{\alpha})$. Let $r \ge k-2$ be such that $d_H(\mathcal{F}_{\mathbf{n},r}, \Phi^{-1}(\mathbf{n})) < \epsilon$, so that there is a point $\boldsymbol{\beta}$ of $\mathcal{F}_{\mathbf{n},r}$ within distance $\epsilon$ of $\boldsymbol{\alpha} \in \Phi^{-1}(\mathbf{n})$. Then $\Phi(\boldsymbol{\beta}) < \mathbf{n} = \Phi(\boldsymbol{\alpha})$ by (3). $\square$

2.3. **Concatenations of repeating blocks.** We will need the following straightforward result about concatentations of repeating blocks of rational infimaxes.

**Lemma 7.** *Let $\boldsymbol{\alpha} \in \Delta'$, and $(\boldsymbol{\alpha}_i)$ be a sequence of rational elements of $\Delta'$ with the property that $\mathcal{I}(\boldsymbol{\alpha}_i) < \mathcal{I}(\boldsymbol{\alpha})$ for all $i$. Write $B_i$ for the repeating block of $\mathcal{I}(\boldsymbol{\alpha}_i)$. Then the sequence*

$$w = B_0\, B_1\, B_2\, \ldots \in \Sigma$$

*satisfies $\sigma^r(w) \le \mathcal{I}(\boldsymbol{\alpha})$ for all $r \ge 0$.*

*Proof.* We will show that, for each $i$,

a) $B_i$ is strictly smaller than the initial length $|B_i|$ subword of $\mathcal{I}(\boldsymbol{\alpha})$; and
b) every proper final subword of $B_i$ is strictly smaller than $B_i$,

from which the result follows. Recall that, by convention, any word is smaller than any of its proper initial subwords, so that the lexicographic order on the set $\mathcal{W}$ of words is total.

For a), since $\overline{B_i} = \mathcal{I}(\boldsymbol{\alpha}_i) < \mathcal{I}(\boldsymbol{\alpha})$, it is enough to show that $B_i$ is not an initial subword of $\mathcal{I}(\boldsymbol{\alpha})$.

Let $\Phi(\boldsymbol{\alpha}_i) = n_0 \ldots n_r\overline{0}$. Since $\mathcal{I}(\boldsymbol{\alpha}_i) < \mathcal{I}(\boldsymbol{\alpha})$, there is some $j \le r$ such that $\Phi(\boldsymbol{\alpha}) = n_0 \ldots n_{j-1}m_j \ldots$, where $m_j < n_j$. Then $\Lambda_{n_j}(k-1) = (k-1)\, 0^{m_j}0 \ldots$, and hence $B_i$ has an initial subword of the form $P_i = \Lambda_{\mathbf{n},j-1}((k-1)\, 0^{m_j}0)$.

Similarly, $\mathcal{I}(\boldsymbol{\alpha})$ has an initial subword $P = \Lambda_{\mathbf{n},j-1}((k-1)\,0^{m_j}\,s\,W)$, where $s > 0$ and $W$ is a word long enough to ensure that $|P| > |P'|$.

Since $\Lambda_{\mathbf{n},j-1} \colon \mathcal{W} \to \mathcal{W}$ is strictly order-preserving, it follows that $P' < P$. Therefore $P'$ is not an initial subword of $\mathcal{I}(\boldsymbol{\alpha})$, and so neither is $B_i$, as required.

For b), suppose for a contradiction that $B_i$ has a proper final subword $W$ with $W > B_i$. Since $\mathcal{I}(\boldsymbol{\alpha}_i) = \overline{B_i}$ is a maximal sequence, $W$ must also be an initial subword of $B_i$. Therefore there are words $U$ and $V$ of the same length with $B_i = WU = VW$. Then $\mathcal{I}(\boldsymbol{\alpha}_i) = \overline{WU} = \overline{VW} > \overline{WV}$, with the inequality coming from the maximality of $\mathcal{I}(\boldsymbol{\alpha}_i)$ and Facts 1c). Therefore $U > V$, so that $\sigma^{|W|}(\mathcal{I}(\boldsymbol{\alpha}_i)) = \overline{UW} > \overline{VW} = \mathcal{I}(\boldsymbol{\alpha}_i)$, contradicting the maximality of $\mathcal{I}(\boldsymbol{\alpha}_i)$.

□

2.4. **Compactification of the space of itineraries.** In this section we describe a compactification of $\mathbb{N}^{\mathbb{N}}$ to a Cantor set $\mathcal{N}$, and extend the map $S \colon \mathbb{N}^{\mathbb{N}} \to \mathcal{M}$, which associates an infimax sequence with each itinerary, over $\mathcal{N}$ as an order-preserving homeomorphism onto its image. The set $\mathcal{N}$ and its topology and order are modelled by the function $g$ whose graph is depicted in Figure 3. In this figure $(L_n)_{n\geq 0}$ is a sequence of mutually disjoint closed subintervals of $[0,1]$, and $g$ maps each $L_n$ affinely and increasingly onto $[0,1]$ and satisfies $g(0) = 0$. The biggest set

$$C = \bigcap_{r=0}^{\infty} g^{-r}\left(\{0\} \cup \bigcup_{n=0}^{\infty} L_n\right)$$

on which all of the iterates of $g$ are defined is a Cantor set, and the points of $C$ correspond bijectively via their itineraries with the union of $\mathbb{N}^{\mathbb{N}}$ and the set of finite words in $\mathbb{N}$ terminated with the symbol $\infty$, which is regarded as the address of 0. The gaps in this Cantor set have points whose itineraries end $\overline{0}$ at their left hand ends, and points whose itineraries are terminated with $\infty$ at their right hand ends.

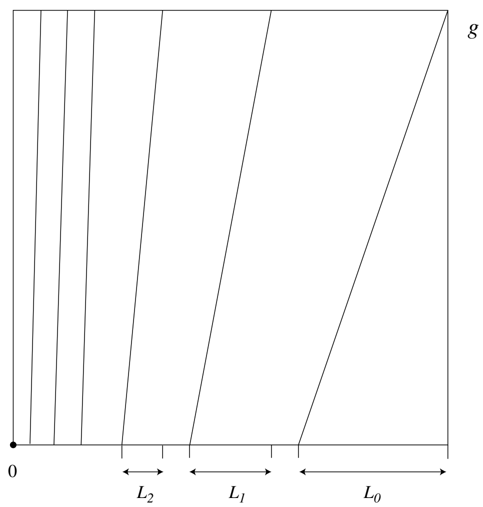


FIGURE 3. Model for the topology and order on $\mathcal{N}$

**Definitions 8** ($\mathcal{N}$, finite type)**.** We write

$$\mathcal{N} = \mathbb{N}^{\mathbb{N}} \cup \{W\infty \,:\, W \text{ is a (possibly empty) word over the alphabet } \mathbb{N}\},$$

Elements of $\mathcal{N}$ of the form $W\infty$ are said to be of *finite type.*

We extend the reverse lexicographic ordering of $\mathbb{N}^{\mathbb{N}}$ to $\mathcal{N}$. Because finite type elements of $\mathcal{N}$ are terminated by $\infty$, every pair $\mathbf{m}, \mathbf{n}$ of distinct elements of $\mathcal{N}$ first disagree at some index $r$ at which both $m_r$ and $n_r$ are either natural numbers or $\infty$: we say that $\mathbf{m} < \mathbf{n}$ if and only if either $m_r = \infty$, or $m_r$ and $n_r$ are both natural numbers with $m_r > n_r$.

We regard two elements of $\mathcal{N}$ as being close if either they agree up to a large index, or if they have large entries up to the point where they first disagree. To define a metric $d$ on $\mathcal{N}$ reflecting this we write, for each pair $\mathbf{m} \neq \mathbf{n}$ of elements of $\mathcal{N}$,

$$\begin{aligned} R(\mathbf{m},\mathbf{n}) &= \min\{r : m_r \neq n_r\}, \\ X(\mathbf{m},\mathbf{n}) &= r + \min\left(\sum_{s\leq r} m_s,\ \sum_{s\leq r} n_s\right) \qquad \text{where } r = R(\mathbf{m},\mathbf{n}), \text{ and} \\ d(\mathbf{m},\mathbf{n}) &= 2^{-X(\mathbf{m},\mathbf{n})}. \end{aligned}$$

Notice that $\min(\sum_{s\leq r} m_s, \sum_{s\leq r} n_s) = \sum_{s<r} m_s + \min(m_r, n_r) < \infty$. It is easily verified that $X(\mathbf{m},\mathbf{p}) \geq \min(X(\mathbf{m},\mathbf{n}), X(\mathbf{n},\mathbf{p}))$ for all $\mathbf{m},\mathbf{n},\mathbf{p} \in \mathcal{N}$, so that $d$ satisfies the triangle inequality. We will use the following property of $d$: if $\mathbf{n} \in \mathcal{N}$ and $n_r \in \mathbb{N}$ (i.e. $\mathbf{n}$ is not a finite type element of length $r+1$ or less), then

$$d(\mathbf{m},\mathbf{n}) < 2^{-(r+\sum_{s\leq r} n_s)} \quad \Longrightarrow \quad \mathbf{m} \text{ has initial subword } n_0 n_1 \dots n_r.$$

Observe that the metric is compatible with the order on $\mathcal{N}$, in the sense that

$$\mathbf{m} \leq \mathbf{n} \leq \mathbf{p} \quad \Longrightarrow \quad d(\mathbf{m},\mathbf{n}) \leq d(\mathbf{m},\mathbf{p}) \quad \text{and} \quad d(\mathbf{n},\mathbf{p}) \leq d(\mathbf{m},\mathbf{p}).$$

**Lemma 9.** *$(\mathcal{N}, d)$ is compact.*

*Proof.* Since every sequence in $\mathcal{N}$ has a monotonic subsequence, it suffices to show that every monotonic sequence converges.

Consider first an increasing sequence $(\mathbf{n}^{(r)})$, and assume without loss of generality that it is not eventually constant. Recalling that the order on $\mathcal{N}$ is reverse lexicographic, it is straightforward to show inductively that for every $s \in \mathbb{N}$ the sequence $\left(n_s^{(r)}\right)_{r\geq 0}$ is eventually defined and takes some constant value $m_s$. Then $\mathbf{n}^{(r)} \to \mathbf{m}$ as $r \to \infty$.

Now let $(\mathbf{n}^{(r)})$ be a decreasing sequence which is not eventually constant. If the sequence $\left(n_s^{(r)}\right)_{r\geq 0}$ is bounded above by some $K_s$ for all $s \in \mathbb{N}$, then the sequence converges by the same argument as in the increasing case. So suppose this is not the case, and let $S \in \mathbb{N}$ be least such that $\left(n_S^{(r)}\right)_{r\geq 0}$ is not bounded above. As in the increasing case, there are natural numbers $m_s$ for $s < S$ and a natural number $R$ such that $n_s^{(r)} = m_s$ for all $r > R$ and $s < S$. Then the sequence $(\mathbf{n}^{(r)})$ converges to $\mathbf{m} = m_0 m_1 \dots m_{S-1}\infty$, since $d(\mathbf{n}^{(r)}, \mathbf{m}) = 2^{-\left(S + n_S^{(r)} + \sum_{s<S} m_s\right)}$ for all $r > R$, and $n_S^{(r)} \to \infty$ as $r \to \infty$. $\square$

**Definitions 10** ($S\colon \mathcal{N} \to \mathcal{M}$, $\mathcal{J}$)**.** Extend the function $S\colon \mathbb{N}^{\mathbb{N}} \to \mathcal{M}$ to $\mathcal{N}$ in the natural way, using $\Lambda_\infty(k-1) = (k-1)\,\overline{0}$:

$$S(n_0\, n_1\, \dots\, n_r\, \infty) = \Lambda_{n_0} \circ \Lambda_{n_1} \circ \dots \circ \Lambda_{n_r}\left((k-1)\,\overline{0}\right).$$

Write $\mathcal{J} = S(\mathcal{N}) \subset \mathcal{M}$ for the image of $S$, the union of the set of infimax sequences with the countable set of sequences just defined.

**Lemma 11.** *$S\colon \mathcal{N} \to \mathcal{J}$ is an order-preserving homeomorphism.*

*Proof.* To show that $S$ is order-preserving, let $\mathbf{m}, \mathbf{n} \in \mathcal{N}$ with $\mathbf{m} < \mathbf{n}$, and write $r = R(\mathbf{m}, \mathbf{n})$ so that $m_r > n_r$ (perhaps $m_r = \infty$). Then there is some digit $\ell \in \{0, \dots, k-1\}$ such that $S(\mathbf{n})$ has initial subword

$$\Lambda_{\mathbf{n},r-1}(\Lambda_{n_r}((k-1)\ell)) = \Lambda_{\mathbf{n},r-1}((k-1)\,0^{n_r}\Lambda_n(\ell)).$$

Now $S(\mathbf{m})$ has initial subword $\Lambda_{\mathbf{n},r-1}((k-1)\,0^{n_r+1})$ because $m_r \ge n_r + 1$, and $(k-1)\,0^{n_r}\Lambda_n(\ell) > (k-1)\,0^{n_r+1}$ since $\Lambda_n(\ell)$ starts with a digit other than 0. Since the $\Lambda_n$ are strictly order-preserving, it follows that $S(\mathbf{m}) < S(\mathbf{n})$ as required.

In particular $S$ is a bijection, and in view of the compactness of $\mathcal{N}$ it only remains to show that $S$ is continuous. Let $\mathbf{n} \in \mathcal{N}$. To show continuity of $S$ at $\mathbf{n}$, we distinguish three cases.

a) Suppose that $\mathbf{n} = n_0\, n_1\, \dots\, n_r\, \infty$ is of finite type, so that $S(\mathbf{n}) = \Lambda_{\mathbf{n},r}((k-1)\,\overline{0})$. For each $N \in \mathbb{N}$, if $d(\mathbf{n}, \mathbf{m}) < 2^{-\left(r + \sum_{s \le r} n_s + N\right)}$ then $\mathbf{m}$ has initial subword $n_0\, n_1\, \dots\, n_r\, M$ for some $M > N$, so $S(\mathbf{m})$ has initial subword $\Lambda_{\mathbf{n},r}((k-1)\,0^N)$. Therefore $S(\mathbf{m}) \to S(\mathbf{n})$ as $\mathbf{m} \to \mathbf{n}$.
b) Suppose that $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ is not of rational type, so that there are arbitrarily large $r$ with $n_r \ne 0$ and hence the length of the word $\Lambda_{\mathbf{n},r}(k-1)$ goes to $\infty$ as $r \to \infty$. Given $r \in \mathbb{N}$, if $d(\mathbf{n}, \mathbf{m}) < 2^{-\left(r + \sum_{s \le r} n_s\right)}$ then $m_s = n_s$ for all $s \le r$: therefore $\Lambda_{\mathbf{m},r}(k-1) = \Lambda_{\mathbf{n},r}(k-1)$ for all $\mathbf{m}$ sufficiently close to $\mathbf{n}$, so that $S(\mathbf{m}) \to S(\mathbf{n})$ as $\mathbf{m} \to \mathbf{n}$.
c) Finally, suppose that $\mathbf{n} = n_0\, n_1\, \dots\, n_r\, \overline{0}$ is of rational type, so that we have $S(\mathbf{n}) = \Lambda_{\mathbf{n},r}(\overline{k-1})$. It can be shown (see for example the proof of Lemma 4 of [6]) that for every $R \ge 0$, the word $\Lambda_{\mathbf{n},r}(\Lambda_0^R((k-1)\,0))$ has initial subword $\Lambda_{\mathbf{n},r}((k-1)^{1+\lfloor R/(k-1)\rfloor})$. Now if $\mathbf{m} \ne \mathbf{n}$ and $d(\mathbf{n}, \mathbf{m}) < 2^{-\left(r + R + \sum_{s \le r} n_s\right)}$, then $\mathbf{m}$ has initial subword $n_0 n_1 \dots n_r 0^{R+T} m$ for some $T \ge 0$ and $m > 0$ (perhaps $m = \infty$); and hence $S(\mathbf{m})$ has initial subword $\Lambda_{\mathbf{n},r}(\Lambda_0^{R+T}((k-1)\,0))$, agreeing with $S(\mathbf{n})$ on a subword of length at least $1 + \lfloor R/(k-1)\rfloor$. Therefore $S(\mathbf{m}) \to S(\mathbf{n})$ as $\mathbf{m} \to \mathbf{n}$.

□

**Lemma 12.** *$(\mathcal{N}, d)$ is a Cantor set.*

*Proof.* It is a compact metric space which is homeomorphic to a subset of $\Sigma$ and so is totally disconnected. It therefore only remains to show that every $\mathbf{n} \in \mathcal{N}$ is the limit of a sequence $\mathbf{m}^{(r)}$ of other elements of $\mathcal{N}$.

If $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ then we can take $m_r^{(r)} = n_r + 1$ and $m_s^{(r)} = n_s$ for all $s \ne r$. On the other hand, if $\mathbf{n} = n_0\, \dots\, n_{R-1}\, \infty$ is of finite type, then we can take $m_s^{(r)} = n_s$ for all $r$ if $0 \le s < R$, and $m_s^{(r)} = r$ for all $r$ if $s \ge R$. □

The final lemma in this section shows that, as suggested by Figure 3, pairs of consecutive elements of $\mathcal{N}$ consist of one element of rational type and one of finite type.

**Lemma 13.** *Let $n_0, n_1, \dots, n_R \in \mathbb{N}$ for some $R \ge 0$. Then $n_0\, \dots\, n_{R-1}\, (n_R{+}1)\, \overline{0} \;<\; n_0\, \dots\, n_{R-1}\, n_R\, \infty$ are consecutive elements of $\mathcal{N}$.*

*On the other hand, every element of $\mathcal{N}$ which is not of rational (respectively finite) type is the limit of a strictly decreasing (respectively strictly increasing) sequence in $\mathcal{N}$.*

*Proof.* It is clear that $n_0\, \dots\, n_{R-1}\, (n_R + 1)\, \overline{0}$ and $n_0\, \dots\, n_{R-1}\, n_R\, \infty$ are consecutive, the former being the largest element of $\mathcal{N}$ starting $n_0\, \dots\, n_{R-1}\, (n_R + 1)$ and the latter the smallest element starting $n_0\, \dots\, n_{R-1}\, n_R$.

If $\mathbf{n}$ is not of finite type then, as in the proof of Lemma 12, there is a sequence $\mathbf{m}^{(r)} \to \mathbf{n}$ with $\mathbf{m}^{(r)} < \mathbf{n}$ for all $r$.

Suppose then that $\mathbf{n}$ is not of rational type. If it is of finite type then, as in the proof of Lemma 12, there is a sequence $\mathbf{m}^{(r)} \to \mathbf{n}$ with $\mathbf{m}^{(r)} > \mathbf{n}$ for all $r$. On the other hand, if $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ then there is an increasing sequence $i_r \to \infty$ with $n_{i_r} > 0$ for each $r$, and we can define a sequence $(\mathbf{m}^{(r)})$ converging to $\mathbf{n}$ from above by taking $m_s^{(r)} = n_s$ for $s \neq i_r$ and $m_{i_r}^{(r)} = n_{i_r} - 1$. □

**Definition 14** (Rational-finite pair)**.** For every $R \geq 0$, and every finite sequence of natural numbers $n_0, \ldots, n_{R-1}, n_R$, the consecutive pair of elements $n_0 \, \ldots \, n_{R-1} \, (n_R + 1) \, \overline{0}$ and $n_0 \, \ldots \, n_{R-1} \, n_R \, \infty$ of $\mathcal{N}$ is called a *rational-finite pair.*

Thus every element of $\mathcal{N}$ of rational type apart from $\overline{0}$, and every element of finite type apart from $\infty$, belongs to a rational-finite pair.

It will be convenient to extend $\Phi\colon \Delta' \to \mathbb{N}^{\mathbb{N}}$ to a (still lower semi-continuous) function $\Phi\colon \Delta \to \mathcal{N}$ by setting $\Phi(\boldsymbol{\alpha}) = \infty$ if $\boldsymbol{\alpha} \in \mathcal{F} = \Delta \setminus \Delta'$; and hence to extend $\mathcal{I}\colon \Delta' \to \mathcal{M}$ to a function $\mathcal{I}\colon \Delta \to \mathcal{M}$ by $\mathcal{I} = S \circ \Phi$. This extension has no dynamical significance (in particular, $\mathcal{I}(\boldsymbol{\alpha}) = (k-1)\overline{0}$ when $\boldsymbol{\alpha} \in \Delta \setminus \Delta'$ is not an $\boldsymbol{\alpha}$-infimax), but will make the statements of some results cleaner.

## 3. Digit frequency sets of symbolic $\beta$-shifts

3.1. **Preliminaries.** Let $w \in \Sigma$. The *symbolic $\beta$-shift* associated to $w$ is the subshift $\sigma\colon X(w) \to X(w)$, where

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

Since the supremum of any shift-invariant subset of $\Sigma$ is a maximal sequence, and since moreover $X(w) = X(\sup X(w))$, it suffices to consider the case where $w \in \mathcal{M}$, which we henceforth assume. We shall also assume that $w$ contains (and hence starts with) the digit $k-1$, since otherwise we could decrease the value of $k$.

**Definitions 15** ($\mathcal{M}^*$, $\mathrm{DF}(w)$ for $w \in \mathcal{M}^*$)**.** Denote by $\mathcal{M}^*$ the set of elements of $\mathcal{M}$ which start with the digit $k-1$, and write

$$\mathrm{DF}(w) = \{\boldsymbol{\alpha} \in \Delta \, : \, X(w) \cap \mathcal{R}(\boldsymbol{\alpha}) \neq \emptyset\} \subset \Delta$$

for each $w \in \mathcal{M}^*$, the set of digit frequencies of elements of $X(w)$.

The following lemma is the fundamental result which connects digit frequency sets to the infimaxes of Section 2.

**Lemma 16.** *Let $w \in \mathcal{M}^*$. Then*

$$\mathrm{DF}(w) = \{\boldsymbol{\alpha} \in \Delta \, : \, \mathcal{I}(\boldsymbol{\alpha}) \leq w\}.$$

*Proof.* Let $\boldsymbol{\alpha} \in \Delta$. First note that if $\alpha_{k-1} = 0$ then $\mathcal{I}(\boldsymbol{\alpha}) = (k-1)\,\overline{0} \leq w$ for all $w \in \mathcal{M}^*$, and $\boldsymbol{\alpha} \in \mathrm{DF}(w)$ since every sequence which doesn't contain the digit $k-1$ belongs to $X(w)$. We can therefore assume that $\alpha_{k-1} \neq 0$, i.e. that $\boldsymbol{\alpha} \in \Delta'$.

If $\mathcal{I}(\boldsymbol{\alpha}) < w$ then, since $\mathcal{I}(\boldsymbol{\alpha}) = \inf \mathcal{M}(\boldsymbol{\alpha})$ (Facts 1a)), there is some $v \in \mathcal{M}(\boldsymbol{\alpha})$ with $v < w$, and hence $v \in X(w)$. Therefore $\boldsymbol{\alpha} \in \mathrm{DF}(w)$.

If $\mathcal{I}(\boldsymbol{\alpha}) > w$ then every $v \in \mathcal{R}(\boldsymbol{\alpha})$ satisfies $w < \mathcal{I}(\boldsymbol{\alpha}) \leq \sup_{r \geq 0} \sigma^r(v)$ by Facts 1b), so that $v \notin X(w)$. Therefore $X(w) \cap \mathcal{R}(\boldsymbol{\alpha}) = \emptyset$, and hence $\boldsymbol{\alpha} \notin \mathrm{DF}(w)$.

Suppose then that $\mathcal{I}(\boldsymbol{\alpha}) = w$. We shall construct an element $v$ of $X(w) \cap \mathcal{R}(\boldsymbol{\alpha})$. We can assume that $\boldsymbol{\alpha}$ is exceptional, and in particular is not rational, since otherwise we can take $v = w$ by Facts 1f). Write $\mathbf{n} = \Phi(\boldsymbol{\alpha})$.

By Lemma 5, $\boldsymbol{\alpha}$ can be approximated arbitrarily closely by rational elements of $\mathrm{DF}(w)$: for the simplex $\Phi^{-1}(\mathbf{n})$, which contains $\boldsymbol{\alpha}$, can be approximated arbitrarily closely by the simplices $\mathcal{F}_{\mathbf{n},r}$, which have rational vertices. Moreover, any element $\boldsymbol{\beta}$ of $\mathcal{F}_{\mathbf{n},r}$ has itinerary $\Phi(\boldsymbol{\beta}) < \mathbf{n}$ by (3), so that $\mathcal{I}(\boldsymbol{\beta}) < \mathcal{I}(\boldsymbol{\alpha}) = w$, and hence $\boldsymbol{\beta} \in \mathrm{DF}(w)$ by the first part of the proof.

The idea of the construction is to pick a sequence $(\boldsymbol{\alpha}_n)$ of rational elements of $\mathrm{DF}(w)$ with $\boldsymbol{\alpha}_n \to \boldsymbol{\alpha}$, and a sequence of positive integers $(M_n)$, such that, writing $B_n$ for the repeating block of $\mathcal{I}(\boldsymbol{\alpha}_n)$, the sequence

$$v = B_1^{M_1}\, B_2^{M_2}\, B_3^{M_3}\, \dots \tag{6}$$

lies in $\mathcal{R}(\boldsymbol{\alpha})$. Since it also lies in $X(w) = X(\mathcal{I}(\boldsymbol{\alpha}))$ by Lemma 7, this will establish that $\boldsymbol{\alpha} \in \mathrm{DF}(w)$ as required.

Pick any strictly decreasing sequence $\epsilon_n \to 0$, and choose the sequence $(\boldsymbol{\alpha}_n)$ so that $d(\boldsymbol{\alpha}_n, \boldsymbol{\alpha}) < \epsilon_n$. Write $\boldsymbol{\alpha}_n = \mathbf{p}_n/q_n$, where $\mathbf{p}_n$ is an integer vector and $q_n = |B_n|$. Choose the positive integers $M_n$ inductively to satisfy $M_1 = 1$ and, for $n > 1$,

$$\frac{\max\left(q_{n+1}, \sum_{\ell=1}^{n-1} M_\ell q_\ell\right)}{\sum_{\ell=1}^{n} M_\ell q_\ell} < \frac{\epsilon_n}{\sqrt{k}}. \tag{7}$$

Define $v \in \Sigma$ by (6). We specify each initial subword of $v$ using a triple of integers $(n, i, j)$ with $n \geq 0$, $0 \leq i < M_{n+1}$, and $0 \leq j < q_{n+1}$: the subword $V(n, i, j)$ of $v$ has length $j + iq_{n+1} + \sum_{\ell=1}^{n} M_\ell q_\ell$: that is,

$$V(n,i,j) = B_1^{M_1}\, \dots\, B_n^{M_n}\, B_{n+1}^{i} B_{n+1}\, [\![j]\!]\,.$$

We shall show that, for every such triple, the digit frequency vector $\mathrm{df}(V(n,i,j))$ of $V(n,i,j)$ is within distance $6\epsilon_n$ of $\boldsymbol{\alpha}$, which will establish that $v \in \mathcal{R}(\boldsymbol{\alpha})$ as required.

To simplify notation, write $\mathbf{a}_n = \sum_{\ell=1}^{n} M_\ell \mathbf{p}_\ell$ and $b_n = \sum_{\ell=1}^{n} M_\ell q_\ell$, so that $\mathbf{a}_n/b_n = \mathrm{df}(V(n,0,0))$. We will use the following simple estimate: if $\mathbf{A}/B, \mathbf{C}/D \in \Delta_k$, where $\mathbf{A}, \mathbf{C} \in \mathbb{N}^k$ and $B, D \in \mathbb{N}$, then $(\mathbf{A}+\mathbf{C})/(B+D) - \mathbf{A}/B = D(\mathbf{C}/D - \mathbf{A}/B)/(B+D)$, and hence

$$d\left(\frac{\mathbf{A}}{B}, \frac{\mathbf{A}+\mathbf{C}}{B+D}\right) \leq \sqrt{k}\, \frac{D}{B+D}. \tag{8}$$

Now

$$\begin{aligned} d(\boldsymbol{\alpha}, \mathrm{df}(V(n,i,j))) \leq{}& d(\boldsymbol{\alpha}, \boldsymbol{\alpha}_n) + d(\boldsymbol{\alpha}_n, \mathrm{df}(V(n,0,0))) \\ &+ d(\mathrm{df}(V(n,0,0)), \mathrm{df}(V(n,i,0))) + d(\mathrm{df}(V(n,i,0)), \mathrm{df}(V(n,i,j))), \end{aligned}$$

and we estimate each term.

For the first, $d(\boldsymbol{\alpha}, \boldsymbol{\alpha}_n) < \epsilon_n$ by choice of the sequence $(\boldsymbol{\alpha}_n)$.

For the second,

$$d(\boldsymbol{\alpha}_n, \mathrm{df}(V(n,0,0))) = d(\mathbf{p}_n/q_n, \mathbf{a}_n/b_n) = d\left(\frac{M_n\mathbf{p}_n}{M_nq_n}, \frac{\mathbf{a}_{n-1} + M_n\mathbf{p}_n}{b_{n-1} + M_nq_n}\right) \leq \sqrt{k}\, \frac{b_{n-1}}{b_{n-1} + M_nq_n} < \epsilon_n \tag{9}$$

using (8) and (7).

For the third, we have

$$\mathrm{df}(V(n,i,0)) = \left(\frac{b_n}{b_n + iq_{n+1}}\right) \frac{\mathbf{a}_n}{b_n} + \left(\frac{iq_{n+1}}{b_n + iq_{n+1}}\right) \frac{\mathbf{p}_{n+1}}{q_{n+1}},$$

which lies on the line segment with endpoints $\mathbf{a}_n/b_n$ and $\boldsymbol{\alpha}_{n+1}$. Since $\mathrm{df}(V(n,0,0)) = \mathbf{a}_n/b_n$, the third term is bounded above by $d(\mathbf{a}_n/b_n, \boldsymbol{\alpha}_n) + d(\boldsymbol{\alpha}_n, \boldsymbol{\alpha}_{n+1})$, which is less than $3\epsilon_n$ by (9) and choice of the sequence $(\boldsymbol{\alpha}_n)$.

Finally, writing $\mathrm{df}(B_{n+1}\llbracket j\rrbracket) = \mathbf{r}_{n,j}/j$, the fourth term is

$$d(\mathrm{df}(V(n,i,0)),\mathrm{df}(V(n,i,j))) = d\left(\frac{\mathbf{a}_n + i\mathbf{p}_{n+1}}{b_n + iq_{n+1}}, \frac{\mathbf{a}_n + i\mathbf{p}_{n+1} + \mathbf{r}_{n,j}}{b_n + iq_{n+1} + j}\right) < \sqrt{k}\,\frac{j}{b_n + iq_{n+1}} \leq \sqrt{k}\,\frac{j}{b_n} < \epsilon_n$$

by (8) and (7), since $j < q_{n+1}$. This completes the proof. □

**Corollary 17.** $\mathrm{DF}(w)$ *is compact for all* $w \in \mathcal{M}^*$.

*Proof.* Immediate from Lemma 16 and the lower semi-continuity of $\mathcal{I} = S \circ \Phi$ (Facts 1d) and e)). □

*Remark* 18. If we were to define the digit frequency set subsequentially, by

$$\mathrm{DF}'(w) = \{\boldsymbol{\alpha} \in \Delta \,:\, X(w) \cap \mathcal{R}'(\boldsymbol{\alpha}) \neq \emptyset\},$$

then the proof of Lemma 16 goes through, using the “primed” versions of Facts 1a) and b) (see Facts 1g)), to show that $\mathrm{DF}'(w) = \{\boldsymbol{\alpha} \in \Delta \,:\, \mathcal{I}(\boldsymbol{\alpha}) \leq w\}$. That is, the digit frequency set isn't sensitive to whether or not it is defined subsequentially. We shall see (Lemma 43) that the digit frequency set $\mathrm{DF}(\beta)$, where $\beta \in (k-1,k]$, can be written as $\mathrm{DF}(w)$ for an appropriate $w \in \mathcal{M}^*$, so that it too can be defined subsequentially without affecting its value.

A first consequence of Lemma 16 is that we can restrict attention to $w \in \mathcal{J}$ rather than all $w \in \mathcal{M}^*$, as expressed by the next lemma.

**Lemma 19.** *Let* $w \in \mathcal{M}^*$*, and let* $s = \max\{t \in \mathcal{J} \,:\, t \leq w\}$*. Then* $\mathrm{DF}(s) = \mathrm{DF}(w)$.

*Proof.* Note that the maximum exists by Lemmas 9 and 11. Moreover, for each $\boldsymbol{\alpha} \in \Delta$ we have, by Lemma 16,

$$\boldsymbol{\alpha} \in \mathrm{DF}(w) \iff \mathcal{I}(\boldsymbol{\alpha}) \leq w \iff \mathcal{I}(\boldsymbol{\alpha}) \in \{t \in \mathcal{J} \,:\, t \leq w\} \iff \mathcal{I}(\boldsymbol{\alpha}) \leq s \iff \boldsymbol{\alpha} \in \mathrm{DF}(s).$$

□

*Remark* 20. Lemma 19 means that, as $w$ increases, the set $\mathrm{DF}(w)$ undergoes bifurcations only as $w$ passes through elements of $\mathcal{J}$. In particular, $\mathrm{DF}(w)$ mode locks as $w$ passes through each complementary gap of the Cantor set $\mathcal{J}$. In fact, if $w \notin \mathcal{J}$ then $s = \max\{t \in \mathcal{J} \,:\, t \leq w\}$ is of rational type by Lemmas 11 and 13, and we will see (Theorem 33) that $\mathrm{DF}(s)$ is a polytope with rational vertices in this case.

In view of the fact that $S\colon \mathcal{N} \to \mathcal{J}$ is an order-preserving homeomorphism (Lemma 11), we can equally well study the sets $\mathrm{DF}(w)$, where $w \in \mathcal{J}$, in terms of itineraries, motivating the following definition.

**Definition 21** ($\mathrm{DF}(\mathbf{n})$ for $\mathbf{n} \in \mathcal{N}$)**.** For each $\mathbf{n} \in \mathcal{N}$, define

$$\mathrm{DF}(\mathbf{n}) := \mathrm{DF}(S(\mathbf{n})).$$

Then Lemma 16 reads

$$\mathrm{DF}(\mathbf{n}) = \{\boldsymbol{\alpha} \in \Delta \,:\, \Phi(\boldsymbol{\alpha}) \leq \mathbf{n}\}. \tag{10}$$

For $\boldsymbol{\alpha} \in \mathrm{DF}(\mathbf{n}) \iff \boldsymbol{\alpha} \in \mathrm{DF}(S(\mathbf{n})) \iff \mathcal{I}(\boldsymbol{\alpha}) \leq S(\mathbf{n}) \iff \Phi(\boldsymbol{\alpha}) \leq \mathbf{n}$, since $\mathcal{I} = S \circ \Phi$ and $S$ is an order-preserving homeomorphism.

*Remark* 22. We have now defined digit frequency sets for three types of objects: numbers $\beta \in (k-1,k]$, digit sequences $w \in \mathcal{M}^*$, and itineraries $\mathbf{n} \in \mathcal{N}$. In each case the collection of digit frequency sets is the same, with the exception that $\mathcal{F}$ can be realised as the digit frequency set of the element $(k-1)\,\overline{0}$ of $\mathcal{M}^*$ and of the element of $\infty$ of $\mathcal{N}$, but not of an element of $(k-1,k]$ – although it is, of course, the digit frequency set of $k-1$. To show that the collections are otherwise the same, it is enough to observe (see Definitions 41 and 42 and Lemma 43) that there are digit frequency set preserving maps

$$\begin{aligned} \beta \mapsto w_\beta \quad &: \quad (k-1, k] \to \mathcal{M}^* \setminus \{(k-1)\,\overline{0}\}, \\ w \mapsto \max\{\mathbf{n} \in \mathcal{N} \,:\, S(\mathbf{n}) \le w\} \quad &: \quad \mathcal{M}^* \setminus \{(k-1)\,\overline{0}\} \to \mathcal{N} \setminus \{\infty\}, \text{ and} \\ \mathbf{n} \mapsto \beta(\mathbf{n}) \quad &: \quad \mathcal{N} \setminus \{\infty\} \to (k-1, k]. \end{aligned}$$

Using (10) and Lemma 13, together with the fact that elements of $\mathcal{N}$ of finite type, other than $\infty$, are not in the image of $\Phi$, we have

$$\mathrm{DF}(n_0\,\ldots\,n_{R-1}\,(n_R+1)\,\overline{0}) = \mathrm{DF}(n_0\,\ldots\,n_{R-1}\,n_R\,\infty) \tag{11}$$

for all $n_0, \ldots, n_R$: the digit frequency set of an element of finite type other than $\infty$ is the same as that of its rational pair. In particular, when describing the different possible structures of $\mathrm{DF}(\mathbf{n})$ as $\mathbf{n}$ varies in $\mathcal{N}$, we can restrict to the case $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$.

### 3.2. **Convexity of the digit frequency set.**

**Definition 23** ($T_n$)**.** For each $n \ge 0$, let $T_n$ be the simplex

$$T_n = \{\boldsymbol{\alpha} \in \Delta \,:\, (n+1)\,\alpha_{k-1} \le \alpha_0\} \subset \Delta,$$

the union of the simplices $\Delta_m$ for $m > n$ together with the face $\alpha_{k-1} = 0$; or, equivalently, the set of $\boldsymbol{\alpha}$ with $\Phi(\boldsymbol{\alpha})_0 > n$ (perhaps $\Phi(\boldsymbol{\alpha})_0 = \infty$).

The following lemma describes the structure of digit frequency sets recursively.

**Lemma 24.** *Let* $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$. *Then*

$$\mathrm{DF}(\mathbf{n}) = T_{n_0} \cup K_{n_0}^{-1}(\mathrm{DF}(\sigma(\mathbf{n}))),$$

*where* $\sigma\colon \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ *is the shift map.*

*Proof.* $T_{n_0} \subset \mathrm{DF}(\mathbf{n})$, since $\boldsymbol{\alpha} \in T_{n_0} \implies \Phi(\boldsymbol{\alpha})_0 > n_0 \implies \Phi(\boldsymbol{\alpha}) < \mathbf{n}$.

If $\boldsymbol{\alpha} \in \mathrm{DF}(\mathbf{n}) \setminus T_{n_0}$ then $\Phi(\boldsymbol{\alpha})_0 = n_0$, and for all $\boldsymbol{\alpha}$ with $\Phi(\boldsymbol{\alpha})_0 = n_0$ we have

$$\boldsymbol{\alpha} \in \mathrm{DF}(\mathbf{n}) \iff \Phi(\boldsymbol{\alpha}) \le \mathbf{n} \iff \sigma(\Phi(\boldsymbol{\alpha})) \le \sigma(\mathbf{n}) \iff \Phi(K_{n_0}(\boldsymbol{\alpha})) \le \sigma(\mathbf{n}) \iff K_{n_0}(\boldsymbol{\alpha}) \in \mathrm{DF}(\sigma(\mathbf{n})).$$

□

We now define useful sequences of subsets and supersets of $\mathrm{DF}(\mathbf{n})$.

**Definition 25** ($L_{\mathbf{n},r}$, $U_{\mathbf{n},r}$)**.** For each $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ and $r \in \mathbb{N}$, write

$$L_{\mathbf{n},r} = \bigcup_{s=0}^{r} K_{n_0}^{-1} \circ \cdots \circ K_{n_{s-1}}^{-1}(T_{n_s}) \qquad \text{and} \qquad U_{\mathbf{n},r} = L_{\mathbf{n},r} \cup A_{\mathbf{n},r}.$$

**Lemma 26.** *Let* $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$. *Then* $L_{\mathbf{n},r} \subset \mathrm{DF}(\mathbf{n}) \subset U_{\mathbf{n},r}$ *and* $L_{\mathbf{n},r} \cap A_{\mathbf{n},r} = \mathcal{F}_{\mathbf{n},r}$ *for all* $r \in \mathbb{N}$. *Moreover* $\mathrm{DF}(\mathbf{n}) = \bigcap_{r \ge 0} U_{\mathbf{n},r}$.

*Proof.* $K_{n_0}^{-1} \circ \cdots \circ K_{n_{s-1}}^{-1}(T_{n_s}) = K_{n_0}^{-1} \circ \cdots \circ K_{n_{s-1}}^{-1}(\mathcal{F}) \cup K_{n_0}^{-1} \circ \cdots \circ K_{n_{s-1}}^{-1}(T_{n_s} \setminus \mathcal{F})$ is the union of $\mathcal{F}_{\mathbf{n},s-1}$ (which is contained in $\mathrm{DF}(\mathbf{n})$ using (3) and that $\mathcal{F} \subset \mathrm{DF}(\mathbf{n})$) and of $\{\boldsymbol{\alpha} \in \Delta \,:\, \Phi(\boldsymbol{\alpha})\,[\![s]\!] = \mathbf{n}\,[\![s]\!]$ and $\Phi(\boldsymbol{\alpha})_s > n_s\}$ (which is contained in $\mathrm{DF}(\mathbf{n})$ by (10)). This establishes the lower bound.

Moreover, since $L_{\mathbf{n},r}$ contains $\{\boldsymbol{\alpha} \in \Delta \,:\, \Phi(\boldsymbol{\alpha})\,[\![s]\!] = \mathbf{n}\,[\![s]\!]$ and $\Phi(\boldsymbol{\alpha})_s > n_s\}$ for each $s$ with $0 \le s \le r$, we have that

$$L_{\mathbf{n},r} \supset \{\boldsymbol{\alpha} \in \Delta \,:\, \Phi(\boldsymbol{\alpha})\,[\![r+1]\!] < \mathbf{n}\,[\![r+1]\!]\}.$$

Since $A_{\mathbf{n},r} = \{\boldsymbol{\alpha} \in \Delta : \Phi(\boldsymbol{\alpha})\,[\![r+1]\!] = \mathbf{n}\,[\![r+1]\!]\} \cup \mathcal{F}_{\mathbf{n},r}$ it follows that $U_{\mathbf{n},r}$ contains the set $\{\boldsymbol{\alpha} \in \Delta : \Phi(\boldsymbol{\alpha})\,[\![r+1]\!] \le \mathbf{n}\,[\![r+1]\!]\}$, which contains $\mathrm{DF}(\mathbf{n})$, establishing the upper bound. Moreover if $\Phi(\boldsymbol{\alpha})\,[\![r+1]\!] = \mathbf{n}\,[\![r+1]\!]$ then $\boldsymbol{\alpha} \notin L_{\mathbf{n},r}$, so that $L_{\mathbf{n},r} \cap A_{\mathbf{n},r} \subset \mathcal{F}_{\mathbf{n},r}$; and points $\boldsymbol{\alpha}$ of $\mathcal{F}_{\mathbf{n},r}$ either lie in $\mathcal{F} \subset T_{n_0} \subset L_{\mathbf{n},r}$ or have $\Phi(\boldsymbol{\alpha})\,[\![r+1]\!] < \mathbf{n}\,[\![r+1]\!]$ by (3), so that $\mathcal{F}_{\mathbf{n},r} \subset L_{\mathbf{n},r} \cap A_{\mathbf{n},r}$ as required.

To show that $\mathrm{DF}(\mathbf{n}) = \bigcap_{r\ge 0} U_{\mathbf{n},r}$, observe that any $\boldsymbol{\alpha} \in \bigcap_{r\ge 0} U_{\mathbf{n},r}$ either lies in $L_{\mathbf{n},r}$ for some $r$, and hence in $\mathrm{DF}(\mathbf{n})$; or lies in $\bigcap_{r\ge 0} A_{\mathbf{n},r} = \Phi^{-1}(\mathbf{n})$, so that $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$ and $\boldsymbol{\alpha} \in \mathrm{DF}(\mathbf{n})$. □

**Theorem 27.** *For all $\mathbf{n} \in \mathcal{N}$, the digit frequency set $\mathrm{DF}(\mathbf{n})$ is convex.*

*Proof.* By Lemma 26 we need only show that $U_{\mathbf{n},r}$ is convex for all $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ and all $r \in \mathbb{N}$, which we do by induction on $r$. In the base case $r = 0$, we have $U_{\mathbf{n},0} = T_{n_0} \cup K_{n_0}^{-1}(\Delta) = T_{n_0-1}$, which is convex (we write $T_{-1} = \Delta$ for convenience).

For $r > 0$, we have $L_{\mathbf{n},r} = T_{n_0} \cup K_{n_0}^{-1}(L_{\sigma(\mathbf{n}),r-1})$ and $A_{\mathbf{n},r} = K_{n_0}^{-1}(A_{\sigma(\mathbf{n}),r-1})$, so that

$$U_{\mathbf{n},r} = T_{n_0} \cup K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1}).$$

Now $U_{\sigma(\mathbf{n}),r-1}$ is convex by the inductive hypothesis, and hence so is $K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1})$. Moreover,

$$\mathrm{Cl}\,(\Delta_{n_0}) \supset K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1}) \supset K_{n_0}^{-1}(\mathrm{DF}(\sigma(\mathbf{n}))) \supset K_{n_0}^{-1}(\mathcal{F}) = \mathcal{F}_{\mathbf{n},0},$$

and $\mathcal{F}_{\mathbf{n},0}$ is also a face of $T_{n_0}$. That is, the simplex $T_{n_0-1}$ is the union of the two simplices $T_{n_0}$ and $\mathrm{Cl}\,(\Delta_{n_0})$ whose intersection is the common face $\mathcal{F}_{\mathbf{n},0}$; and $U_{\mathbf{n},r}$ is the union of $T_{n_0}$ and the convex subset $K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1})$ of $\mathrm{Cl}\,(\Delta_{n_0})$ which contains $\mathcal{F}_{\mathbf{n},0}$. It follows that $U_{\mathbf{n},r}$ is convex as required, since any line segment joining a point of $T_{n_0}$ to a point of $K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1})$ passes through $\mathcal{F}_{\mathbf{n},0}$, and is therefore the join of a segment in $T_{n_0}$ and a segment in $K_{n_0}^{-1}(U_{\sigma(\mathbf{n}),r-1})$. □

*Remark* 28. As a consequence, we obtain the following property of the itinerary map $\Phi$: if $\ell \subset \Delta$ is a line segment with endpoints $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, and $\boldsymbol{\gamma} \in \ell$, then $\Phi(\boldsymbol{\gamma}) \le \max(\Phi(\boldsymbol{\alpha}), \Phi(\boldsymbol{\beta}))$. For if not then, since $\max(\Phi(\boldsymbol{\alpha}), \Phi(\boldsymbol{\beta}))$ and $\Phi(\boldsymbol{\gamma})$ are not the elements of a rational-finite pair (the only finite type element in the image of $\Phi$ is $\infty$), we can choose $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ with $\max(\Phi(\boldsymbol{\alpha}), \Phi(\boldsymbol{\beta})) < \mathbf{n} < \Phi(\boldsymbol{\gamma})$. Then $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathrm{DF}(\mathbf{n})$ but $\boldsymbol{\gamma} \notin \mathrm{DF}(\mathbf{n})$, contradicting the convexity of $\mathrm{DF}(\mathbf{n})$.

The convexity of $\mathrm{DF}(\mathbf{n})$ is in turn an immediate consequence of this "line segment" property. An alternative approach to establishing that $\mathrm{DF}(\mathbf{n})$ is convex is to prove the line segment property directly from the dynamics of the continued fraction map $K$.

3.3. **Extreme points of the digit frequency set.** By Theorem 27, $\mathrm{DF}(\mathbf{n})$ is determined by the set of its extreme points, which we now describe. Observe first that since $\mathcal{F} \subset \mathrm{DF}(\mathbf{n})$, the vertices $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{k-2}$ of $\mathcal{F}$ are extreme points of $\mathrm{DF}(\mathbf{n})$ for all $\mathbf{n}$.

**Definitions 29** ($\mathrm{EP}(\mathbf{n})$, $\mathrm{E}(\mathbf{n})$, non-trivial extreme points)**.** Let $\mathrm{EP}(\mathbf{n})$ denote the set of extreme points of $\mathrm{DF}(\mathbf{n})$, and

$$\mathrm{E}(\mathbf{n}) = \mathrm{EP}(\mathbf{n}) \setminus \{\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_{k-2}\}$$

the set of *non-trivial* extreme points.

The next lemma translates the inductive expression for $\mathrm{DF}(\mathbf{n})$ given by Lemma 24 into an analogous one for $\mathrm{E}(\mathbf{n})$. Notice that in the case $k = 2$ the condition "$n_r = 0$ for $1 \le r \le k-2$" is always true.

**Lemma 30.** *Let $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$. Then*

$$\mathrm{E}(\mathbf{n}) = \begin{cases} K_{n_0}^{-1}(\mathrm{E}(\sigma(\mathbf{n}))) & \text{if } n_r = 0 \text{ for } 1 \le r \le k-2, \\ K_{n_0}^{-1}(\mathrm{E}(\sigma(\mathbf{n})) \cup \{\mathbf{e}_{k-2}\}) & \text{otherwise.} \end{cases}$$

*Proof.* By Lemma 24, $\mathrm{DF}(\mathbf{n})$ is the union of the simplex $T_{n_0}$ and the image of $\mathrm{DF}(\sigma(\mathbf{n}))$ under the projective homeomorphism $K_{n_0}^{-1}$. Since the intersection of these sets is exactly $\mathcal{F}_{\mathbf{n},0}$, which has vertices $\mathbf{e}_1,\ldots,\mathbf{e}_{k-2}$ and $K_{n_0}^{-1}(\mathbf{e}_{k-2})$, the non-trivial extreme points of $\mathrm{DF}(\mathbf{n})$ are the images under $K_{n_0}^{-1}$ of the non-trivial extreme points of $\mathrm{DF}(\sigma(\mathbf{n}))$, together perhaps with the point $K_{n_0}^{-1}(\mathbf{e}_{k-2})$.

Now if $K_{n_0}^{-1}(\mathbf{e}_{k-2}) = ((n_0+1)/(n_0+2), 0, \ldots, 0, 1/(n_0+2))$ is on a line segment joining two other points of $\mathrm{DF}(\mathbf{n})$, then these points must have coordinates $(1-x, 0, \ldots, 0, x)$ and $(1-y, 0, \ldots, 0, y)$ with $x < 1/(n_0+2) < y$. Since $\mathbf{e}_0 = (1, 0, \ldots, 0, 0) \in \mathrm{DF}(\mathbf{n})$, it follows that $K_{n_0}^{-1}(\mathbf{e}_{k-2})$ is *not* an extreme point of $\mathrm{DF}(\mathbf{n})$ if and only if $K_{n_0}^{-1}(\mathrm{DF}(\sigma(\mathbf{n})))$ contains a point of the form $(1-y, 0, \ldots, 0, y)$ for some $y$ with $1/(n_0+2) < y \le 1/(n_0+1)$ (the latter inequality coming from the fact that $(1-y, 0, \ldots, 0, y)$ can only be in $K_{n_0}^{-1}(\Delta)$ if $(1-y)/y \ge n_0$).

Since $K_{n_0}(1-y, 0, \ldots, 0, y) = \left(0, \ldots, 0, \frac{1}{y} - (n_0+1), (n_0+2) - \frac{1}{y}\right)$ it follows that

$$
\begin{aligned}
K_{n_0}^{-1}(\mathbf{e}_{k-2}) \notin \mathrm{E}(\mathbf{n}) \quad &\iff \quad \exists z \in (0,1],\ (0, 0, \ldots, 0, 1-z, z) \in \mathrm{DF}(\sigma(\mathbf{n})) \\
&\iff \quad \exists z \in (0,1],\ \Phi(0, 0, \ldots, 0, 1-z, z) \le \sigma(\mathbf{n}).
\end{aligned}
$$

Now $\Phi(0, 0, \ldots, 0, 1-z, z)_r = 0$ for $0 \le r \le k-3$, so that

$$\exists z \in (0,1],\ \Phi(0, 0, \ldots, 0, 1-z, z) \le \sigma(\mathbf{n}) \implies n_r = 0 \text{ for } 1 \le r \le k-2.$$

Conversely, if $n_r = 0$ for $1 \le r \le k-2$, then let $z = 1/(n_{k-1}+2) \in (0,1]$. We have

$$K^{k-2}(0, 0, \ldots, 0, 1-z, z) = (1-z, 0, 0, \ldots, 0, z) \in \Delta_{\lfloor (1-z)/z \rfloor} = \Delta_{n_{k-1}+1},$$

so that $\Phi(0, 0, \ldots, 0, 1-z, z) \le \sigma(\mathbf{n})$.

Therefore $K_{n_0}^{-1}(\mathbf{e}_{k-2}) \notin \mathrm{E}(\mathbf{n})$ if and only if $n_r = 0$ for $1 \le r \le k-2$, as required. □

The following theorem describes the set of non-trivial extreme points of $\mathrm{DF}(\mathbf{n})$ non-inductively: this set consists of the points $\Upsilon_{\mathbf{n},s}(\mathbf{e}_{k-2})$ for those indices $s$ which are not followed by $k-2$ zeroes in the itinerary $\mathbf{n}$, together with a subset of the vertices of $\Phi^{-1}(\mathbf{n})$. In the regular case, this subset is exactly the one-point set $\Phi^{-1}(\mathbf{n})$; while in the exceptional case, since $\Phi^{-1}(\mathbf{n})$ is a $j$-simplex for some $j \le k-2$, it contains at most $k-1$ points.

**Definition 31** ($\mathrm{E}'(\mathbf{n})$)**.** For each $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$, write

$$\mathrm{E}'(\mathbf{n}) = \{K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \,:\, s \ge 0,\ n_{s+t} \ne 0 \text{ for some } 1 \le t \le k-2\}.$$

*Remark* 32. Each point of $\mathrm{E}'(\mathbf{n})$ is listed once only: that is, whenever $s < t$ we have $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \ne K_{n_0}^{-1} \circ \cdots \circ K_{n_t}^{-1}(\mathbf{e}_{k-2})$. For otherwise we would have $K_{n_{s+1}}^{-1} \circ \cdots \circ K_{n_t}^{-1}(\mathbf{e}_{k-2}) = \mathbf{e}_{k-2}$; but it follows immediately from (2) that $(K_{n_{s+1}}^{-1} \circ \cdots \circ K_{n_t}^{-1}(\mathbf{e}_{k-2}))_{k-1} > 0$, which is a contradiction.

**Theorem 33.** *Let* $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$*. Then*

$$\mathrm{E}'(\mathbf{n}) \ \subset\ \mathrm{E}(\mathbf{n}) \ \subset\ \left(\mathrm{E}'(\mathbf{n}) \cup \mathrm{Vert}\,\Phi^{-1}(\mathbf{n})\right).$$

*Moreover,*

*a)* $\mathrm{E}(\mathbf{n})$ *is finite (i.e.* $\mathrm{DF}(\mathbf{n})$ *is a polytope) if and only if* $k = 2$ *or* $\mathbf{n}$ *is of rational type.*

*b) If* $\mathbf{n}$ *is regular, so that* $\Phi^{-1}(\mathbf{n}) = \{\boldsymbol{\alpha}\}$ *for some* $\boldsymbol{\alpha}$*, then* $\mathrm{E}(\mathbf{n}) = \mathrm{E}'(\mathbf{n}) \cup \{\boldsymbol{\alpha}\}$*, and* $\boldsymbol{\alpha}$ *is an accumulation point of* $\mathrm{E}(\mathbf{n})$ *if* $\mathrm{E}(\mathbf{n})$ *is infinite.*

*c) The points of* $\mathrm{E}'(\mathbf{n})$ *are rational, while those of* $\mathrm{Vert}\,\Phi^{-1}(\mathbf{n})$ *are non-rational unless* $\mathbf{n}$ *is of rational type.*

*Remarks* 34.

a) When $k=2$, $\mathrm{E}'(\mathbf{n})$ is always empty and $\mathbf{n}$ is always regular, so there is one non-trivial extreme point, the point of $\Phi^{-1}(\mathbf{n})$, together with the trivial extreme point $(1,0)$.
b) Let $n_r = 2^{2^{3r}}$. Then for each $k \geq 3$, the set $\mathrm{E}(\mathbf{n})$ has $k-1$ accumulation points. For it is shown in [6] that $\Phi^{-1}(\mathbf{n})$ is a simplex of dimension $k-2$. The easy part of the proof of Borovikov's theorem (see [19] Section 3.3 for a proof in English) shows that each of the $k-1$ vertices of $\Phi^{-1}(\mathbf{n})$ is an accumulation point of vertices of the simplices $A_{\mathbf{n},r}$ – and each of these vertices belongs to $\mathrm{E}(\mathbf{n})$, since $n_r > 0$ for all $r$.
c) The authors do not know whether or not $\mathrm{E}(\mathbf{n}) \neq \mathrm{E}'(\mathbf{n}) \cup \operatorname{Vert}\Phi^{-1}(\mathbf{n})$ is possible in the case when $\mathbf{n}$ is exceptional; nor whether or not it is possible for the points of $\mathrm{E}(\mathbf{n})$ to accumulate at a point of $\Phi^{-1}(\mathbf{n})$ other than a vertex (they can't accumulate at a point not in $\Phi^{-1}(\mathbf{n})$, since $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \in A_{\mathbf{n},s} \to \Phi^{-1}(\mathbf{n})$ as $s \to \infty$).

*Proof.* A straightforward induction on $R$ using Lemma 30 gives that, for each $R \geq 1$,

$$\begin{aligned}\mathrm{E}(\mathbf{n}) \quad &= \quad \{K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \,:\, 0 \leq s < R,\ n_{s+t} \neq 0 \text{ for some } 1 \leq t \leq k-2\} \\ &\cup \quad K_{n_0}^{-1} \circ \cdots \circ K_{n_{R-1}}^{-1}(\mathrm{E}(\sigma^R(\mathbf{n})).\end{aligned}$$

Now $K_{n_0}^{-1} \circ \cdots \circ K_{n_{R-1}}^{-1}(\mathrm{E}(\sigma^R(\mathbf{n})) \subset A_{\mathbf{n},R-1}$, and $\bigcap_{r\geq 0} A_{\mathbf{n},r} = \Phi^{-1}(\mathbf{n})$. Therefore

$$\mathrm{E}(\mathbf{n}) \supset \mathrm{E}'(\mathbf{n}) = \{K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \,:\, s \geq 0,\ n_{s+t} \neq 0 \text{ for some } 1 \leq t \leq k-2\},$$

and any elements of $\mathrm{E}(\mathbf{n})$ not in $\mathrm{E}'(\mathbf{n})$ are contained in $\Phi^{-1}(\mathbf{n})$. Since $\Phi^{-1}(\mathbf{n})$ is a simplex which is contained in $\mathrm{DF}(\mathbf{n})$, any remaining extreme points of $\mathrm{DF}(\mathbf{n})$ must be vertices of $\Phi^{-1}(\mathbf{n})$.

We now prove the remaining statements of the theorem.

a) By Remark 32, $\mathrm{E}(\mathbf{n})$ is finite if and only if there are only finitely many $s \geq 0$ with the property that $n_{s+t} \neq 0$ for some $1 \leq t \leq k-2$. This is always the case when $k=2$, but for $k \geq 3$ is true if and only if $n_s = 0$ for all sufficiently large $s$; i.e., if and only if $\mathbf{n}$ is of rational type.
b) By Remark 28, if $\mathbf{n}$ is regular then it is not possible for the unique point of $\Phi^{-1}(\mathbf{n})$ (which has itinerary $\mathbf{n}$) to belong to a line segment whose endpoints are in $\mathrm{DF}(\mathbf{n})\setminus\Phi^{-1}(\mathbf{n})$ (and so have itinerary less than $\mathbf{n}$). This point is therefore an extreme point of $\mathrm{DF}(\mathbf{n})$ as required. Moreover, $\Phi^{-1}(\mathbf{n})$ is necessarily an accumulation point of $\mathrm{E}(\mathbf{n})$ when $\mathrm{E}(\mathbf{n})$ is infinite, since $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(\mathbf{e}_{k-2}) \in A_{\mathbf{n},s} \to \Phi^{-1}(\mathbf{n})$ as $s \to \infty$.
c) The points of $\mathrm{E}'(\mathbf{n})$ are clearly rational. No point with itinerary $\mathbf{n}$ can be rational unless $\mathbf{n}$ is of rational type.

□

**Corollary 35.** *Let* $\mathbf{m}, \mathbf{n} \in \mathcal{N}$ *with* $\mathbf{m} \neq \mathbf{n}$. *Then* $\mathrm{DF}(\mathbf{m}) = \mathrm{DF}(\mathbf{n})$ *if and only if* $\mathbf{m}$ *and* $\mathbf{n}$ *form a rational-finite pair.*

*Proof.* If $\mathbf{m}$ and $\mathbf{n}$ form a rational-finite pair then $\mathrm{DF}(\mathbf{m}) = \mathrm{DF}(\mathbf{n})$ by (11). Otherwise there are both rational and non-rational elements of $\mathcal{N}$ between $\mathbf{m}$ and $\mathbf{n}$. Since DF is increasing by definition, it follows from Theorem 33a) (or, in the case $k=2$, from Theorem 33c)) that $\mathrm{DF}(\mathbf{m}) \neq \mathrm{DF}(\mathbf{n})$ as required. □

**Examples 36.** a) Let $\mathbf{n} = 2\,1\,0\,1\,\overline{0}$. In the case $k = 3$, the digit frequency set $\mathrm{DF}(\mathbf{n})$ has two trivial extreme points, $(1, 0, 0)$ and $(0, 1, 0)$, together with three non-trivial extreme points

$$\begin{aligned} K_2^{-1}(0,1,0) &= (3/4,\ 0,\ 1/4), \\ K_2^{-1}K_1^{-1}K_0^{-1}(0,1,0) &= (5/8,\ 1/8,\ 2/8), \quad \text{and} \\ \Phi^{-1}(\mathbf{n}) = K_2^{-1}K_1^{-1}K_0^{-1}K_1^{-1}(0,0,1) &= (4/9,\ 3/9,\ 2/9). \end{aligned}$$

Hence $\mathrm{DF}(\mathbf{n})$ is a pentagon (Figure 1). The point $K_2^{-1}K_1^{-1}(0,1,0) = (2/5,\ 2/5,\ 1/5)$ is not an extreme point since $n_2 = 0$ — and indeed it can be checked that it lies on the line segment joining the extreme points $(0, 1, 0)$ and $(4/9,\ 3/9,\ 2/9)$.

The same itinerary $\mathbf{n}$ in the case $k = 4$ gives $|\,\mathrm{EP}(\mathbf{n})| = 7$: in addition to the three trivial extreme points $(1, 0, 0, 0)$, $(0, 1, 0, 0)$, and $(0, 0, 1, 0)$, we have

$$\begin{aligned} K_2^{-1}(0,0,1,0) &= (3/4,\ 0,\ 0,\ 1/4), \\ K_2^{-1}K_1^{-1}(0,0,1,0) &= (2/5,\ 2/5,\ 0,\ 1/5), \\ K_2^{-1}K_1^{-1}K_0^{-1}(0,0,1,0) &= (2/5,\ 1/5,\ 1/5,\ 1/5), \quad \text{and} \\ \Phi^{-1}(\mathbf{n}) = K_2^{-1}K_1^{-1}K_0^{-1}K_1^{-1}(0,0,0,1) &= (5/8,\ 1/8,\ 0,\ 2/8). \end{aligned}$$

Thus $\mathrm{DF}(\mathbf{n})$ is a polyhedron with 7 vertices. In this case we do include the point $K_2^{-1}K_1^{-1}(0, 0, 1, 0)$, since it is not true that both $n_2 = 0$ and $n_3 = 0$.

b) When $k = 3$ and $\mathbf{n}$ is exceptional, the boundary of $\mathrm{DF}(\mathbf{n})$ contains the exceptional interval, both ends of which are accumulations of extreme points. Figure 4, drawn in the $(\alpha_0, \alpha_2)$-plane, illustrates the situation: it depicts $\mathrm{DF}(\mathbf{n})$, and its extreme points, in the case where $n_r = r^3$ for $0 \le r \le 25$, and $n_r = 0$ for $r > 25$. Notice that, if $\mathbf{N} \in \mathbb{N}^{\mathbb{N}}$ is given by $N_r = r^3$ for all $r$, then the first 25 extreme points in $\mathrm{E}'(\mathbf{N})$ provided by Theorem 33 are exactly the points of $\mathrm{E}'(\mathbf{n})$.

It is not known whether or not $\mathbf{N}$ is exceptional, although experimental evidence suggests that it is. However, itineraries which are known to be exceptional, such as $n_r = 2^{2^{3r}}$, grow too quickly for it to be feasible to produce plots similar to Figure 4.

Constrast Figure 4 with Figure 5, which shows an approximation to $\mathrm{DF}(\mathbf{n})$ when $n_r = r^2$, which is known to be regular. Here the sequence of rational extreme points limits on the unique (non-rational) $\boldsymbol{\alpha}$ with $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$.

c) Let $k = 3$, and pick $\boldsymbol{\alpha} \in \Delta'$ with $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$. In this example we consider the way in which the rational extreme points $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(0, 1, 0)$ (with $n_{s+1} \neq 0$) are ordered around the boundary of $\mathrm{DF}(\mathbf{n})$. Each such point either lies on the *even* segment of $\partial\,\mathrm{DF}(\mathbf{n})$, which joins $(1, 0, 0)$ to $\boldsymbol{\alpha}$ and does not contain $(0, 1, 0)$; or on the *odd* segment, which joins $(0, 1, 0)$ to $\boldsymbol{\alpha}$ and does not contain $(1, 0, 0)$. We claim that the extreme points $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(0, 1, 0)$ with $s$ even (respectively odd) are contained in the even (respectively odd) segment, and move monotonically along the segment towards $\boldsymbol{\alpha}$ as $s$ increases. In other words, in pictures such as those of Figures 1, 4, and 5, the points with $s$ even move counter-clockwise around the boundary starting at the bottom right vertex $(1, 0, 0)$, while those with $s$ odd move clockwise starting at the bottom left vertex $(0, 1, 0)$.

The claim can be proved in the case where $\boldsymbol{\alpha}$ is rational, with $\Phi(\boldsymbol{\alpha}) = n_0\, n_1\, \ldots\, n_r\, \overline{0}$, by induction on $r$, using Lemma 24 and the fact that each $K_n^{-1}$ is orientation-reversing. The result then follows in the non-rational case since the extreme points $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(0, 1, 0)$ on the boundary of $\mathrm{DF}(\mathbf{n})$ are the same points as those on the boundary of $\mathrm{DF}(\mathbf{m})$, where $\mathbf{m} = n_0\, n_1\, \ldots\, n_{s+1}\, \overline{0}$.

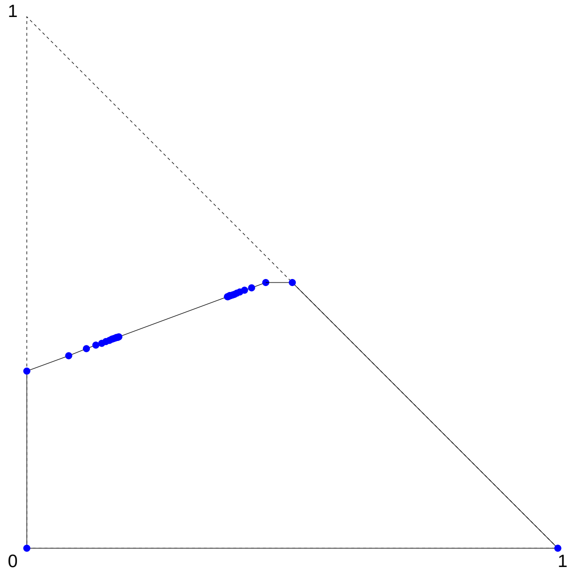


FIGURE 4. An approximation to $\mathrm{DF}(\mathbf{n})$ for exceptional $\mathbf{n}$ in the case $k = 3$

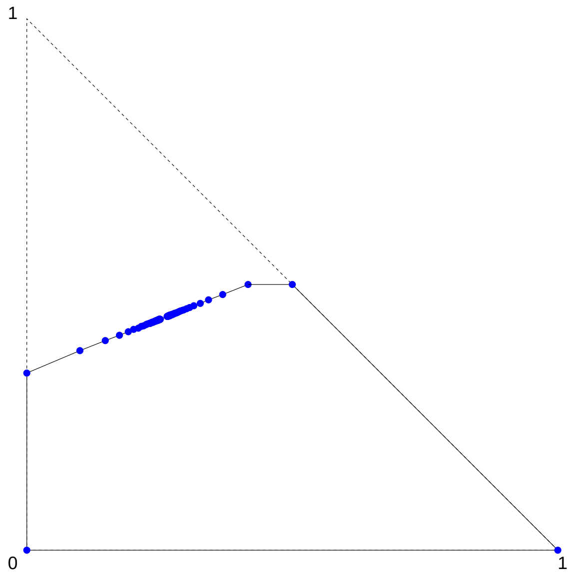


FIGURE 5. An approximation to $\mathrm{DF}(\mathbf{n})$ when $k = 3$ in the regular case $n_r = r^2$

In particular, if $\boldsymbol{\alpha}$ is regular then it is a limit of rational extreme points of $\mathrm{DF}(\mathbf{n})$ from both sides (as in Figure 5), unless there is some $s \in \mathbb{N}$ such that $n_{s+2i} = 0$ for all $i \in \mathbb{N}$ (so that none of the points $K_{n_0}^{-1} \circ \cdots \circ K_{n_{s+2i-1}}^{-1}(0,1,0)$ is an extreme point). By Facts 1i), such an $s$ exists if and only if $K^r(\boldsymbol{\alpha})$ lies on the boundary of $\Delta$ for some $r \in \mathbb{N}$. In this case, by the convexity of $\mathrm{DF}(\sigma^r(\mathbf{n}))$ and the fact that $\mathcal{F} \subset \mathrm{DF}(\sigma^r(\mathbf{n}))$, the boundary of $\mathrm{DF}(\sigma^r(\mathbf{n}))$ contains a segment $I$ with endpoints $K^r(\boldsymbol{\alpha})$ and either $(0,1,0)$ or $(1,0,0)$. This segment is contained in a face of $\Delta$, and so has rational direction. By Lemma 24, $K_{n_0}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1}(I)$ is a segment of the boundary of $\mathrm{DF}(\mathbf{n})$ which has one endpoint at $\boldsymbol{\alpha}$ and the other endpoint at a rational boundary point of $\mathrm{DF}(\mathbf{n})$.

In summary, if $\boldsymbol{\alpha}$ is regular then there is a segment of the boundary of $\mathrm{DF}(\mathbf{n})$ which contains $\boldsymbol{\alpha}$ if and only if every other entry of $\Phi(\boldsymbol{\alpha})$ is eventually zero; and in this case, the segment is necessarily contained in a line of rational direction which passes through rational vectors, so that the components of $\boldsymbol{\alpha}$ are rationally dependent. See Figure 6, which depicts $\mathrm{DF}(\mathbf{n})$ for $\mathbf{n} = 1\,1\,\overline{1\,0}$. The irrational extreme point $\boldsymbol{\alpha}$ is a limit of rational extreme points from the odd side only, since $n_s = 0$ for all odd $s \geq 3$. In fact $\boldsymbol{\alpha} = ((1+\sqrt{5})/6, (3-\sqrt{5})/6, 1/3)$, and the rational extreme point adjacent to $\boldsymbol{\alpha}$ is $(2/3, 0, 1/3)$.

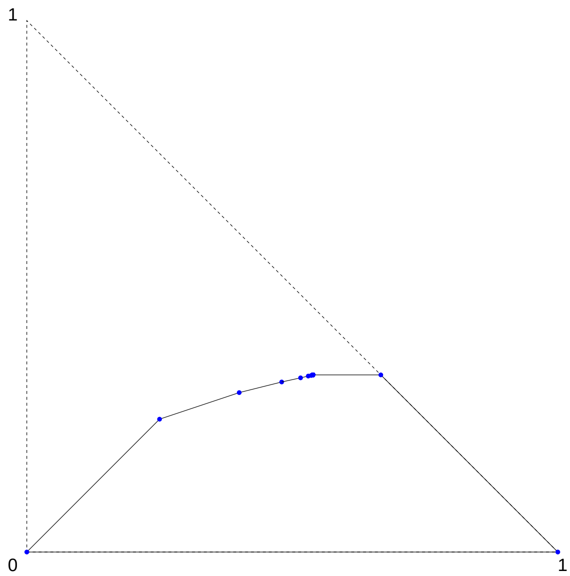


FIGURE 6. An approximation to $\mathrm{DF}(\mathbf{n})$ where $\mathbf{n} = 1\,1\,\overline{1\,0}$

### 3.4. **The digit frequency set varies continuously.**

**Theorem 37.** *The function* $\mathrm{DF}\colon \mathcal{N} \to \mathcal{C}(\Delta)$ *is continuous.*

*Proof.* Let $\mathbf{n} \in \mathcal{N}$. We shall show that DF is continuous at $\mathbf{n}$, and we consider first the case where $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ is not of finite type.

Let $\epsilon > 0$. By Lemma 5, and since $A_{\mathbf{n},r} \to \Phi^{-1}(\mathbf{n})$ as $r \to \infty$, there is some $r$ with $d_H(\mathcal{F}_{\mathbf{n},r}, A_{\mathbf{n},r}) < \epsilon$. Then for any $\mathbf{m} \in \mathcal{N}$ which is close enough to $\mathbf{n}$ that $\mathbf{m}\,[\![r]\!] = \mathbf{n}\,[\![r]\!]$, we have, by Lemma 26,

$$\mathcal{F}_{\mathbf{n},r} \subset L_{\mathbf{n},r} \subset \mathrm{DF}(\mathbf{m}) \subset L_{\mathbf{n},r} \cup A_{\mathbf{n},r} \quad \text{and} \quad \mathcal{F}_{\mathbf{n},r} \subset L_{\mathbf{n},r} \subset \mathrm{DF}(\mathbf{n}) \subset L_{\mathbf{n},r} \cup A_{\mathbf{n},r}.$$

Therefore the symmetric difference

$$\mathrm{DF}(\mathbf{m}) \bigtriangleup \mathrm{DF}(\mathbf{n}) \subset A_{\mathbf{n},r} \subset B_\epsilon(\mathcal{F}_{\mathbf{n},r}) \subset B_\epsilon(\mathrm{DF}(\mathbf{m}) \cap \mathrm{DF}(\mathbf{n})),$$

so that $d_H(\mathrm{DF}(\mathbf{m}), \mathrm{DF}(\mathbf{n})) < \epsilon$ as required.

Next consider the case where $\mathbf{n} = n_0 \,\ldots\, n_{R-2}\, n_{R-1}\, \infty \in \mathcal{N} \setminus \mathbb{N}^{\mathbb{N}}$ (with $R \geq 1$) is of finite type. Now any $\mathbf{m} \in \mathcal{N}$ with $\mathbf{m} \neq \mathbf{n}$ and $d(\mathbf{m}, \mathbf{n}) < 2^{-\left(R + \sum_{s \leq R-1} n_s\right)}$ is of the form $\mathbf{m} = n_0 \,\ldots\, n_{R-2}\, n_{R-1}\, m_R \,\ldots$ for some $m_R \in \mathbb{N}$. It therefore suffices to show that for all $\epsilon > 0$ there is some $M$ such that every $\mathbf{m} \in \mathcal{N}$ of the form $\mathbf{m} = n_0 \,\ldots\, n_{R-2}\, n_{R-1}\, m_R \,\ldots$ with $m_R \geq M$ has $d_H(\mathrm{DF}(\mathbf{m}), \mathrm{DF}(\mathbf{n})) < \epsilon$.

By (11) we have $\mathrm{DF}(\mathbf{n}) = \mathrm{DF}(n_0\, \dots\, n_{R-2}\,(n_{R-1}+1)\,\overline{0})$. Applying Lemma 24 $R-1$ times to $\mathrm{DF}(\mathbf{n})$ and $\mathrm{DF}(\mathbf{m})$ and using the continuity of the maps $K_n^{-1}$, we can suppose that $R = 1$. We therefore need to show that if $\mathbf{m} = n_0\, m_1\, \dots$ with $m_1$ sufficiently large, then $d_H(\mathrm{DF}(\mathbf{m}), \mathrm{DF}((n_0+1)\,\overline{0})) < \epsilon$.

Now $\mathrm{DF}((n_0+1)\,\overline{0}) = T_{n_0}$, while Lemma 26 gives

$$L_{\mathbf{m},0} = T_{n_0} \subset \mathrm{DF}(\mathbf{m}) \subset T_{n_0} \cup K_{n_0}^{-1}(T_{m_1-1}) = U_{\mathbf{m},1}.$$

Since $T_{m_1-1} \to \mathcal{F}$ as $m_1 \to \infty$ and $K_{n_0}^{-1}(\mathcal{F}) \subset T_{n_0}$, the result follows.

The remaining case $\mathbf{n} = \infty$ is straightforward since $\mathrm{DF}(\infty) = \mathcal{F}$ and $\mathrm{DF}(\mathbf{m}) \subset T_{m_0-1} \to \mathcal{F}$ as $m_0 \to \infty$. □

Although $\mathrm{DF}(\mathbf{n})$ varies continuously with $\mathbf{n}$, the same is not true of the set $\mathrm{EP}(\mathbf{n})$ of extreme points of $\mathrm{DF}(\mathbf{n})$. See Figures 4 and 7, which depict respectively $\mathrm{DF}(\mathbf{n})$ and $\mathrm{DF}(\mathbf{m})$, in the case $k = 3$, for elements $\mathbf{n}$ and $\mathbf{m}$ of $\mathbb{N}^{\mathbb{N}}$ which agree on their first 26 entries and are therefore very close to each other. The two digit frequency sets are also very close to each other, but the sets of extreme points are far apart. In these examples, $n_r = r^3$ for $0 \le r \le 25$ and $n_r = 0$ for $r > 25$; while $m_r = n_r$ for all $r$ except $r = 26$, for which $m_r = 100$.

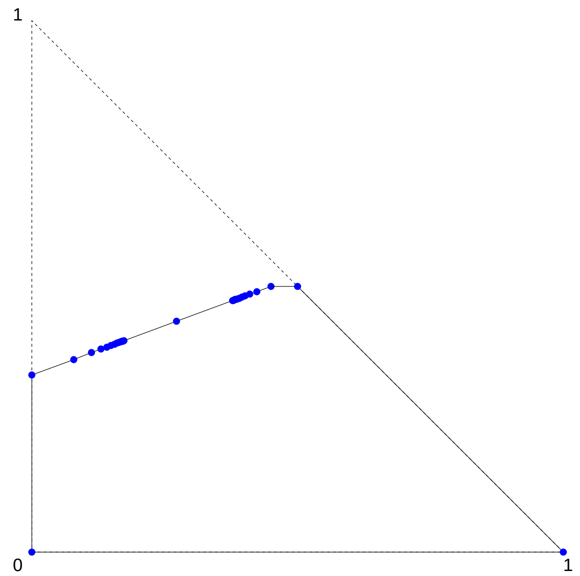


FIGURE 7. A rational digit frequency set close to an exceptional one

The proof of the following theorem shows how such examples can be constructed formally in the case $k = 3$.

**Theorem 38.** *Let $k = 3$. The function $\mathrm{EP}\colon \mathcal{N} \to \mathcal{C}(\Delta)$ is discontinuous at $\mathbf{n} \in \mathcal{N}$ if and only if $\mathbf{n}$ is either exceptional or of finite type.*

*Proof.* Continuity in the case where $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ is regular is proved similarly to continuity of DF (Theorem 37). If $\mathbf{n} \in \mathcal{N}$ satisfies $\mathbf{m}\,[\![r]\!] = \mathbf{n}\,[\![r]\!]$ then $\mathrm{EP}(\mathbf{m}) \triangle \mathrm{EP}(\mathbf{n}) \subset A_{\mathbf{n},r}$, and $A_{\mathbf{n},r}$ contains points of both $\mathrm{EP}(\mathbf{m})$ and $\mathrm{EP}(\mathbf{n})$. Since $A_{\mathbf{n},r}$ converges to the point $\Phi^{-1}(\mathbf{n})$ as $r \to \infty$, the result follows.

Next, consider the case where $\mathbf{n} = n_0\, \dots\, n_r\, \infty$ is of finite type, and set $\mathbf{n}^{(i)} = n_0\, \dots\, n_r\, i\, i\, \overline{0}$ for each $i \ge 1$. Then $\mathbf{n}^{(i)} \to \mathbf{n}$ as $i \to \infty$, and it suffices to show that $\mathrm{EP}(\mathbf{n}^{(i)})$ does not converge to $\mathrm{EP}(\mathbf{n})$.

Recall that $\mathrm{EP}(\mathbf{n}) = \mathrm{EP}(n_0 \dots n_{r-1}\,(n_r+1)\,\overline{0})$ by (11). One of the elements of $\mathrm{EP}(\mathbf{n}^{(i)})$ is $\Phi^{-1}(\mathbf{n}^{(i)}) = K_{n_0}^{-1} \circ \cdots \circ K_{n_r}^{-1} \circ K_i^{-1} \circ K_i^{-1}(0,0,1)$, and these points converge as $i \to \infty$ to the point $L := K_{n_0}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1} \left(\frac{n_r+1}{n_r+3}, \frac{1}{n_r+3}, \frac{1}{n_r+3}\right)$. Since $\mathrm{EP}(\mathbf{n})$ has finitely many elements, it only remains to show that none of them is equal to this limit $L$. It is clearly impossible for an extreme point $K_{n_0}^{-1} \circ \cdots \circ K_{n_s}^{-1}(0,1,0)$ (where $s \le r-1$) to coincide with $L$, since $(0,1,0)$ is not in the $K$-orbit of any interior point of $\Delta$. The only remaining non-trivial element of $\mathrm{EP}(\mathbf{n})$ is

$$K_{n_0}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1} \circ K_{n_r+1}^{-1}(0,0,1) = K_{n_0}^{-1} \circ \cdots \circ K_{n_{r-1}}^{-1} \left(\frac{n_r+1}{n_r+2}, 0, \frac{1}{n_r+2}\right),$$

which is also distinct from $L$.

The case $\mathbf{n} = \infty$ can be treated similarly by considering $\mathbf{n}^{(i)} = i\,i\,\overline{0}$ and using $\mathrm{DF}(\infty) = \mathcal{F}$.

Finally, then, consider the case where $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$ is exceptional, and let $\ell = \Phi^{-1}(\mathbf{n})$ be the exceptional interval in the boundary of $\mathrm{DF}(\mathbf{n})$. Observe first that the points of $\mathrm{EP}(\mathbf{n})$ can only accumulate on the endpoints of $\ell$. For the rational elements of $\mathrm{EP}(\mathbf{n})$ are vertices of the 2-simplices $A_{\mathbf{n},r}$, which converge to $\ell$, so all accumulation points must be in $\ell$; and an accumulation in the interior of $\ell$ would contradict the convexity of $\mathrm{DF}(\mathbf{n})$.

Let $\boldsymbol{\alpha}$ be the midpoint of $\ell$. The 2-simplices $A_{\mathbf{n},r}$ contain $\ell$ in their interior (since they have rational vertices, exactly two of which lie in $\mathrm{DF}(\mathbf{n})$), and every point of $A_{\mathbf{n},r}$ has itinerary starting $n_0 \dots n_{r-3}$ (Section 2.2). For each $r \ge 0$, pick a rational point $\boldsymbol{\alpha}^{(r)}$ of $A_{\mathbf{n},r}$ with $d(\boldsymbol{\alpha}, \boldsymbol{\alpha}^{(r)}) < 2^{-r}$, and let $\mathbf{n}^{(r)} = \Phi(\boldsymbol{\alpha}^{(r)})$. Then $\mathbf{n}^{(r)} \to \mathbf{n}$, $\boldsymbol{\alpha}^{(r)} \in \mathrm{EP}(\mathbf{n}^{(r)})$, and $\boldsymbol{\alpha}^{(r)} \to \boldsymbol{\alpha}$. Since $\boldsymbol{\alpha}$ is bounded away from $\mathrm{EP}(\mathbf{n})$, this establishes the discontinuity of EP at $\mathbf{n}$, as required. □

*Remarks* 39.

a) In the final paragraph of the proof, the rational points $\boldsymbol{\alpha}^{(r)}$ can be chosen all to be contained in, or all to be disjoint from, $\mathrm{DF}(\mathbf{n})$. Therefore EP is discontinuous from both sides at exceptional itineraries.

b) The proofs of continuity of EP in the regular non-finite case, and of discontinuity in the finite case, work for all $k \ge 3$ (with minor modifications in the finite case). The proof of discontinuity in the exceptional case does not generalise so easily, principally because of our ignorance of the set of accumulation points of $\mathrm{E}(\mathbf{n})$ in higher dimensions (Remark 34c)).

## 4. Application to $\beta$-expansions

### 4.1. **Reinterpretation of results on symbolic $\beta$-shifts.**

The results of Section 3 will now be applied to digit frequencies of $\beta$-expansions. We start by recalling some notation and a key fact from Section 1.

Fix throughout an integer $k \ge 2$ and work with $\beta \in (k-1, k]$, so that digit sequences of $\beta$-expansions lie in $\Sigma = \{0, \dots, k-1\}^{\mathbb{N}}$, which we order lexicographically. For each $\beta$, write $\mathrm{DF}(\beta)$ for the set of all digit frequencies of (greedy) $\beta$-expansions of numbers $x \in [0,1]$, a subset of the standard $(k-1)$-simplex $\Delta$. Note that $\mathrm{DF}(k) = \Delta$ and $\mathrm{DF}(k-1) = \mathcal{F}$, the face $\alpha_{k-1} = 0$ of $\Delta$.

The set $Z_\beta$ of all digit sequences $d_\beta(x)$ of $\beta$-expansions of $x \in [0,1]$ is given by

$$Z_\beta = \{v \in \Sigma \,:\, \sigma^r(v) < w_\beta \text{ for all } r \in \mathbb{N}\} \cup \{d_\beta(1)\},$$

where $w_\beta = \lim_{x \nearrow 1} d_\beta(x)$.

We need the following elementary facts about the sequence $w_\beta$:

**Lemma 40.**

a) *The function $\beta \mapsto w_\beta$ from $(k-1, k]$ to $\Sigma$ is strictly increasing.*

*b)* $\{w_\beta : \beta \in (k-1,k]\}$ *is equal to the set of elements* $w$ *of* $\mathcal{M}^*$ *which are not of the form* $w = v\,\overline{0}$ *for any word* $v$*. In particular, for every* $\mathbf{n} \in \mathcal{N} \setminus \{\infty\}$ *there is a unique* $\beta \in (k-1,k]$ *with* $w_\beta = S(\mathbf{n})$*.*

*Proof.* Part a) can be proved in exactly the same way as Lemma 3 of [14]. The first statement of part b) is a translation of the well known result (see for example corollary 7.2.10 of [13]) that an element $w$ of $\mathcal{M}^*$ is equal to $d_\beta(1)$ for some $\beta \in (k-1,k)$ if and only if it is not periodic and not equal to $(k-1)\,\overline{0}$. Finally, if $\mathbf{n} \in \mathcal{N} \setminus \{\infty\}$ then $S(\mathbf{n})$ is an element of $\mathcal{M}^*$ which is in the image of $\Lambda_{n_0}$ and hence is not of the form $v\,\overline{0}$: therefore $S(\mathbf{n}) = w_\beta$ for some $\beta$, which is unique by a). □

**Definition 41.** Define $\beta \colon \mathcal{N} \setminus \{\infty\} \to (k-1,k]$ by

$$\beta(\mathbf{n}) = \text{the unique } \beta \text{ with } w_\beta = S(\mathbf{n}).$$

This is a strictly increasing function by Lemmas 11 and 40a).

The following definition and lemma make the connection between $\mathrm{DF}(\beta)$ for $\beta \in (k-1,k]$, and $\mathrm{DF}(\mathbf{n})$ for $\mathbf{n} \in \mathcal{N}$. The condition $w_0 = k-1$ in the definition is to ensure that $S(\infty) = (k-1)\,\overline{0} \le w$, so that the maximum is defined.

**Definition 42.** Let $\mathbf{n} \colon \{w \in \Sigma : w_0 = k-1\} \to \mathcal{N}$ be the function defined by

$$\mathbf{n}(w) = \max\{\mathbf{m} \in \mathcal{N} : S(\mathbf{m}) \le w\}.$$

**Lemma 43.** *Let* $\beta \in (k-1,k]$*. Then* $\mathrm{DF}(\beta) = \mathrm{DF}(\mathbf{n}(w_\beta))$*.*

*Proof.* $\mathrm{DF}(\beta)$ is the set of digit frequencies of elements of

$$Z_\beta = \{v \in \Sigma : \sigma^r(v) < w_\beta \text{ for all } r \in \mathbb{N}\} \cup \{d_\beta(1)\},$$

while, by Lemma 19, $\mathrm{DF}(\mathbf{n}(w_\beta))$ is the set of digit frequencies of elements of

$$X(w_\beta) = \{v \in \Sigma : \sigma^r(v) \le w_\beta \text{ for all } r \in \mathbb{N}\}.$$

We distinguish two cases.

a) Suppose that $f_\beta^r(1) \neq 0$ for all $r \in \mathbb{N}$, so that $w_\beta = d_\beta(1)$. Then $Z_\beta \subset X(w_\beta)$, since $w_\beta \in X(w_\beta)$. On the other hand, any element $v$ of $X(w_\beta) \setminus Z_\beta$ has $\sigma^r(v) = w_\beta = d_\beta(1)$ for some $r \ge 0$, so that the digit frequency of $v$, if it exists, is equal to that of $d_\beta(1)$. The two digit frequency sets are therefore equal.

b) Suppose that $f_\beta^r(1) = 0$ for some $r \in \mathbb{N}$, so that $d_\beta(1) = d_1 \dots d_r\,\overline{0}$ and $w_\beta = \overline{W}$ is periodic. The only element of $Z_\beta \setminus X(w_\beta)$ is $d_\beta(1)$, which has digit frequency $(1,0,\dots,0)$, the same as the digit frequency of $\overline{0} \in X(w_\beta)$. On the other hand, any element $v$ of $X(w_\beta) \setminus Z_\beta$ satisfies $\sigma^r(v) = w_\beta$ for some $r \in \mathbb{N}$, and hence has the same (rational) digit frequency as $w_\beta$. However the sequence $W\,0\,W\,W\,0\,W\,W\,W\,0\,\dots$ lies in $Z_\beta$ and has the same digit frequency as $w_\beta$. The two digit frequency sets are therefore equal.

□

Using this lemma we can interpret the results of Section 3 in terms of $\beta$-expansions. Before doing so, we define intervals $I_{n_0\dots n_R} \subset (k-1,k]$ associated to each rational-finite pair.

**Definition 44** ($I_{n_0\dots n_R}$)**.** Given $R \ge 0$ and $n_0,\dots,n_R \in \mathbb{N}$ write

$$I_{n_0\dots n_R} = \left[\beta(n_0\,\dots\,n_{R-1}\,(n_R+1)\,\overline{0}),\ \beta(n_0\,\dots\,n_{R-1}\,n_R\,\infty)\right] \subset (k-1,k].$$

**Theorem 45.** *Let* $k \ge 3$*. Then*

*a)* $\mathrm{DF}(\beta)$ *is a compact convex set of dimension* $k-1$ *for all* $\beta \in (k-1,k]$*.*

*b)* $\mathrm{DF}(\beta)$ *has countably many extreme points, of which all but at most* $k-1$ *are rational. There exist* $\beta$ *for which the set of extreme points accumulates on* $k-1$ *non-rational points.*
*c) The extension* $\mathrm{DF}\colon [k-1,k] \to \mathcal{C}(\Delta)$ *is continuous and increasing.*
*d) The* $I_{n_0\dots n_R}$ *are mutually disjoint non-trivial closed subintervals of* $(k-1,k]$ *whose union has full Lebesgue measure, on each of which the digit frequency set is a constant polytope with rational vertices.*
*e) Let* $\mathcal{X} = (k-1,k] \setminus \bigcup_{n\geq 1}\bigcup_{W\in\mathbb{N}^n} I_W$*. The function* $\beta \mapsto w_\beta$ *restricts to a bijection*

$$\mathcal{X} \to \{S(\mathbf{n}) \,:\, \mathbf{n} \text{ is not of rational or finite type}\}.$$

*In particular,* $\mathrm{DF}$ *is injective on* $\mathcal{X}$*, and* $\mathrm{DF}(\mathcal{X})$ *does not contain any polytopes.*
*f) The set* $\mathrm{DF}([k-1,k]) \subset \mathcal{C}(\Delta)$ *is homeomorphic to a compact interval.*

*Proof.* a) is a restatement of Corollary 17 and Theorem 27 (the digit frequency set having dimension $k-1$ since it strictly contains $\mathcal{F}$), and b) is immediate from Theorem 33 and Remark 34b), in each case using Lemma 43.

For c), consider first $\mathrm{DF}\colon (k-1,k) \to \mathcal{C}(\Delta)$. The functions $\beta \mapsto w_\beta$, $w \mapsto \mathbf{n}(w)$, and $\mathbf{n} \mapsto \mathrm{DF}(\mathbf{n})$ are all increasing, the first by Lemma 40 and the other two by definition. Therefore $\beta \mapsto \mathrm{DF}(\beta) = \mathrm{DF}(\mathbf{n}(w_\beta))$ is also increasing. To show that it is continuous, fix $\beta \in (k-1,k)$ and $\epsilon > 0$, and let $\mathbf{n} = \mathbf{n}(w_\beta)$. Since $\mathbf{n} \mapsto \mathrm{DF}(\mathbf{n})$ is continuous by Theorem 37, we can find $\mathbf{m}, \mathbf{p} \in \mathcal{N}$ with $\mathbf{m} < \mathbf{n} < \mathbf{p}$ and with $d_H(\mathrm{DF}(\mathbf{m}), \mathrm{DF}(\mathbf{p})) < \epsilon$. (If $\mathbf{n}$ is of rational type then we take $\mathbf{p}$ to be the corresponding element of finite type, with $\mathrm{DF}(\mathbf{p}) = \mathrm{DF}(\mathbf{n})$; and if not, there are $\mathbf{p} > \mathbf{n}$ arbitrarily close to $\mathbf{n}$. Similarly if $\mathbf{n}$ is of finite type then we take $\mathbf{m}$ to be the corresponding element of rational type; and if not, there are $\mathbf{m} < \mathbf{n}$ arbitrarily close to $\mathbf{n}$.) Then $d_H(\mathrm{DF}(\beta), \mathrm{DF}(\gamma)) < \epsilon$ for all $\gamma \in (\beta(\mathbf{m}), \beta(\mathbf{p}))$.

Since $\mathrm{DF}(k-1) = \mathcal{F}$ and $\mathrm{DF}(k) = \Delta$, the extension to $[k-1,k]$ is clearly increasing. That $\mathrm{DF}(\beta) \to \mathcal{F}$ as $\beta \searrow k-1$ is a consequence of the fact that $\mathrm{DF}(\beta) \subset T_{n_0-1}$ if $\mathbf{n}(\beta)$ begins with $n_0$; and that $\mathrm{DF}(\beta) \to \Delta$ as $\beta \nearrow k$ follows from the observation, using Theorem 33, that if $\mathbf{n}(\beta)$ begins with $0^R$, where $R \geq k-2$, then every non-trivial extreme point of $\mathrm{DF}(\beta)$ lies in $K_0^{-(R+2-k)}(\Delta)$, which converges Hausdorff to $\{\mathbf{e}_{k-1}\}$ as $R \to \infty$.

For d), the intervals $I_{n_0\dots n_R}$ are clearly closed and non-trivial since $\mathbf{n} \mapsto \beta(\mathbf{n})$ is strictly increasing. They are mutually disjoint because $n_0 \,\dots\, n_{R-1}\,(n_R+1)\,\overline{0}$ and $n_0 \,\dots\, n_{R-1}\, n_R\, \infty$ are consecutive elements of $\mathcal{N}$. By Theorem 33 and (11), $\mathrm{DF}(\beta) = \mathrm{DF}(n_0 \,\dots\, n_{R-1}\,(n_R+1)\,\overline{0})$ is a constant polytope on each interval. That the union of the intervals has Lebesgue measure 1 is a consequence of Theorem 49 below.

Now suppose that $\beta$ is in the complement $\mathcal{X}$ of the union of these intervals. Then, by Lemma 13, for every $n_0 \dots n_R$, either we have $S(n_0 \,\dots\, n_{R-1}\,(n_R+1)\,\overline{0}) > w_\beta$, or there is some $\mathbf{m} \in \mathcal{N}$ with $S(n_0 \,\dots\, n_{R-1}\, n_R\, \infty) < S(\mathbf{m}) < w_\beta$. Therefore $\mathbf{n}(w_\beta)$ is not of rational or finite type and, using Lemma 13 again, $w_\beta = S(\mathbf{n}(w_\beta))$. Therefore the image of $\mathcal{X}$ under $\beta \mapsto w_\beta$ is contained in the set of infimax sequences which are not of rational or finite type. On the other hand, every such sequence $S(\mathbf{n})$ is equal to $w_{\beta(\mathbf{n})}$ where $\beta(\mathbf{n}) \in \mathcal{X}$. Since $\beta \mapsto w_\beta$ is strictly increasing, it follows that it is a bijection from $\mathcal{X}$ to the set of infimax sequences which are not of rational or finite type. Moreover, by Theorem 33, $\mathrm{DF}(\beta)$ is not a polytope for $\beta \in \mathcal{X}$.

If $\beta, \gamma \in \mathcal{X}$ with $\beta < \gamma$, then there is some $\mathbf{n}$ of rational type with $w_\beta < S(\mathbf{n}) < w_\gamma$, and hence there is some $\beta'$ between $\beta$ and $\gamma$ with $\mathrm{DF}(\beta')$ a polytope. This establishes the injectivity of DF on $\mathcal{X}$.

By parts d) and e), collapsing each interval $I_{n_0\dots n_R}$ to a point gives a compact interval on which DF descends to a continuous injection, so that the image of DF is a compact interval as required. $\square$

*Remark* 46. One way to see the effect of exceptional elements on digit frequency sets is to define a *forcing relation* $\leq$ on $\Delta$ by

$$\boldsymbol{\alpha} \leq \boldsymbol{\alpha}' \iff \forall \beta \in (k-1, k],\ \boldsymbol{\alpha}' \in \mathrm{DF}(\beta) \implies \boldsymbol{\alpha} \in \mathrm{DF}(\beta).$$

By Lemma 43 and (10), we have

$$\boldsymbol{\alpha} \leq \boldsymbol{\alpha}' \iff \Phi(\boldsymbol{\alpha}) \leq \Phi(\boldsymbol{\alpha}')$$

(if $\Phi(\boldsymbol{\alpha}) > \Phi(\boldsymbol{\alpha}')$ then pick $\mathbf{n} \in \mathcal{N}$ with $\Phi(\boldsymbol{\alpha}') < \mathbf{n} < \Phi(\boldsymbol{\alpha})$ and let $\beta = \beta(\mathbf{n})$: then $\boldsymbol{\alpha}' \in \mathrm{DF}(\beta)$ but $\boldsymbol{\alpha} \notin \mathrm{DF}(\beta)$).

The relation $\leq$ is therefore reflexive, transitive, and total, but is not antisymmetric when $k \geq 3$. In order to make it into a total order, it is necessary to identify each exceptional simplex in $\Delta$ to a point.

**Example 47.** Some parts of Theorem 45 are not true in the case $k = 2$, when $\mathrm{DF}(\beta)$ is a subset of the interval $\Delta = \{(\alpha_0, \alpha_1) \in \mathbb{R}^2_{\geq 0} : \alpha_0 + \alpha_1 = 1\}$, which we identify with $[0, 1]$ using the homeomorphism $(\alpha_0, \alpha_1) \mapsto \alpha_1$. Since $\mathrm{DF}(\beta)$ is compact and convex, and $0 = \delta(d_\beta(0)) \in \mathrm{DF}(\beta)$ for all $\beta \in (1, 2]$, the set $\mathrm{DF}(\beta) = [0, \mathrm{rhe}(\beta)]$ is determined by its right hand endpoint $\mathrm{rhe}(\beta)$, which is the digit frequency of the Sturmian sequence $S(\mathbf{n}(w_\beta))$.

Figure 8 is a graph of $\mathrm{rhe}(\beta)$ against $\beta$, showing how $\mathrm{rhe}(\beta)$ locks on each rational value (see also [12], where the inverse of this function was studied). For instance, the itinerary of the point $(1/2, 1/2) \in \Delta$ is $1\,\overline{0}$, so that $\mathrm{rhe}(\beta) = 1/2$ if and only if $\beta \in I_0 = [\beta(1\,\overline{0}), \beta(0\,\infty)]$; that is, if and only if $w_\beta \in [S(1\,\overline{0}), S(0\,\infty)] = [\overline{10}, 1\overline{10}]$.

Now $w_\beta = \overline{10}$ when $\beta^2 - \beta - 1 = 0$, and $w_\beta = 1\overline{10}$ when $\beta^3 - \beta^2 - 2\beta + 1 = 0$. This gives the endpoints (approximately 1.618 and 1.802) of the interval $\{\beta \in (1, 2] : \mathrm{rhe}(\beta) = 1/2\}$.

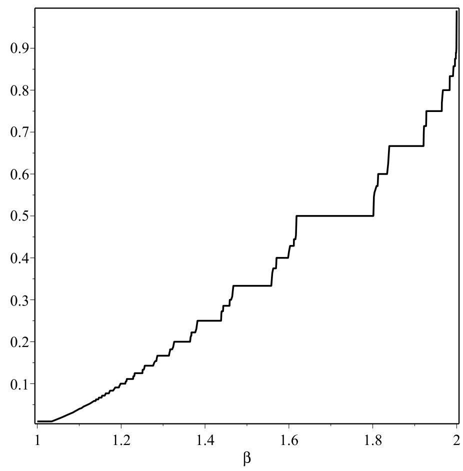


FIGURE 8. The right hand endpoint of $\mathrm{DF}(\beta)$ when $k = 2$

**Example 48.** Here we do analogous calculations to those of Example 47 in the cases $k = 3$ and $k = 4$ (compare with Example 36 and Figure 1). When $k = 3$ we have $\mathrm{DF}(\beta) = \mathrm{DF}(2\,1\,0\,1\,\overline{0})$ if and only if $\beta \in I_{2\,1\,0\,0}$, i.e. if and only if

$$w_\beta \in [S(2\,1\,0\,1\,\overline{0}),\ S(2\,1\,0\,0\,\infty)] = [\overline{200120011},\ 200\overline{120012000}].$$

Now $w_\beta = \overline{200120011}$ when $\beta \simeq 2.190055$ (a root of $\beta^9 - 2\beta^8 - \beta^5 - 2\beta^4 - \beta - 2$), and $w_\beta = 200\overline{120012000}$ when $\beta \simeq 2.19019$ (a root of $\beta^{12} - 2\beta^{11} - \beta^8 - 2\beta^7 - 2\beta^4 + 1$). Thus $\mathrm{DF}(\beta)$ locks on the pentagon of Example 36a) and Figure 1 for $\beta$ between these two values.

On the other hand, when $k=4$ we have $\mathrm{DF}(\beta)=\mathrm{DF}(2\,1\,0\,1\,\overline{0})$ if and only if

$$w_\beta \in [S(2\,1\,0\,1\,\overline{0}),\ S(2\,1\,0\,0\,\infty)] = [\overline{30013000},\ 300\overline{13001}],$$

so that the polyhedron with 7 vertices of Example 36b) is equal to $\mathrm{DF}(\beta)$ for $\beta$ between roots of $\beta^8-3\beta^7-\beta^4-3\beta^3-1$ and $\beta^8-3\beta^7-\beta^4-4\beta^3+3\beta^2-1$ (approximately 3.0688 and 3.0690).

4.2. **Typical phenomena.** In this section we shall show that, from the point of view of the parameter $\beta$, the typical digit frequency set is of rational type (that is, a polytope with rational vertices). By constrast, we then show that from the point of view of the digit frequency sets themselves, the generic example is non-rational and regular (that is, having a single accumulation of rational vertices); and moreover, the non-rational extreme point is generically *totally irrational* (its components are independent over the rationals).

**Theorem 49.** *Let $k\ge 2$. Then $\mathrm{DF}(\beta)$ is a polytope with rational vertices for Lebesgue a.e. $\beta\in(k-1,k]$.*

*Proof.* Suppose first that $k\ge 3$. For each $\beta\in(k-1,k]$, let $\mathbf{p}(\beta)\in\Delta$ be the *normal* digit frequency for $\beta$-expansions, which is realised by $d_\beta(x)$ for Lebesgue a.e. $x\in[0,1]$. It is given by

$$p_i(\beta)=\int_{i/\beta}^{(i+1)/\beta} h_\beta \qquad \text{for } 0\le i<k-1, \quad \text{and} \qquad p_{k-1}(\beta)=\int_{(k-1)/\beta}^{1} h_\beta,$$

where $h_\beta\colon[0,1]\to\mathbb{R}^+$ is the density of Parry's measure of maximal entropy [14]. Now $h_\beta$ is a decreasing function, from which it follows that

$$\frac{p_{k-1}(\beta)}{p_0(\beta)} \le \frac{1-(k-1)/\beta}{1/\beta} = \beta-(k-1). \tag{12}$$

By a theorem of Schmeling [17], the sequence $w_\beta$ has digit frequency $\mathbf{p}(\beta)$ for Lebesgue a.e. $\beta$ in $(k-1,k]$. It therefore suffices to prove that $w_\beta$ does *not* have digit frequency $\mathbf{p}(\beta)$ whenever $\mathrm{DF}(\beta)$ is not a polytope with rational vertices.

Suppose therefore that $\mathrm{DF}(\beta)$ is not a polytope with rational vertices, so that, by Theorem 45d) and e), $w_\beta=\mathcal{I}(\boldsymbol{\alpha})$ for some $\boldsymbol{\alpha}$ which is not of rational type. If the digit frequency of $w_\beta$ exists then it is equal to $\boldsymbol{\alpha}$ by Facts 1f), so that it is only necessary to show that $\boldsymbol{\alpha}\ne\mathbf{p}(\beta)$.

Let $\mathbf{n}=\Phi(\boldsymbol{\alpha})$. Then $\boldsymbol{\alpha}\in\Delta_{n_0}$, which means by definition that $\alpha_{k-1}/\alpha_0>1/(n_0+1)$ (we can assume that $\alpha_0>0$, since otherwise it is immediate that $\boldsymbol{\alpha}\ne\mathbf{p}(\beta)$). We shall show that $1/(n_0+1)>\beta-(k-1)$, which will establish the result by comparison with (12). This statement is immediate if $n_0=0$ (we cannot have $\beta=k$, since $\mathrm{DF}(\beta)$ is not a polytope with rational vertices), so we suppose $n_0\ge 1$.

Since $\mathbf{n}<n_0\,\overline{0}$, we have $w_\beta=\mathcal{I}(\boldsymbol{\alpha})<S(n_0\,\overline{0})=\overline{(k-1)\,0^{n_0}}$, and hence $\beta<\beta'$, where $w_{\beta'}=\overline{(k-1)\,0^{n_0}}$. Now $\beta'$ is the unique root in $(k-1,k)$ of the function $f(x)=x^{n_0+1}-(k-1)x^{n_0}-1$. Since this function is increasing in $(k-1,k)$, showing that $f((k-1)+1/(n_0+1))\ge 0$ will establish that $\beta<\beta'\le(k-1)+1/(n_0+1)$.

Now

$$\begin{aligned} f\left(k-1+\frac{1}{n_0+1}\right) &= \left(k-1+\frac{1}{n_0+1}\right)^{n_0+1}-(k-1)\left(k-1+\frac{1}{n_0+1}\right)^{n_0}-1\\ &= \frac{1}{n_0+1}\left(k-1+\frac{1}{n_0+1}\right)^{n_0}-1\\ &> \frac{(k-1)^{n_0}}{n_0+1}-1 \ \ge\ \frac{2^{n_0}}{n_0+1}-1 \ \ge\ 0 \end{aligned}$$

as required.

The case $k = 2$ follows from the observation in [8] (page 398) that the set of $\beta \in (1, 2]$ for which $w_\beta$ is an irrational infimax (i.e. a Sturmian sequence) has zero measure, by theorem C of [17] and standard results on Sturmian sequences. $\square$

Before embarking on the proof that the generic digit frequency set is non-rational and regular, we introduce some notation for the various spaces which will be involved. We fix throughout the integer $k \geq 3$.

An element $\boldsymbol{\alpha}$ of $\Delta'$ is said to be *completely irrational* if there is no non-trivial relationship of the form $\sum_{i=0}^{k-1} m_i\alpha_i = 0$ for integers $m_i$. It is said to be of *infinite type* if every component of $K^r(\boldsymbol{\alpha})$ is strictly positive for all $r \geq 0$.

We define the following subsets of $\Delta'$:

$$\begin{aligned}
\mathrm{Rat}_\Delta &= \text{the set } \Delta' \cap \mathbb{Q}^k \text{ of rational elements,}\\
\mathrm{CI}_\Delta &= \text{the set of completely irrational elements,}\\
\mathrm{In}_\Delta &= \text{the set of infinite type elements,}\\
\mathrm{Reg}_\Delta &= \text{the set of regular elements, and}\\
\mathcal{O}_\Delta &= \text{the set of elements whose itinerary contains infinitely many words } 1^{2k-3}.
\end{aligned}$$

The images of these sets under the itinerary map $\Phi\colon \Delta' \to \mathbb{N}^\mathbb{N}$ are denoted with subscripts $S$ (for "sequence"). Thus, for example, $\mathrm{Rat}_S = \Phi(\mathrm{Rat}_\Delta)$ is the set of elements of $\mathbb{N}^\mathbb{N}$ which end $\overline{0}$; $\mathcal{O}_S = \Phi(\mathcal{O}_\Delta)$ is the set of elements of $\mathbb{N}^\mathbb{N}$ which contain infinitely many distinct words $1^{2k-3}$; and, by Facts 1i), $\mathrm{In}_S = \Phi(\mathrm{In}_\Delta)$ is the set of elements $\mathbf{n}$ of $\mathbb{N}^\mathbb{N}$ which have the property that, for all $r \geq 0$, there is some $s \geq 0$ with $n_{r+s(k-1)} \neq 0$. Facts 1h) states that $\mathcal{O}_S \subset \mathrm{Reg}_S$ is a dense $G_\delta$ subset of $\mathbb{N}^\mathbb{N}$.

Let $\mathcal{D} = \mathrm{DF}(\mathbb{N}^\mathbb{N})$ be the set of all digit frequency sets for $\beta \in (k-1, k]$ with the Hausdorff topology, which is homeomorphic to a half-open interval. The images of the above subsets of $\mathbb{N}^\mathbb{N}$ under DF will be denoted with a subscript $D$ (for "digit"). Thus, for example, $\mathrm{Rat}_D$ is the set of digit frequency sets which are polytopes, and $\mathrm{Reg}_D$ is contained in the set of digit frequency sets which are either polytopes or have a single non-rational extreme point.

The complements of these sets in $\Delta'$, in $\mathbb{N}^\mathbb{N}$, or in $\mathcal{D}$, as appropriate, are denoted with a superscript $c$. Thus, for example, $\mathrm{Rat}_D^c$ is the set of non-polytope digit frequency sets.

Recall that a function $f\colon X \to Y$ is called *quasi-open* if, for every open set $U \subset X$, the image $f(U) \subset Y$ has interior.

**Lemma 50.**

*a)* $\mathrm{CI}_\Delta$ *and* $\mathrm{In}_\Delta$ *are dense* $G_\delta$ *subsets of* $\Delta'$*, with* $\mathrm{CI}_\Delta \subset \mathrm{In}_\Delta$*.*

*b) The itinerary map* $\Phi\colon \Delta' \to \mathbb{N}^\mathbb{N}$ *is quasi-open; its restriction* $\Phi\colon \mathrm{In}_\Delta \to \mathrm{In}_S$ *is continuous; and its restriction* $\Phi\colon \mathrm{In}_\Delta \cap \mathcal{O}_\Delta \to \mathrm{In}_S \cap \mathcal{O}_S$ *is a homeomorphism.*

*c)* $\mathrm{In}_S$ *is a dense* $G_\delta$ *subset of* $\mathbb{N}^\mathbb{N}$*.*

*d) The restriction* $\mathrm{DF}\colon \mathrm{Rat}_S^c \to Rat_D^c$ *is a homeomorphism.*

*Proof.*

a) $\mathrm{CI}_\Delta$ is the countable intersection of the open dense subsets $\sum_{i=0}^{k-1} m_i\alpha_i \neq 0$ of the Baire space $\Delta'$, and is therefore dense $G_\delta$. Similarly for

$$\mathrm{In}_\Delta = \bigcap_{R\geq 0}\ \bigcap_{n_0,\dots,n_R\in\mathbb{N}} \left(\Delta' \setminus K_{n_0}^{-1} \circ \cdots \circ K_{n_R}^{-1}(\partial\Delta)\right),$$

where $\partial\Delta$ is the union of the faces of $\Delta$.

The image of a completely irrational vector under a projective homeomorphism with integer coefficients is again completely irrational, and in particular has no zero components, from which it follows that $\mathrm{CI}_\Delta \subset \mathrm{In}_\Delta$.

b) Let $U \subset \Delta'$ be open, pick $\boldsymbol{\alpha} \in U \cap \mathrm{Rat}_\Delta$, and write $\mathbf{n} = \Phi(\boldsymbol{\alpha})$. Because $\mathrm{Rat}_\Delta \subset \mathrm{Reg}_\Delta$, we have $A_{\mathbf{n},r} \subset U$ for sufficiently large $r$. Therefore $\Phi(U)$ contains the (open) cylinder set determined by the block $n_0 \ldots n_r$ for sufficiently large $r$, establishing that $\Phi$ is quasi-open as required.

To show that the restriction $\Phi\colon \mathrm{In}_\Delta \to \mathrm{In}_S$ is continuous, observe that the hyperplanes $\alpha_0 = n\alpha_{k-1}$ on which $K\colon \Delta' \to \Delta'$ is discontinuous are contained in $\mathrm{In}_\Delta^c$. It follows that if $\boldsymbol{\alpha} \in \mathrm{In}_\Delta$ and $r > 0$, then there is a neighbourhood $U$ of $\boldsymbol{\alpha}$ in $\Delta'$ such that $K^s(U)$ is disjoint from these hyperplanes for $0 \le s \le r$. Then $\Phi(U)$ is contained in the cylinder set determined by the first $r$ symbols of $\Phi(\boldsymbol{\alpha})$, which establishes continuity.

In particular, the restriction $\Phi\colon \mathrm{In}_\Delta \cap \mathcal{O}_\Delta \to \mathrm{In}_S \cap \mathcal{O}_S$ is continuous. It is also bijective because points of $\Phi(\mathcal{O}_\Delta) = \mathcal{O}_S$ have unique preimages under $\Phi$. It therefore only remains to show that it is open. For this it is required to show that if $U \subset \Delta'$ is open and $\mathbf{n} \in \Phi(U \cap \mathrm{In}_\Delta \cap \mathcal{O}_\Delta)$, then there is an open subset $V$ of $\mathbb{N}^\mathbb{N}$ with $\mathbf{n} \in V \cap \mathrm{In}_S \cap \mathcal{O}_S \subset \Phi(U \cap \mathrm{In}_\Delta \cap \mathcal{O}_\Delta)$. As above, the fact that $\mathbf{n}$ is regular means that $A_{\mathbf{n},r} \subset U$ for sufficiently large $r$, so that the cylinder set $V$ determined by $n_0 \ldots n_r$ satisfies $V \subset \Phi(U)$. Therefore

$$V \cap \mathrm{In}_S \cap \mathcal{O}_S \;\subset\; \Phi(U) \cap \Phi(\mathrm{In}_\Delta) \cap \Phi(\mathcal{O}_\Delta) \;=\; \Phi(U \cap \mathrm{In}_\Delta \cap \mathcal{O}_\Delta),$$

with the final equality holding since points of $\Phi(\mathcal{O}_\Delta)$ have a unique $\Phi$-preimage in $\Delta'$. This completes the proof, since clearly $\mathbf{n} \in V \cap \mathrm{In}_S \cap \mathcal{O}_S$.

c)
$$\mathrm{In}_S = \bigcap_{r=0}^{\infty} \{\mathbf{n} \in \mathbb{N}^\mathbb{N} \,:\, n_{r+s(k-1)} \ne 0 \text{ for some } s \ge 0\},$$

a countable intersection of open dense subsets of the Baire space $\mathbb{N}^\mathbb{N}$.

d) $\mathrm{DF}\colon \mathrm{Rat}_S^c \to \mathrm{Rat}_D^c$ is continuous by Theorem 37, injective by Corollary 35, and surjective by definition. It therefore only remains to show that it is open. Since the cylinder sets form a basis for the topology of $\mathbb{N}^\mathbb{N}$, it suffices to show that for each cylinder set $C$, the set $\mathrm{DF}(C \setminus \mathrm{Rat}_S)$ is open in $\mathrm{Rat}_D^c$. Suppose, then, that $C$ is determined by the block $n_0 \ldots n_r$. Write $\ell = n_0 \ldots n_r \infty$ and $r = n_0 \ldots n_r \overline{0}$, so that $C = [\ell, r]_\mathcal{N} \cap \mathbb{N}^\mathbb{N}$, where $[a,b]_\mathcal{N} := \{\mathbf{n} \in \mathcal{N} \,:\, a \le \mathbf{n} \le b\}$. Since DF is continuous and order-preserving on $\mathcal{N}$ and $\mathrm{DF}(\ell), \mathrm{DF}(r) \in \mathrm{Rat}_D$, we have that

$$\mathrm{DF}(C \setminus \mathrm{Rat}_S) = [\mathrm{DF}(\ell), \mathrm{DF}(r)]_D \setminus \mathrm{Rat}_D = (\mathrm{DF}(\ell), \mathrm{DF}(r))_D \setminus \mathrm{Rat}_D$$

is open in $\mathrm{Rat}_D^c$ as required.

□

We can now prove that a generic digit frequency set has a single limiting extreme point which is completely irrational.

**Theorem 51.** *The set* $\mathrm{Reg}_D \cap \mathrm{CI}_D$ *contains a dense* $G_\delta$ *subset of* $\mathcal{D}$.

*Proof.* In this proof references a), b), c), and d) are to the parts of Lemma 50, while numerical references 1) and 2) are to the following straightforward facts about subspaces $A \subset B \subset X$ of a metric space $X$:

1) If $A$ is a dense (respectively $G_\delta$) subset of $X$, then it is a dense (respectively $G_\delta$) subset of $B$.
2) If $B$ is a dense (respectively $G_\delta$) subset of $X$, and $A$ is a dense (respectively $G_\delta$) subset of $B$, then $A$ is a dense (respectively $G_\delta$) subset of $X$.

By c) and Facts 1h), $\mathcal{O}_S \cap \mathrm{In}_S$ is a dense $G_\delta$ subset of $\mathbb{N}^\mathbb{N}$. Because $\mathcal{O}_S \subset \mathrm{Reg}_S$, its preimage $\Phi^{-1}(\mathcal{O}_S \cap \mathrm{In}_S)$ is equal to $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta$.

Now $\mathcal{O}_S \cap \mathrm{In}_S$ is a $G_\delta$ subset of $\mathrm{In}_S$ by 1), so by the continuity of $\Phi\colon \mathrm{In}_\Delta \to \mathrm{In}_S$, its preimage $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta$ is a $G_\delta$ subset of $\mathrm{In}_\Delta$, and hence, by a) and 2), of $\Delta'$. On the other hand, it follows from the quasi-openness of $\Phi\colon \Delta' \to \mathbb{N}^\mathbb{N}$ and the denseness of $\mathcal{O}_S \cap \mathrm{In}_S$ in $\mathbb{N}^\mathbb{N}$ that $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta$ is dense in $\Delta'$. That is, $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta$ is a dense $G_\delta$ subset of $\Delta'$. By a), $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta \cap \mathrm{CI}_\Delta = \mathcal{O}_\Delta \cap \mathrm{CI}_\Delta$ is also a dense $G_\delta$ subset of $\Delta'$.

In particular, by 1), $\mathcal{O}_\Delta \cap \mathrm{CI}_\Delta$ is a dense $G_\delta$ subset of $\mathcal{O}_\Delta \cap \mathrm{In}_\Delta$. Therefore, by b), $\Phi(\mathcal{O}_\Delta \cap \mathrm{CI}_\Delta)$ is a dense $G_\delta$ subset of $\mathcal{O}_S \cap \mathrm{In}_S$, and so also of $\mathbb{N}^\mathbb{N}$ by 2); but $\Phi(\mathcal{O}_\Delta \cap \mathrm{CI}_\Delta) = \Phi(\mathcal{O}_\Delta) \cap \Phi(\mathrm{CI}_\Delta) = \mathcal{O}_S \cap \mathrm{CI}_S$ since $\mathcal{O}_S \subset \mathrm{Reg}_S$.

Now $\mathcal{O}_S \cap \mathrm{CI}_S$ is a dense $G_\delta$ subset of $\mathrm{Rat}_S^c$ by 1), and hence $\mathrm{DF}(\mathcal{O}_S \cap \mathrm{CI}_S)$ is a dense $G_\delta$ subset of $\mathrm{Rat}_D^c$ by d). Since $\mathrm{Rat}_D^c$ is a dense $G_\delta$ subset of $\mathcal{D}$ (it has countable complement), it follows by 2) that the subset $\mathrm{DF}(\mathcal{O}_S \cap \mathrm{CI}_S)$ of $\mathrm{Reg}_D \cap \mathrm{CI}_D$ is dense $G_\delta$ in $\mathcal{D}$ as required. $\square$

The following lemma provides explicit elements of $\mathrm{Reg}_S \cap \mathrm{CI}_S$, and hence of $\mathrm{Reg}_D \cap \mathrm{CI}_D$.

**Lemma 52.** *Suppose that $k \geq 3$. For every $n > 0$, the element $\overline{n}$ of $\mathbb{N}^\mathbb{N}$ lies in $\mathrm{Reg}_S \cap \mathrm{CI}_S$.*

*Proof.* The itinerary $\overline{n}$ lies in $\mathrm{Reg}_S$ since it is bounded and contains no zeroes. It is therefore only necessary to prove that if $\Phi(\boldsymbol{\alpha}) = \overline{n}$, then $\boldsymbol{\alpha}$ is completely irrational.

Let $A$ be the $k$ by $k$ matrix which is the Abelianization of the substitution $\Lambda_n$: that is, $A_{0,k-2} = n+1$, $A_{0,k-1} = n$, $A_{i,i-1} = 1$ for $1 \leq i \leq k-1$, $A_{k-1,k-1} = 1$, and all other entries are zero. Then $\boldsymbol{\alpha} = \mathbf{v}/||\mathbf{v}||_1$, where $\mathbf{v}$ is the Perron-Frobenius eigenvector for $A$ normalized so that $v_{k-1} = 1$.

The eigenvector equation $A\mathbf{v} = \lambda\mathbf{v}$ gives $v_i = \lambda^{k-2-i}(\lambda - 1)$ for $0 \leq i \leq k-2$. Therefore if $m_i \in \mathbb{Z}$ for $0 \leq i \leq k-1$, then $\sum_{i=0}^{k-1} m_i v_i$ is a polynomial in $\lambda$ of degree at most $k-1$ with integer coefficients. Hence if $\boldsymbol{\alpha}$ is not completely irrational, then the degree of $\lambda$ is less than $k$, and therefore the characteristic polynomial $p(x) = x^k - x^{k-1} - (n+1)x + 1$ of $A$ is reducible over $\mathbb{Z}$. Conversely, if $q(\lambda) = 0$ for some non-zero integer polynomial $q(x) = \sum_{i=0}^{k-1} q_i x^i$ of degree less than $k$, then we can construct integers $m_i = \sum_{j=k-1-i}^{k-1} q_j$, not all zero, with $\sum_{i=0}^{k-1} m_i v_i = 0$. Therefore $\boldsymbol{\alpha}$ is completely irrational if and only if $p(x)$ is irreducible over $\mathbb{Z}$.

Let $r(x) = x^k p(1/x) = x^k - (n+1)x^{k-1} - x + 1$ which is irreducible if and only if $p$ is. The Perron criterion for irreducibility applies to $r$ to give the required result when $n + 1 > 3$. A slight variation in the first part of the usual proof of Perron's criterion (for example, in the proof of Theorem 2.2.5(a) of [15], take $g(x) = -(n+1)x^{k-1} + 1$ which has all $k-1$ roots in the unit disk and satisfies $|g| > |r - g|$ on the unit circle, implying by Rouché's theorem that $r$ also has $k-1$ roots in the unit disk) gives the result for $n = 1$ and $n = 2$ also. $\square$

4.3. **Generic smoothness at non-rational extreme points.** Figures 4 and 5 suggest that, in the case $k = 3$, non-rational (i.e. limiting) extreme points $\boldsymbol{\alpha}$ of digit frequency sets $\mathrm{DF}(\mathbf{n})$ are *smooth*: that is, that there is a unique line $L$ through $\boldsymbol{\alpha}$ in the plane of $\mathrm{DF}(\mathbf{n})$ such that $\mathrm{DF}(\mathbf{n}) \setminus L$ is connected. In this section we show that this property holds for all $\mathbf{n}$ in $\mathcal{P}_S$, the set of itineraries which contain infinitely many words 1 1 1 1 1. By arguments analogous to those of Section 4.2, $\mathcal{P}_S$ and its counterparts $\mathcal{P}_\Delta$ and $\mathcal{P}_D$ are generic subsets of $\mathbb{N}^\mathbb{N}$, $\Delta$, and $\mathcal{D}$ respectively. Note that, by Facts 1h), we have $\mathcal{P}_S \subset \mathrm{Reg}_S$.

We restrict to the case $k = 3$ throughout. Let $\boldsymbol{\alpha} \in \Delta$ have itinerary $\Phi(\boldsymbol{\alpha}) = \mathbf{n}$. Recall from Section 2.2 that we write, for each $r \geq 0$,

$$\begin{aligned} \Lambda_{\mathbf{n},r} &= \Lambda_{n_0} \circ \Lambda_{n_1} \circ \cdots \circ \Lambda_{n_r}, \\ \Upsilon_{\mathbf{n},r} &= K_{n_0}^{-1} \circ K_{n_1}^{-1} \circ \cdots \circ K_{n_r}^{-1} \colon \Delta \to \Delta, \qquad \text{and} \\ A_{\mathbf{n},r} &= \Upsilon_{\mathbf{n},r}(\Delta). \end{aligned}$$

Thus $A_{\mathbf{n},r}$ is a triangle, with rational vertices labelled $\boldsymbol{\alpha}_r^{(0)} = \Upsilon_{\mathbf{n},r}(1,0,0)$, $\boldsymbol{\alpha}_r^{(1)} = \Upsilon_{\mathbf{n},r}(0,1,0)$, and $\boldsymbol{\alpha}_r^{(2)} = \Upsilon_{\mathbf{n},r}(0,0,1)$. The vertices $\boldsymbol{\alpha}_r^{(0)}$ and $\boldsymbol{\alpha}_r^{(1)}$ are contained in $\mathrm{DF}(\mathbf{n})$; on the other hand, since $\Phi^{-1}(\mathbf{n}) \subset A_{\mathbf{n},r}$, we have $\boldsymbol{\alpha}_r^{(2)} \notin \mathrm{DF}(\mathbf{n})$.

By (4) and (5), the triangles $A_{\mathbf{n},r}$ evolve according to

$$\begin{aligned} \boldsymbol{\alpha}_{r+1}^{(0)} &= \boldsymbol{\alpha}_r^{(1)}, \\ \boldsymbol{\alpha}_{r+1}^{(1)} &= \frac{(n_{r+1}+1)L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(2)}\boldsymbol{\alpha}_r^{(2)}}{(n_{r+1}+1)L_r^{(0)} + L_r^{(2)}}, \qquad \text{and} \\ \boldsymbol{\alpha}_{r+1}^{(2)} &= \frac{n_{r+1}L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(2)}\boldsymbol{\alpha}_r^{(2)}}{n_{r+1}L_r^{(0)} + L_r^{(2)}}, \end{aligned}$$

where $L_r^{(i)} = |\Lambda_{\mathbf{n},r}(i)|$. Therefore both $\boldsymbol{\alpha}_{r+1}^{(1)}$ and $\boldsymbol{\alpha}_{r+1}^{(2)}$ lie on the edge of $A_{\mathbf{n},r}$ with endpoints $\boldsymbol{\alpha}_r^{(0)}$ and $\boldsymbol{\alpha}_r^{(2)}$, and cut this edge in the ratios $(n_{r+1}+1)L_r^{(0)} : L_r^{(2)}$ and $n_{r+1}L_r^{(0)} : L_r^{(2)}$ respectively. See Figure 9, in which $A_{\mathbf{n},r}$ is shown with dashed edges and $A_{\mathbf{n},r+1}$ with solid edges.

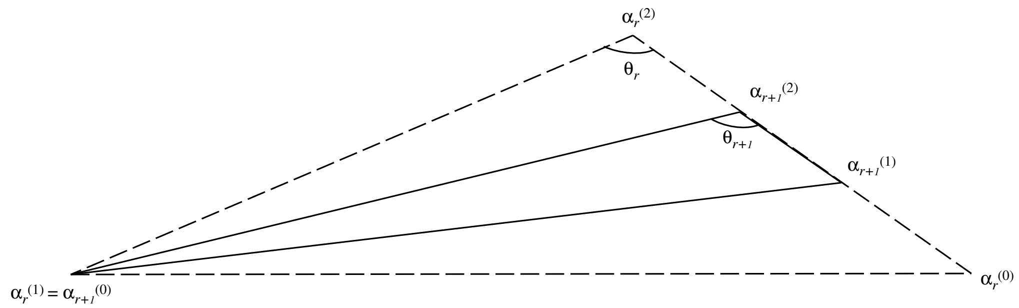


FIGURE 9. Evolution of the triangles $A_{\mathbf{n},r}$

As in the figure, let $\theta_r$ denote the angle of the triangle $A_{\mathbf{n},r}$ at the vertex $\boldsymbol{\alpha}_r^{(2)}$. Then $\theta_{r+1} \geq \theta_r$ for all $r$, with equality if and only if $n_{r+1} = 0$ (i.e. if and only if $\boldsymbol{\alpha}_{r+1}^{(2)} = \boldsymbol{\alpha}_r^{(2)}$). Generic smoothness at non-rational extreme points is a consequence of the following lemma.

**Lemma 53.** *Suppose that* $\mathbf{n} \in \mathcal{P}_S$*, and moreover that* $n_0 = n_1 = 1$*. Then* $\theta_r \to \pi$ *as* $r \to \infty$*.*

*Proof.* Suppose for a contradiction that $\theta_r \to \theta = \pi - 2\epsilon$ as $r \to \infty$, where $\epsilon > 0$. A direct calculation using $n_0 = n_1 = 1$ gives $\boldsymbol{\alpha}_1^{(0)} = (2/3, 0, 1/3)$, $\boldsymbol{\alpha}_1^{(1)} = (1/4, 1/2, 1/4)$, and $\boldsymbol{\alpha}_1^{(2)} = (1/3, 1/3, 1/3)$, so that $\cos\theta_1 < 0$, and hence $\theta > \pi/2$.

Now pick any $r$ with $n_{r+s} = 1$ for $0 \le s \le 4$. Using $\Lambda_1(0) = 1$, $\Lambda_1(1) = 2\,0\,0$ and $\Lambda_1(2) = 2\,0$, we have

$$\begin{aligned} L^{(0)}_{r+s} &= L^{(1)}_{r+s-1}, \\ L^{(1)}_{r+s} &= L^{(2)}_{r+s-1} + 2L^{(0)}_{r+s-1}, \qquad \text{and} \\ L^{(2)}_{r+s} &= L^{(2)}_{r+s-1} + L^{(0)}_{r+s-1} \end{aligned}$$

for $0 \le s \le 4$. Writing $(a, b, c) = (L^{(0)}_{r-1}, L^{(1)}_{r-1}, L^{(2)}_{r-1})$, this gives

$$\begin{aligned} (L^{(0)}_{r+2}, L^{(1)}_{r+2}, L^{(2)}_{r+2}) &= (a + 2b + c, 5a + b + 3c, 3a + b + 2c) \qquad \text{and} \\ (L^{(0)}_{r+3}, L^{(1)}_{r+3}, L^{(2)}_{r+3}) &= (5a + b + 3c, 5a + 5b + 4c, 4a + 3b + 3c). \end{aligned}$$

Therefore both $n_{r+3}L^{(0)}_{r+2}/L^{(2)}_{r+2} = L^{(0)}_{r+2}/L^{(2)}_{r+2}$ and $n_{r+4}L^{(0)}_{r+3}/L^{(2)}_{r+3} = L^{(0)}_{r+3}/L^{(2)}_{r+3}$ are greater than 1/3. That is, $\boldsymbol{\alpha}^{(2)}_{r+3}$ is at least 1/3 of the way along the edge of $A_{\mathbf{n},r+2}$ from $\boldsymbol{\alpha}^{(2)}_{r+2}$ to $\boldsymbol{\alpha}^{(0)}_{r+2}$; and similarly $\boldsymbol{\alpha}^{(2)}_{r+4}$ is at least 1/3 of the way along the edge of $A_{\mathbf{n},r+3}$ from $\boldsymbol{\alpha}^{(2)}_{r+3}$ to $\boldsymbol{\alpha}^{(0)}_{r+3}$.

Let $\varphi_r$ and $\psi_r$ denote the angles of the triangle $A_{\mathbf{n},r}$ at the vertices $\boldsymbol{\alpha}^{(0)}_r$ and $\boldsymbol{\alpha}^{(1)}_r$ respectively. Then $\varphi_r + \psi_r = \pi - \theta_r > 2\epsilon$ for all $r$; moreover (see Figure 9), $\psi_{r+1} > \varphi_r$ for all $r$, so that if $\psi_r < \epsilon$ then $\psi_{r+1} > \epsilon$.

We can therefore pick an index $r$ within each block of symbols 1 1 1 1 1 in $\mathbf{n}$ with the property that both $\psi_r > \epsilon$ and $n_{r+1}L^{(0)}_r/L^{(2)}_r > 1/3$. Figure 10 shows the corresponding triangle $A_{\mathbf{n},r}$ and the relevant edges of $A_{\mathbf{n},r+1}$. Let $u$, $v$, $w$, and $\gamma_r$ denote the lengths and angle indicated in the figure. The condition $n_{r+1}L^{(0)}_r/L^{(2)}_r > 1/3$ gives that $(u + v)/v > 4/3$.

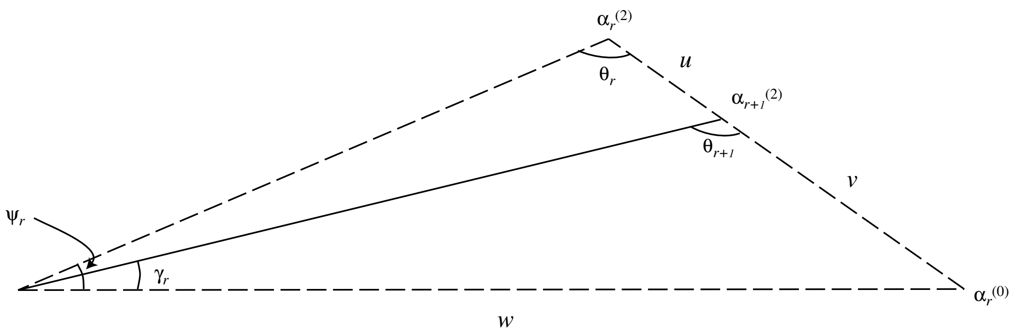


FIGURE 10. Angles in the triangles $A_{\mathbf{n},r}$ and $A_{\mathbf{n},r+1}$

Applying the sine rule to the two triangles with base $w$ gives

$$\frac{\sin\psi_r}{\sin\gamma_r} = \frac{u+v}{v}\,\frac{\sin\theta_r}{\sin\theta_{r+1}} > \frac{4}{3},$$

since $\pi/2 < \theta_r < \theta_{r+1} < \pi$. Therefore $\sin\psi_r - \sin\gamma_r > \frac{1}{4}\sin\psi_r > \frac{1}{4}\sin\epsilon$. Since $0 < \gamma_r < \psi_r < \pi/2$, this gives

$$\theta_{r+1} - \theta_r = \psi_r - \gamma_r > \sin\psi_r - \sin\gamma_r > \frac{1}{4}\,\sin\epsilon.$$

Therefore $\theta_r$ increases by at least $\frac{1}{4}\sin\epsilon$ as the index $r$ passes through each block 1 1 1 1 1 in $\mathbf{n}$; since there are infinitely many such blocks, this gives the required contradiction. $\square$

**Theorem 54.** *Let* $\mathbf{n} \in \mathcal{P}_S$. *Then* $\boldsymbol{\alpha} = \Phi^{-1}(\mathbf{n})$ *is a smooth extreme point of* $\mathrm{DF}(\mathbf{n})$.

*Proof.* We can suppose without loss of generality that $n_0 = n_1 = 1$, so that the hypotheses of Lemma 53 are satisfied. For if not, let $\mathbf{m} = \sigma^r(\mathbf{n})$ for some $r$ with $n_r = n_{r+1} = 1$. Then, by Lemma 24, $\mathrm{DF}(\mathbf{n})$ is the union of a polygon with the image of $\mathrm{DF}(\mathbf{m})$ under a projective homeomorphism, and hence $\boldsymbol{\alpha}$ is a smooth extreme point of $\mathrm{DF}(\mathbf{n})$ if and only if $\Phi^{-1}(\mathbf{m})$ is a smooth extreme point of $\mathrm{DF}(\mathbf{m})$.

Suppose for a contradiction that $\boldsymbol{\alpha}$ is not a smooth extreme point, so that there are distinct lines $L_1$ and $L_2$ through $\boldsymbol{\alpha}$ which do not disconnect $\mathrm{DF}(\mathbf{n})$. Let $\Theta < \pi$ be the angle between $L_1$ and $L_2$ in the sector which contains $\mathrm{DF}(\mathbf{n})$. Then for all $r$, we have that $\boldsymbol{\alpha}_r^{(0)}$ and $\boldsymbol{\alpha}_r^{(1)}$ are contained in this sector, while $\boldsymbol{\alpha}_r^{(2)}$ is contained in the opposite sector (in order that $A_{\mathbf{n},r}$ contains $\boldsymbol{\alpha}$). It follows that $\theta_r < \Theta$ for all $r$, contradicting Lemma 53. □

*Remark* 55. Since $\mathcal{P}_S \subset \mathrm{Reg}_S$, Theorem 54 says nothing about smoothness of extreme points in the exceptional case. In fact, a similar but simpler argument can be used to show that the endpoints of an exceptional interval are always smooth extreme points, provided that the itinerary $\mathbf{n}$ has only finitely many zeroes.

4.4. **Subsequential digit frequencies of** $w_\beta$**.** Let $\beta > 1$ be such that the digit frequency set $\mathrm{DF}(\beta)$ is not a polytope. Then $\mathrm{DF}(\beta) = \mathrm{DF}(\mathbf{n})$, where $\mathbf{n} = \mathbf{n}(w_\beta)$, by Lemma 43. Since $\mathrm{DF}(\beta)$ is not a polytope, $\mathbf{n}$ is not of rational or finite type, so that $S(\mathbf{n}) = w_\beta$ by Definition 42 and Lemma 13.

In the regular case, when $\mathrm{DF}(\beta)$ has exactly one non-rational extreme point $\boldsymbol{\alpha}$, the sequence $S(\mathbf{n})$ has digit frequency $\boldsymbol{\alpha}$ by Theorem 33 and Facts 1f). That is, $w_\beta$ has well-defined digit frequency $\boldsymbol{\alpha}$.

In the exceptional case the digit frequency of $w_\beta = S(\mathbf{n})$ is not well defined, again by Facts 1f). In this case the interesting object is the set of subsequential digit frequencies of $w_\beta$. Let $\boldsymbol{\alpha}^{(\beta,s)}$ be the rational element of $\Delta$ giving the digit frequency of the initial subword $w_\beta\,[\![s]\!]$ of $w_\beta$, and define $F_\beta$ to be the set of limits of convergent subsequences of $\left(\boldsymbol{\alpha}^{(\beta,s)}\right)_{s\geq 1}$.

$F_\beta$ is necessarily contained in the exceptional simplex $\Phi^{-1}(\mathbf{n})$. To see this, observe that for each $r \geq 0$, the vertices of the simplex $A_{\mathbf{n},r}$ are the digit frequencies of the words $\Lambda_{\mathbf{n},r}(i)$ for $0 \leq i \leq k-1$, and $S(\mathbf{n})$ is a concatenation of these words. Therefore any subword of $S(\mathbf{n})$ which is a concatenation of the $\Lambda_{\mathbf{n},r}(i)$ has digit frequency contained in $A_{\mathbf{n},r}$, and hence an arbitrary initial subword of length $s$ has digit frequency within distance $L/s$ of $A_{\mathbf{n},r}$, where $L$ is the maximum of the lengths of the words $\Lambda_{\mathbf{n},r}(i)$. It follows that $F_\beta \subset A_{\mathbf{n},r}$ for all $r$, and so $F_\beta \subset \Phi^{-1}(\mathbf{n}) = \bigcap_{r\geq 0} A_{\mathbf{n},r}$.

A natural and seemingly difficult question is whether or not it is always the case that $F_\beta = \Phi^{-1}(\mathbf{n})$. The proof of Theorem 57 below, which treats the case $k = 3$, depends strongly on the fact that the exceptional simplex is one-dimensional, and so does not generalise to higher values of $k$.

We will need a preliminary result, that bounded itineraries are regular when $k = 3$. Since whether itineraries are regular or exceptional is connected with their rate of growth, this result appears obvious at first sight; but care has to be taken when there are many zeroes in the itinerary.

**Lemma 56.** *Let* $k = 3$ *and* $\mathbf{n} \in \mathbb{N}^{\mathbb{N}}$. *Then* $\mathbf{n}$ *is regular in each of the following two cases:*

*a) there are only finitely many* $r$ *for which both* $n_r > 0$ *and* $n_{r+1} > 0$*;*

*b)* $\mathbf{n}$ *is bounded.*

*Proof.* Given a non-negative $3 \times 3$ matrix $A$, let $f_A$ denote its projective action on $\Delta$: in other words, $f_A(\boldsymbol{\alpha}) = A\boldsymbol{\alpha}/||A\boldsymbol{\alpha}||_1$. For each $n \geq 0$, let

$$A_n = \begin{pmatrix} 0 & n+1 & n \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

be the abelianization of the substitution $\Lambda_n$, so that $f_{A_n} = K_n^{-1}\colon \Delta \to \Delta$.

By a theorem of Birkhoff [3], if $A$ is strictly positive then $f_A$ contracts the Hilbert metric $\delta$ on the interior $\mathring{\Delta}$ of $\Delta$ by a factor $(\sqrt{d(A)}-1)/(\sqrt{d(A)}+1)$, where

$$d(A) = \max_{1\le i,j,l,m\le 3} \frac{a_{il}a_{jm}}{a_{im}a_{jl}}$$

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.

Moreover (Lemma 30 of [6]) the matrices $A_n$, while not strictly positive, have the property that they do not expand the Hilbert metric: $\delta(f_{A_n}(\boldsymbol{\alpha}), f_{A_n}(\boldsymbol{\beta})) \le \delta(\boldsymbol{\alpha},\boldsymbol{\beta})$ for all $\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathring{\Delta}$.

Recall that

$$\Phi^{-1}(\mathbf{n}) = \bigcap_{r\ge 0} K_{n_0}^{-1}\circ K_{n_1}^{-1}\circ\cdots\circ K_{n_r}^{-1}(\Delta) = \bigcap_{r\ge 0} f_{A_{n_0}}\circ f_{A_{n_1}}\circ\cdots\circ f_{A_{n_r}}(\Delta).$$

In order to prove that an itinerary $\mathbf{n}$ is regular, it therefore suffices to find a constant $C$ and infinitely many disjoint subwords $n_r\dots n_{r+s}$ of $\mathbf{n}$, each having the property that the product $A_{n_r\dots n_{r+s}} = A_{n_r}\,A_{n_{r+1}}\,\cdots\,A_{n_{r+s}}$ is strictly positive and satisfies $d(A_{n_r\dots n_{r+s}})\le C$.

a) Suppose that there are only finitely many $r$ for which both $n_r>0$ and $n_{r+1}>0$. Since rational itineraries are regular, we can assume that $\mathbf{n}$ has infinitely many non-zero entries, so that it has a tail of the form $p_1\,0^{k_1}\,p_2\,0^{k_2}\,\dots$, where the $p_i$ and the $k_i$ are strictly positive. Moreover, we can assume that infinitely many of the integers $k_i$ are even, since otherwise $\Phi^{-1}(\mathbf{n})$ would be contained in the (one-dimensional) faces of $\Delta$ by Facts 1i), and hence $\mathbf{n}$ would be regular (if it were exceptional then $\Phi^{-1}(\mathbf{n})$ would be a non-trivial interval and hence would contain rational points, a contradiction).

   There are therefore infinitely many disjoint subwords of $\mathbf{n}$ which either have the form $p\,0^{2r}\,q\,0^{2s}$ or have the form $p\,0^{2r}\,q\,0^{2s-1}$, where $p$, $q$, $r$, and $s$ are positive integers. Now a straightforward induction gives that

$$A_0^{2r} = \begin{pmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ r & r & 1\end{pmatrix} \qquad\text{and}\qquad A_0^{2r-1} = \begin{pmatrix} 0 & 1 & 0\\ 1 & 0 & 0\\ r-1 & r & 1\end{pmatrix}.$$

   Therefore

$$A_p\,A_0^{2r}\,A_q\,A_0^{2s} = \begin{pmatrix} 1+p+pr+ps+pqrs & p+pr+ps+pqr+pqrs & p+pqr\\ qs & 1+q+qs & q\\ 1+r+s+qrs & 1+r+s+qr+qrs & 1+qr\end{pmatrix},$$

   a strictly positive matrix, each of whose entries is bounded below by, but no more than five times than, the corresponding entry in the matrix

$$B = \begin{pmatrix} pqrs & pqrs & pqr\\ qs & qs & q\\ qrs & qrs & qr\end{pmatrix}.$$

   Since $d(B)=1$ for all $p$, $q$, $r$, and $s$, it follows that $d(A_pA_0^{2r}A_qA_0^{2s})$ is bounded above by 25. By a similar calculation the same is true of $A_pA_0^{2r}A_qA_0^{2s-1}$ for all positive $p$, $q$, $r$, and $s$, which establishes the result.

b) Let $\mathbf{n}$ be bounded. By a direct calculation, if $n_r>0$ and $n_{r+1}>0$, then the matrix $A_{n_r}A_{n_{r+1}}A_{n_{r+2}}$ is strictly positive. Since $\mathbf{n}$ is bounded, there are only finitely many possible values for this matrix, and hence there is a constant $C$ such that $d(A_{n_r}A_{n_{r+1}}A_{n_{r+2}})\le C$ whenever $n_r$ and $n_{r+1}$ are both positive. This establishes that $\mathbf{n}$ is regular when there are infinitely many such values of $r$; and if there are only finitely many, then regularity follows from a).

□

**Theorem 57.** *Let $\beta \in (2,3)$ be such that $\mathbf{n} = \mathbf{n}(w_\beta)$ is exceptional. Then $F_\beta = \Phi^{-1}(\mathbf{n})$.*

*Proof.* Since $\beta \in (2,3)$ we have $k = 3$. Let $\ell = \Phi^{-1}(\mathbf{n})$ be the exceptional interval, and $L$ denote the length of $\ell$.

The distance between $\boldsymbol{\alpha}^{(\beta,s)}$ and $\boldsymbol{\alpha}^{(\beta,s+1)}$ is at most $1/s$, and so it is enough to prove that the two endpoints $\mathbf{v}_1$ and $\mathbf{v}_2$ of $\ell$ lie in $F_\beta$. We shall show that for all $\epsilon > 0$ there are natural numbers $r_1$ and $r_2$ such that $\boldsymbol{\alpha}_{r_i}^{(2)} \in B_\epsilon(\mathbf{v}_i)$ for each $i$, which establishes the result since $\boldsymbol{\alpha}_r^{(2)}$ is the digit frequency of the initial subword $\Lambda_{\mathbf{n},r}(2)$ of $S(\mathbf{n}) = w_\beta$.

Suppose without loss of generality that $\epsilon < L/4$, and let $R$ be large enough that $d_H(A_{\mathbf{n},r}, \ell) < \epsilon/8$ for all $r \geq R$. In particular,

$$\text{For all } r \geq R \text{ and for each } a \in \{1,2\}, \text{ there exists } j \in \{0,1,2\} \text{ with } \boldsymbol{\alpha}_r^{(j)} \in B_{\epsilon/8}(\mathbf{v}_a). \tag{13}$$

Pick $r \geq R$ with $n_{r+1} \geq 4L/\epsilon$, which is possible since the exceptional itinerary $\mathbf{n}$ is unbounded by Lemma 56. Using (4) and (5), we have that

$$\boldsymbol{\alpha}_{r+1}^{(1)} = \frac{(n_{r+1}+1)L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(2)}\boldsymbol{\alpha}_r^{(2)}}{(n_{r+1}+1)L_r^{(0)} + L_r^{(2)}} \qquad \text{and}$$

$$\boldsymbol{\alpha}_{r+1}^{(2)} = \frac{n_{r+1}L_r^{(0)}\boldsymbol{\alpha}_r^{(0)} + L_r^{(2)}\boldsymbol{\alpha}_r^{(2)}}{n_{r+1}L_r^{(0)} + L_r^{(2)}}$$

lie on the line segment joining $\boldsymbol{\alpha}_r^{(2)}$ to $\boldsymbol{\alpha}_r^{(0)}$. Since $d_H(A_{\mathbf{n},r}, \ell) < \epsilon/8$, this segment has length $\Lambda < L + \epsilon/4 < 2L$. The distances of $\boldsymbol{\alpha}_{r+1}^{(1)}$ and $\boldsymbol{\alpha}_{r+1}^{(2)}$ from $\boldsymbol{\alpha}_r^{(0)}$ are therefore given by $L_r^{(2)}\Lambda/((n_{r+1}+1)L_r^{(0)}+L_r^{(2)})$ and $L_r^{(2)}\Lambda/(n_{r+1}L_r^{(0)} + L_r^{(2)})$. Subtracting these gives $d(\boldsymbol{\alpha}_{r+1}^{(1)}, \boldsymbol{\alpha}_{r+1}^{(2)}) < \Lambda/n_{r+1} < 2L/(4L/\epsilon) = \epsilon/2$. It follows from (13), using $L > 4\epsilon$, that $\boldsymbol{\alpha}_{r+1}^{(0)}$ is within $\epsilon/8$ of one of the endpoints of $\ell$, say $\mathbf{v}_1$; while both $\boldsymbol{\alpha}_{r+1}^{(1)}$ and $\boldsymbol{\alpha}_{r+1}^{(2)}$ lie in $B_\epsilon(\mathbf{v}_2)$.

Since $\boldsymbol{\alpha}_{r+1}^{(2)} \in B_\epsilon(\mathbf{v}_2)$, it remains to find $r'$ with $\boldsymbol{\alpha}_{r'}^{(2)} \in B_\epsilon(\mathbf{v}_1)$. We shall show that, if $\boldsymbol{\alpha}_s^{(0)} \in B_{\epsilon/8}(\mathbf{v}_1)$ and $\boldsymbol{\alpha}_s^{(1)}, \boldsymbol{\alpha}_s^{(2)} \in B_\epsilon(\mathbf{v}_2)$ for some $s$, then

a) if $n_{s+1} = 0$ then the same conditions hold when $s$ is replaced with $s+2$; and
b) if $n_{s+1} > 0$ then $\boldsymbol{\alpha}_{s+1}^{(2)} \in B_\epsilon(\mathbf{v}_1)$.

This will complete the proof, since there is some $p \geq 1$ for which $n_{r+2p} > 0$; for otherwise, by Facts 1i), the exceptional interval $\ell$ would be contained in one of the (one-dimensional) faces of $\Delta$, contradicting the fact that it contains no rational points.

For a), observe that $\boldsymbol{\alpha}_{s+1}^{(0)} = \boldsymbol{\alpha}_s^{(1)} \in B_\epsilon(\mathbf{v}_2)$ and, since $n_{s+1} = 0$, we have $\boldsymbol{\alpha}_{s+1}^{(2)} = \boldsymbol{\alpha}_s^{(2)} \in B_\epsilon(\mathbf{v}_2)$. By (13), we have $\boldsymbol{\alpha}_{s+1}^{(1)} \in B_{\epsilon/8}(\mathbf{v}_1)$.

Then $\boldsymbol{\alpha}_{s+2}^{(0)} = \boldsymbol{\alpha}_{s+1}^{(1)} \in B_{\epsilon/8}(\mathbf{v}_1)$, and both $\boldsymbol{\alpha}_{s+2}^{(1)}$ and $\boldsymbol{\alpha}_{s+2}^{(2)}$ lie on the line segment joining $\boldsymbol{\alpha}_{s+1}^{(2)}$ to $\boldsymbol{\alpha}_{s+1}^{(0)}$, which is contained in $B_\epsilon(\mathbf{v}_2)$, as required.

For b), we have as in a) that $\boldsymbol{\alpha}_{s+1}^{(0)} = \boldsymbol{\alpha}_s^{(1)} \in B_\epsilon(\mathbf{v}_2)$. Both $\boldsymbol{\alpha}_{s+1}^{(1)}$ and $\boldsymbol{\alpha}_{s+1}^{(2)}$ lie on the line segment joining $\boldsymbol{\alpha}_s^{(2)} \in B_\epsilon(\mathbf{v}_2)$ to $\boldsymbol{\alpha}_s^{(0)} \in B_{\epsilon/8}(\mathbf{v}_1)$, and $\boldsymbol{\alpha}_{s+1}^{(1)}$ is closer than $\boldsymbol{\alpha}_{s+1}^{(2)}$ to $\boldsymbol{\alpha}_s^{(0)}$. Therefore $\boldsymbol{\alpha}_{s+1}^{(1)} \in$

$B_{\epsilon/8}(\mathbf{v}_1)$ by (13). Calculating the ratio of the distances from $\boldsymbol{\alpha}_{s+1}^{(1)}$ and $\boldsymbol{\alpha}_{s+1}^{(2)}$ to $\boldsymbol{\alpha}_s^{(0)}$ gives

$$\frac{d(\boldsymbol{\alpha}_s^{(0)}, \boldsymbol{\alpha}_{s+1}^{(2)})}{d(\boldsymbol{\alpha}_s^{(0)}, \boldsymbol{\alpha}_{s+1}^{(1)})} = \frac{(n_{s+1}+1)L_s^{(0)} + L_s^{(2)}}{n_{s+1}L_s^{(0)} + L_s^{(2)}} < \frac{n_{s+1}+1}{n_{s+1}} \leq 2$$

since $n_{s+1} \geq 1$. Therefore $d(\boldsymbol{\alpha}_s^{(0)}, \boldsymbol{\alpha}_{s+1}^{(2)}) < \epsilon/2$, so that $d(\boldsymbol{\alpha}_{s+1}^{(2)}, \mathbf{v}_1) < \epsilon/2 + \epsilon/8 < \epsilon$ as required. □

4.5. **Calculations for specific values of $\beta$.** We finish by addressing the practical problem of computing the function $\beta \mapsto \mathbf{n}(w_\beta)$, so that digit frequency sets of specific numbers $\beta$ can be determined.

We first define a left inverse $\Gamma_n$ of each $\Lambda_n : \Sigma \to \Sigma$. Informally, to determine $\Gamma_n(w)$ we repeatedly remove $\Lambda_n$-images of digits from the front of $w$ until we are no longer able to do so: at that stage we complete $\Gamma_n(w)$ with $\overline{0}$ if the remaining block is smaller than anything in the image of $\Lambda_n$, and with $\overline{k-1}$ if the remaining block is larger than anything in the image of $\Lambda_n$.

**Definition 58** ($\Gamma_n : \Sigma \to \Sigma$)**.** For each $n \in \mathbb{N}$, define $\Gamma_n : \Sigma \to \Sigma$ as follows. Let $w \in \Sigma$. Then

- If $w_0 = 0$ then $\Gamma_n(w) = \overline{0}$.
- If $1 \leq w_0 \leq k-2$ then $\Gamma_n(w) = (w_0 - 1)\,\Gamma_n(\sigma(w))$.
- If $w = (k-1)\,0^{n+1}\,v$ then $\Gamma_n(w) = (k-2)\,\Gamma_n(v)$.
- If $w = (k-1)\,0^n\,v$ with $v_0 > 0$ then $\Gamma_n(w) = (k-1)\,\Gamma_n(v)$.
- If $w_0 = k-1$ and there is some $1 \leq r \leq n$ with $w_r \neq 0$ then $\Gamma_n(w) = \overline{k-1}$.

Notice that $\Gamma_n$ is increasing (the five cases in its definition are listed in order of increasing $w$, and the five corresponding outputs are also in increasing order) and continuous (if $w$ and $w'$ agree to $r(n+2)$ digits for any $r$, then $\Gamma_n(w)$ and $\Gamma_n(w')$ agree to $r$ digits).

The following lemma gives a recursive algorithm for calculating $\mathbf{n}(w)$ in the rational or finite case, or for reading off successive entries of $\mathbf{n}(w)$ in the general case.

**Lemma 59.** *Let $w \in \Sigma$ with $w_0 = k-1$. If $w = (k-1)\,\overline{0}$ then $\mathbf{n}(w) = \infty$. Otherwise, let $n \geq 0$ be least such that $w \geq (k-1)\,0^n\,\overline{1}$. Then $\mathbf{n}(w) = n\,\mathbf{n}(\Gamma_n(w))$.*

*Proof.* Recall that $\mathbf{n}(w) = \max\{\mathbf{m} \in \mathcal{N} : S(\mathbf{m}) \leq w\}$. Since $S(\infty) = (k-1)\,\overline{0}$ and $S(N\,\infty) = (k-1)\,0^N\,\overline{1}$ for each $N \in \mathbb{N}$, we have that $\mathbf{n}(w) = \infty$ if and only if $w = (k-1)\,\overline{0}$. We henceforth assume that this is not the case.

Since $n\,\infty$ is the smallest element of $\mathcal{N}$ which starts with $n$, it is immediate that $n_0(w) = n$, where $n \geq 0$ is least such that $w \geq S(n\,\infty) = (k-1)\,0^n\,\overline{1}$.

Then, observing that $\Gamma_n(S(n\,\mathbf{m})) = \Gamma_n(\Lambda_n(S(\mathbf{m}))) = S(\mathbf{m})$, since $\Gamma_n$ is a left inverse of $\Lambda_n$, we have

$$\begin{aligned}
\sigma(\mathbf{n}(w)) &= \max\{\mathbf{m} \in \mathcal{N} : S(n\,\mathbf{m}) \leq w\} \\
&= \max\{\mathbf{m} \in \mathcal{N} : \Gamma_n(S(n\,\mathbf{m})) \leq \Gamma_n(w)\} \\
&= \max\{\mathbf{m} \in \mathcal{N} : S(\mathbf{m}) \leq \Gamma_n(w)\} \\
&= \mathbf{n}(\Gamma_n(w))
\end{aligned}$$

which establishes the result, once we have proved the second equality above.

For this equality we need to show that, for all $\mathbf{m} \in \mathcal{N}$, we have $S(n\,\mathbf{m}) \leq w$ if and only if $\Gamma_n(S(n\,\mathbf{m})) \leq \Gamma_n(w)$. The "only if" statement is immediate since $\Gamma_n$ is increasing. For the other direction, we use again that $\Gamma_n$ is increasing to observe that if $\Gamma_n(S(n\,\mathbf{m})) \leq \Gamma_n(w)$ then either $S(n\,\mathbf{m}) \leq w$ as required, or $\Gamma_n(w) = \Gamma_n(S(n\,\mathbf{m})) = S(\mathbf{m})$. In the latter case, there are two possibilities:

a) $w = \Lambda_n(v)$ is in the image of $\Lambda_n$. Then $S(\mathbf{m}) = \Gamma_n(w)$ becomes $S(\mathbf{m}) = v$, so that $S(n\,\mathbf{m}) = \Lambda_n(v) = w$ as required.

b) $w$ is not in the image of $\Lambda_n$, so that $\Gamma_n(w)$ ends with $\overline{0}$ or with $\overline{k-1}$ by definition of $\Gamma_n$. Since $\Gamma_n(w) = S(\mathbf{m})$, the only possibilities are that $\Gamma_n(w) = \overline{k-1}$ and $\mathbf{m} = \overline{0}$; or that $\Gamma_n(w) = (k-1)\,\overline{0}$ and $\mathbf{m} = \infty$.

If the former of these two possibilities holds, then we have $S(n\,\mathbf{m}) = \overline{(k-1)\,0^n}$, while $\Gamma_n(w) = \overline{k-1}$ gives either $w = \overline{(k-1)\,0^n}$, or $w = ((k-1)\,0^n)^s\,(k-1)\,0^r\,v$ for some $s \geq 0$, $r < n$, and $v \in \Sigma$ with $v_0 > 0$. It follows that $S(n\,\mathbf{m}) \leq w$ as required.

In the latter case, on the other hand, we have $S(n\,\mathbf{m}) = (k-1)\,0^n\,\overline{1}$, while $\Gamma_n(w) = (k-1)\,\overline{0}$ gives either $w = (k-1)\,0^n\,\overline{1}$ – in which case $S(n\,\mathbf{m}) \leq w$ as required – or $w = (k-1)\,0^n\,1^s\,0\,v$ for some $s \geq 1$ and $v \in \Sigma$, which contradicts $w \geq (k-1)\,0^n\,\overline{1}$.

□

**Example 60.** Let $\beta = 2.1901$ so that $k = 3$. Calculating the orbit entries $f_\beta^r(1)$ for $0 \leq r \leq 12$, we find (and could establish rigorously) that $w_\beta$ starts with the digits $2\,0\,0\,1\,2\,0\,0\,1\,2\,0\,0\,0\,0\ldots$. Applying the algorithm of Lemma 59, we have

- $2\,0\,0\,\overline{1} \leq w_\beta < 2\,0\,\overline{1}$. Therefore $\mathbf{n}(w_\beta) = 2\,\mathbf{n}(\Gamma_2(w_\beta))$.
- $\Gamma_2(w_\beta) = 2\,0\,2\,0\,1\,\overline{0}$, so that $2\,0\,\overline{1} \leq \Gamma_2(w_\beta) < 2\,\overline{1}$. Therefore $\mathbf{n}(w_\beta) = 2\,1\,\mathbf{n}(\Gamma_1\Gamma_2(w_\beta))$.
- $\Gamma_1\Gamma_2(w_\beta) = 2\,2\,\overline{0}$, so that $2\,\overline{1} \leq w_\beta$. Therefore $\mathbf{n}(w_\beta) = 2\,1\,0\,\mathbf{n}(\Gamma_0\Gamma_1\Gamma_2(w_\beta))$.
- $\Gamma_0\Gamma_1\Gamma_2(w_\beta) = 2\,1\,\overline{0}$, so that $2\,0\,\overline{1} \leq \Gamma_0\Gamma_1\Gamma_2(w_\beta) < 2\,\overline{1}$. Therefore $\mathbf{n}(w_\beta) = 2\,1\,0\,1\,\mathbf{n}(\Gamma_1\Gamma_0\Gamma_1\Gamma_2(w_\beta))$.
- $\Gamma_1\Gamma_0\Gamma_1\Gamma_2(w_\beta) = \overline{2}$, so that $\mathbf{n}(\Gamma_1\Gamma_0\Gamma_1\Gamma_2(w_\beta)) = \overline{0}$.

In summary, $\mathbf{n}(w_\beta) = 2\,1\,0\,1\,\overline{0}$, from which we can compute $\mathrm{DF}(\beta)$ as in Example 36a).

**Example 61.** Consider the example with $d_\beta(1) = 2\,1\,2\,1\,\overline{0}$, and hence $w_\beta = \overline{2\,1\,2\,0}$, which was treated in Section 1 using Markov partition techniques. The algorithm of Lemma 59 gives $\mathbf{n}(w_\beta) = 0\,1\,1\,\overline{0}$, so that, by Theorem 33, the non-trivial extreme points of $\mathrm{DF}(\beta)$ are

$$\begin{aligned} K_0^{-1}(0,1,0) &= (1/2,\,0,\,1/2), \\ K_0^{-1}K_1^{-1}(0,1,0) &= (0,\,2/3,\,1/3), \text{ and} \\ K_0^{-1}K_1^{-1}K_1^{-1}(0,0,1) &= (1/4,\,1/4,\,1/2), \end{aligned}$$

in agreement with the Markov partition calculation.

## References


1. V. Berthé and V. Delecroix, *Beyond substitutive dynamical systems: S-adic expansions*, arXiv:1309.3960 (2013).
2. A. Besicovitch, *On the sum of digits of real numbers represented in the dyadic system. (On sets of fractional dimensions II.).*, Math. Ann. **110** (1934), 321–330 (English).
3. G. Birkhoff, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc. **85** (1957), 219–227. MR 0087058 (19,296a)
4. V. Borovikov, *On the intersection of a sequence of simplexes*, Uspehi Matem. Nauk (N.S.) **7** (1952), no. 6(52), 179–180. MR 0053505 (14,784d)
5. P. Boyland, A. de Carvalho, and T. Hall, *New rotation sets in a family of torus homeomorphisms*, arXiv:1410.7727 (2014).
6. ———, *Symbol ratio minimax sequences in the lexicographic order*, Ergodic Theory Dynam. Systems (in press), DOI: http://dx.doi.org/10.1017/etds.2014.44.
7. H. Bruin and S. Troubetzkoy, *The Gauss map on a class of interval translation mappings*, Israel J. Math. **137** (2003), 125–148. MR 2013352 (2004j:37007)
8. D. P. Chi and D. Kwon, *Sturmian words, β-shifts, and transcendence*, Theoret. Comput. Sci. **321** (2004), no. 2-3, 395–404. MR 2076154 (2005e:68195)

9. H. Eggleston, *The fractional dimension of a set defined by decimal properties*, Quart. J. Math., Oxford Ser. **20** (1949), 31–36. MR 0031026 (11,88e)
10. P. Erdös, I. Joó, and V. Komornik, *Characterization of the unique expansions* $1 = \sum_{i=1}^{\infty} q^{-n_i}$ *and related problems*, Bull. Soc. Math. France **118** (1990), no. 3, 377–390. MR 1078082 (91j:11006)
11. D. Fried, *The geometry of cross sections to flows*, Topology **21** (1982), no. 4, 353–371. MR 670741 (84d:58068)
12. D. Kwon, *A devil's staircase from rotations and irrationality measures for Liouville numbers*, Math. Proc. Cambridge Philos. Soc. **145** (2008), no. 3, 739–756. MR 2464787 (2009k:37091)
13. M. Lothaire, *Algebraic combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002, A collective work by Jean Berstel, Dominique Perrin, Patrice Seebold, Julien Cassaigne, Aldo De Luca, Steffano Varricchio, Alain Lascoux, Bernard Leclerc, Jean-Yves Thibon, Veronique Bruyere, Christiane Frougny, Filippo Mignosi, Antonio Restivo, Christophe Reutenauer, Dominique Foata, Guo-Niu Han, Jacques Desarmenien, Volker Diekert, Tero Harju, Juhani Karhumaki and Wojciech Plandowski, With a preface by Berstel and Perrin. MR 1905123 (2003i:68115)
14. W. Parry, *On the $\beta$-expansions of real numbers*, Acta Math. Acad. Sci. Hungar. **11** (1960), 401–416. MR 0142719 (26 #288)
15. V. Prasolov, *Polynomials*, Algorithms and Computation in Mathematics, vol. 11, Springer-Verlag, Berlin, 2010, Translated from the 2001 Russian second edition by Dimitry Leites, Paperback edition [of MR2082772]. MR 2683151 (2011g:12001)
16. A. Rényi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungar. **8** (1957), 477–493. MR 0097374 (20 #3843)
17. J. Schmeling, *Symbolic dynamics for $\beta$-shifts and self-normal numbers*, Ergodic Theory Dynam. Systems **17** (1997), no. 3, 675–694. MR 1452189 (98c:11080)
18. N. Sidorov, *Almost every number has a continuum of $\beta$-expansions*, Amer. Math. Monthly **110** (2003), no. 9, 838–842. MR 2024754 (2004i:11085)
19. G. Winkler, *Choquet order and simplices with applications in probabilistic models*, Lecture Notes in Mathematics, vol. 1145, Springer-Verlag, Berlin, 1985. MR 808401 (88f:46019)
20. K. Ziemian, *Rotation sets for subshifts of finite type*, Fund. Math. **146** (1995), no. 2, 189–201. MR 1314983 (96b:58072)

Department of Mathematics, University of Florida, 372 Little Hall, Gainesville, FL 32611-8105, USA
*E-mail address*: `boyland@ufl.edu`

Departamento de Matemática Aplicada, IME-USP, Rua Do Matão 1010, Cidade Universitária, 05508-090 São Paulo SP, Brazil
*E-mail address*: `andre@ime.usp.br`

Department of Mathematical Sciences, University of Liverpool, Liverpool L69 7ZL, UK
*E-mail address*: `tobyhall@liv.ac.uk`