

AN INFINITE INTERVAL VERSION OF THE α -KAKUTANI EQUIDISTRIBUTION PROBLEM

BY

MARK POLLICOTT*

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
e-mail: masdbl@warwick.ac.uk

AND

BENEDICT SEWELL**

Alfréd Rényi Institute of Mathematics, 13-15 Reáltanoda St., Budapest, Hungary
e-mail: sewell@renyi.hu

ABSTRACT

In this article we extend results of Kakutani, Adler–Flatto, Smilansky and others on the classical α -Kakutani equidistribution result for sequences arising from finite partitions of the interval. In particular, we describe a generalization of the equidistribution result to infinite partitions. In addition, we give discrepancy estimates, extending results of Drmota–Infusino [8].

1. Introduction

Uniform distribution of sequences of numbers $(x_n)_{n=1}^\infty$ in the unit interval has been an important area of interest for over a century. For example, it was shown by Weyl [22] that, for any irrational α , the sequence $x_n = \alpha n \pmod{1}$ is uniformly distributed and Hardy and Littlewood showed that, for almost all $\lambda > 1$, the sequence $x_n = \lambda^n \pmod{1}$ is uniformly distributed [14]. In this note we consider another simple family of examples based on subdividing intervals. Before introducing the original motivating example, we first fix our terminology: a **partition** $\mathcal{P}$ is a set of closed, positive-length intervals, which have pairwise disjoint interiors and cover $[0, 1]$ up to a set of Lebesgue measure zero.

* Partly supported by ERC-Advanced Grant 833802-Resonances and EPSRC grant EP/T001674/1.

** Partly supported by the Alfréd Rényi Young Researcher fund.

Received April 7, 2021 and in revised form March 12, 2022

1.1. THE ORIGINAL α -KAKUTANI EQUIDISTRIBUTION RESULT. In 1973 Araki posed a problem which lead to Kakutani to prove an elegant equidistribution result [17]. (An interesting historical background is presented in [1].) For clarity, we give a description of his original partition scheme, which we generalise in the next section.

Definition (α -Kakutani scheme): For a fixed $0 < \alpha < 1$, the α -**Kakutani scheme** is a sequence of partitions $(\mathcal{P}_n)_{n=0}^\infty$ defined inductively:

- $\mathcal{P}_0 = \{[0, 1]\}$ is the trivial partition; and
- $\mathcal{P}_{n+1}$ is obtained from $\mathcal{P}_n$ by taking each interval of maximal length and subdividing it into two smaller intervals in the ratio $\alpha : 1 - \alpha$.

Example 1: Figure 1 shows the first seven partitions for the choice $\alpha = 1/3$. Notice that $\mathcal{P}_5$ is obtained by splitting two maximal length intervals in $\mathcal{P}_4$ simultaneously (each of length $2/9$). By contrast, the choice $\alpha = 1/2$ gives the trivial dyadic splitting.

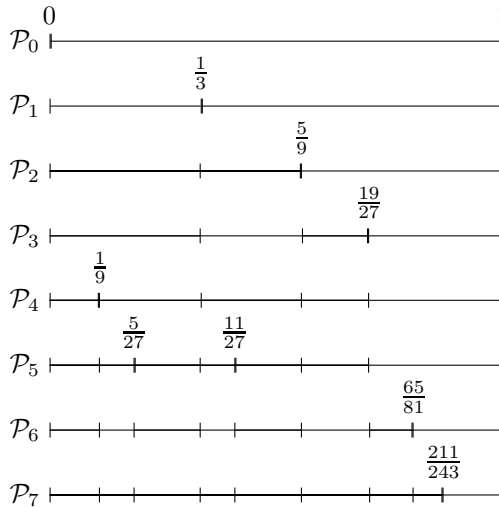


Figure 1. The first seven partitions $(\mathcal{P}_n)_{n=0}^7$ of the $\frac{1}{3}$ -Kakutani scheme.

Consider the set of endpoints at the n th stage of this process, $E_n = \bigcup_{I \in \mathcal{P}_n} \partial I$. Kakutani's result is the following.

THEOREM (Kakutani): *For all $\alpha \in (0, 1)$, the set E_n becomes uniformly distributed as $n \rightarrow \infty$.*

1.2. INTERVAL SUBSTITUTIONS USING MULTIPLE INTERVALS. A natural generalisation of the α -Kakutani scheme, first introduced by Volčič in [20], is to alter the above process by splitting intervals of maximal length according to a fixed finite partition consisting of $N \geq 2$ subintervals, say. That is, at each stage, one splits all intervals of maximal length into N pieces whose lengths have a certain fixed ratio, $\alpha_1 : \alpha_2 : \dots : \alpha_N$, where the α_i sum to 1.

Example 2: In Figure 2, we have the first seven partitions of the interval substitution scheme in which one splits maximal intervals according to the partition $\{[0, \frac{1}{2}], [\frac{1}{2}, \frac{2}{3}], [\frac{2}{3}, 1]\}$, i.e., with ratio $\frac{1}{2} : \frac{1}{6} : \frac{1}{3}$. By contrast, the α -Kakutani scheme corresponds to splitting according to the partition $\{[0, \alpha], [\alpha, 1]\}$.

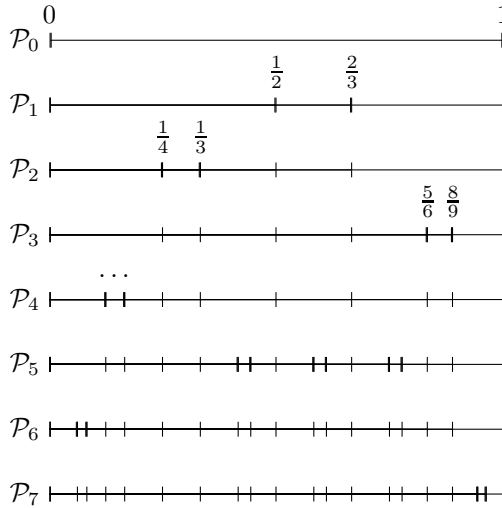


Figure 2. The first seven partitions $(\mathcal{P}_n)_{n=0}^7$ of the interval substitution scheme where one splits maximal-length intervals according to the partition $\mathcal{P}_1 = \{[0, \frac{1}{2}], [\frac{1}{2}, \frac{2}{3}], [\frac{2}{3}, 1]\}$.

One particular family of these schemes, that of so-called LS-sequences, which include the ϕ -Kakutani scheme for $\phi = \frac{1}{2}(\sqrt{5}-1)$, has received particular attention in the context of low discrepancy sequences: see, e.g., [3, 5, 6, 15, 16, 21].

1.3. INTERVAL SUBSTITUTIONS USING INFINITELY MANY INTERVALS. In this note, we continue the process and ask in what way the result above still holds if at every stage we insert an *infinite* partition $\mathcal{P}$ into each maximal-length subinterval. We denote by $\hat{E}_n$ the finite set of endpoints of those intervals which have been split up to the $(n+1)$ -st stage, i.e.,

$$\hat{E}_n := \left\{ \min(I), \max(I) : I \in \bigcup_{i=0}^n \mathcal{P}_i \setminus \mathcal{P}_{n+1} \right\}.$$

Example 3: In Figure 3 we depict $\mathcal{P}_n$ and $\hat{E}_n$ (for $n \leq 7$) for the infinite substitution scheme generated by $\mathcal{P} = \{[0, \frac{1}{2}]\} \cup \{[1 - \frac{1}{2} \cdot 3^{-n}, 1 - \frac{1}{6} \cdot 3^{-n}]\}_{n=0}^{\infty}$.

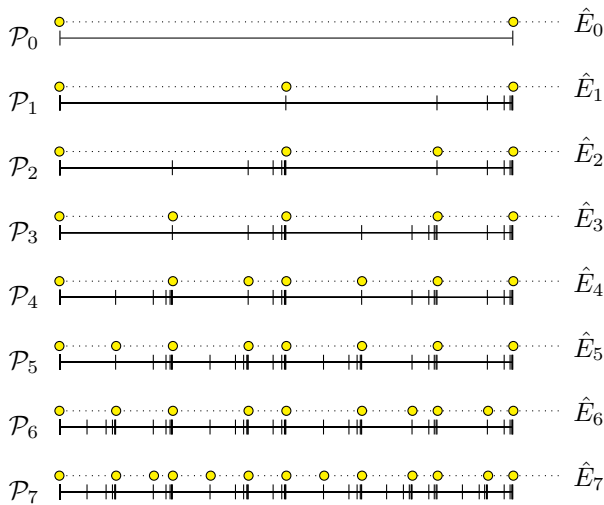


Figure 3. An illustration of $(\mathcal{P}_n)_{n=0}^7$ and $(\hat{E}_n)_{n=0}^7$ for the example generated by the partition $\mathcal{P} = \{[1, \frac{1}{2}]\} \cup \{[1 - \frac{1}{2} \cdot 3^{-n}, 1 - \frac{1}{6} \cdot 3^{-n}]\}_{n=0}^{\infty}$. Here the tick marks (which accumulate on certain points in the interval) denote the elements of E_n and the suspended yellow circles denote the elements of $\hat{E}_n$.

For simplicity, we restrict our attention to the set of left endpoints, which we shall denote by L_n , although we could equally well have chosen the right endpoints, midpoints, etc.

Our main result is the following generalisation of Kakutani's equidistribution theorem. Let $\|I\|$ denote the length of an interval I .

THEOREM 1: *Let $\mathcal{P}$ be a countable partition. Then, provided that*

$$-\sum_{I \in \mathcal{P}} \|I\| \log \|I\| < \infty,$$

the set L_n becomes uniformly distributed as $n \rightarrow \infty$.

A brief outline of this note: In section 2, we give a new dynamical viewpoint of the problem and in section 3, we apply renewal theory to prove Theorem 1 in two cases. In section 4 we use a generating function to estimate the discrepancy in the rank one case. In section 5 we use methods of analytic number theory to estimate the discrepancy in the higher rank case, with a generic Diophantine-type assumption.

Our interest in this problem, and the starting point for our analysis, began with the very elegant work of Smilansky [19].

2. Partitions and similarities

Our approach to Theorem 1 is to express the elements of the partition $\mathcal{P}$ in terms of the images of similarities. The refinements into finer partitions, $\mathcal{P}_n$, can then be expressed in terms of words formed from the index set of $\mathcal{P}$.

That is, one can write $\mathcal{P} = \{T_i[0, 1]\}_{i \in \mathcal{I}}$, where each $T_i : [0, 1] \rightarrow [0, 1]$ is an orientation preserving similarity with contraction ratio $\alpha_i > 0$. We see that $\mathcal{P}$ being a partition is equivalent to the following:

- $T_i[0, 1) \cap T_j[0, 1) = \emptyset$ for $i \neq j$; and
- $\sum_{i \in \mathcal{I}} \alpha_i = 1$.

Example 4: Given $0 = t_0 < t_1 < \dots$ with $(t_n) \rightarrow 1$, the partition $\{[t_{n-1}, t_n]\}_{n \in \mathbb{N}}$ is equal to $\{T_n[0, 1]\}_{n \in \mathbb{N}}$, where

$$T_n(x) = (t_n - t_{n-1})x + \sum_{k=0}^{n-1} t_k.$$

Note that 1 is not contained in any $T_n[0, 1]$, and a mild adaptation gives examples for which 0 is not contained in any of these images.

Example 5: Setting $t_n = 1 - \frac{1}{6}3^{-n}$ for $n \geq 1$ gives rise to $\mathcal{P}_1$ in Figure 3.

We now explain how this can be used to give an explicit description of the splitting process.

Definition ((T_i) -refinement): Given a partition $\mathcal{P} = \{S_k[0, 1]\}_k$, where the $\{S_k\}_k$ are orientation preserving similarities, the (T_i) -**refinement** of $\mathcal{P}$ is the refinement obtained by taking all intervals of maximal length in $\mathcal{P}$ and replacing them by subintervals in the following manner: if $S[0, 1] \in \mathcal{P}$ has maximal length in $\mathcal{P}$, then it is replaced by the elements of the set

$$\{S \circ T_i[0, 1] \mid i \in \mathcal{I}\}.$$

Definition (Interval substitution scheme): The **interval substitution scheme** generated by $\{T_i\}_{i \in \mathcal{I}}$ is the sequence of partitions $(\mathcal{P}_n)_{n=0}^\infty$ defined as follows:

- $\mathcal{P}_0$ is the trivial partition, $\mathcal{P}_0 = \{[0, 1]\}$; and
- $\mathcal{P}_{n+1}$ is the (T_i) -refinement of $\mathcal{P}_n$.

This gives a convenient presentation of the partitions.

Example 6: The α -Kakutani scheme is the interval substitution scheme generated by the pair $T_1 : x \mapsto \alpha x$ and $T_2 : x \mapsto (1 - \alpha)x + \alpha$.

Example 7: Similarly, the interval substitution scheme generated by the triple $T_1 : x \mapsto x/2$, $T_2 : x \mapsto (x + 3)/6$ and $T_3 : x \mapsto (x + 2)/3$ gives the sequence of partitions depicted in Figure 2.

We now associate to the sequence of partitions $(\mathcal{P}_n)_{n=0}^\infty$ a sequence of families of left endpoints of split intervals, $(L_n)_{n=0}^\infty$.

Definition (L_n) : Given an interval substitution scheme $(\mathcal{P}_n)_{n=0}^\infty$ generated by similarities $(T_i)_{i \in \mathcal{I}}$, we define the finite sets L_n ($n \geq 0$) to be

$$L_n = \bigcup_{k=0}^n \bigcup_{I \in \mathcal{P}_k \setminus \mathcal{P}_{k+1}} \min(I).$$

Remark: One can consider a generalisation of the above process by dropping the assumption that the $(T_i)_i$ have to be affine. We will not consider this more general setting, but we note it is easy to give superficial examples where E_n (or L_n) is not uniformly distributed: take, for example, $T_1(x) = \sqrt{x}/2$, $T_2(x) = (x + 1)/2$.

Considering the interval substitution scheme generated by $\{T_i\}_{i \in \mathcal{I}}$, it follows inductively that every interval appearing in the process is obtained by applying a sequence of maps from $\{T_i\}_i$ to $[0, 1]$, and so each is naturally described by a finite word in $\mathcal{I}$. It is convenient to introduce the following notation.

Definition ($W(\mathcal{I})$, $*$, $\alpha_{\mathbf{v}}$, $T_{\mathbf{v}}$): Given a countable set $\mathcal{I}$, the word set $W(\mathcal{I})$ is the semigroup consisting of all words in $\mathcal{I}$: i.e.,

$$W(\mathcal{I}) = \{\emptyset\} \cup \bigcup_{n=1}^{\infty} \mathcal{I}^n,$$

where $\emptyset$ denotes the empty word (unique word of length zero), and the semigroup operation $*$: $W(\mathcal{I}) \times W(\mathcal{I}) \rightarrow W(\mathcal{I})$ denotes concatenation of words, for which $\emptyset$ acts as the identity:

$$(n_1, \dots, n_k) * (m_1, \dots, m_j) = (n_1, \dots, n_k, m_1, \dots, m_j); \quad \mathbf{v} * \emptyset = \emptyset * \mathbf{v} = \mathbf{v}.$$

Furthermore, for ease of notation, we extend the definitions of α_i and T_i to the whole of $W(\mathcal{I})$: For the word $\mathbf{v} = (i_1, \dots, i_k) \in \mathcal{I}^k$, define

$$\alpha_{\mathbf{v}} := \prod_{j=1}^k \alpha_{i_j}, \quad T_{\mathbf{v}} := T_{i_1} \circ \dots \circ T_{i_k},$$

and also define $\alpha_{\emptyset} = 1$ and $T_{\emptyset} = \text{Id}_{[0,1]}$.

That is, given any interval I which appears in the process, $I = T_{\mathbf{v}}[0, 1]$ for some word $\mathbf{v} \in W(\mathcal{I})$,¹ and will be split between $\mathcal{P}_n$ and $\mathcal{P}_{n+1}$ (i.e., $I \in \mathcal{P}_n \setminus \mathcal{P}_{n-1}$) precisely when n satisfies

$$\alpha_{\mathbf{v}} = \max_{I \in \mathcal{P}_n} \{\|I\|\},$$

and consequently its left endpoint $T_{\mathbf{v}}(0)$ will appear in L_n , if not already present in L_{n-1} .

Rather than using n in $\{L_n\}_{n \geq 1}$ to parameterise this process, we want to reparameterise this family to reflect the lengths of the maximal intervals, and rewrite it as (X_{λ}) as follows.

Definition: For $\lambda > 1$, let $X_{\lambda} = \emptyset$, and for $\lambda \in (0, 1]$, let

$$X_{\lambda} := L_{n(\lambda)}, \quad \text{where } n(\lambda) := \max\{n \geq 0 : \exists I \in \mathcal{P}_n, \|I\| \geq \lambda\}.$$

¹ Similarly, every word $\mathbf{v} \in \mathcal{I}$ gives rise to an interval in some $\mathcal{P}_n$.

I.e., given $\lambda \leq 1$, $\mathcal{P}_{n(\lambda)+1}$ is the first partition for which all intervals have lengths strictly smaller than λ . From the previous discussion, one obtains a convenient, dynamical formula for X_λ :

$$X_\lambda = \{T_v(0) : v \in W(\mathcal{I}), \alpha_v \geq \lambda\}.$$

As $\lambda \rightarrow 0^+$, $n(\lambda) \rightarrow \infty$ and so the uniform distribution of L_n as $n \rightarrow \infty$ is equivalent to that of X_λ as $\lambda \rightarrow 0^+$. As an aside, this is equivalent to the (weak-*) convergence of the probability measure μ_λ ,

$$\mu_\lambda = \frac{1}{|X_\lambda|} \sum_{x \in X_\lambda} \delta_x,$$

to the Lebesgue measure (which we will henceforth denote as Leb), as $\lambda \rightarrow 0^+$, where δ_x denotes the Dirac delta measure at x , and $|\cdot|$ denotes cardinality (of a finite set).

3. Proof of Theorem 1

This section is devoted to proving the main result, which we can conveniently rephrase in the following way.

THEOREM 1: *Provided that $-\sum_{I \in \mathcal{P}} \|I\| \log \|I\| < \infty$, the measures μ_λ converge to the Lebesgue measure Leb as Borel measures as $\lambda \rightarrow 0^+$, i.e., for any interval $J \subset [0, 1]$, we have $\mu_\lambda(J) \rightarrow \text{Leb}(J) = \|J\|$.*

We first prove the convergence for a given interval of the form $T_v[0, 1)$. For each of these elementary sets, their μ_λ -measure is intimately related to the asymptotics of $|X_\lambda|$ as $\lambda \rightarrow 0^+$. To proceed, we can relate $|X_\lambda|$ to the set of words $|A_\lambda|$, where

$$A_\lambda := \{w \in W(\mathcal{I}) | \alpha_w \geq \lambda\},$$

and exploit a natural renewal equation for the quantity $\lambda|A_\lambda|$.

LEMMA 1: *One of the following two cases hold: Either there is a (unique) symbol in $1 \in \mathcal{I}$, say, such that $T_1(0) = 0$; in which case, for all $\lambda > 0$,*

$$(1) \quad |A_\lambda| = \sum_{k=0}^{\infty} |X_{\lambda\alpha_1^{-k}}| \iff |X_\lambda| = |A_\lambda| - |A_{\lambda/\alpha_1}|;$$

or no element of $\{T_w\}_{w \in W(\mathcal{I})}$ fixes 0, and $|X_\lambda| = |A_\lambda|$ for all $\lambda > 0$.

Proof of Lemma 1. Let $\mathbf{w}, \mathbf{v} \in A_\lambda$ satisfy $T_{\mathbf{w}}(0) = T_{\mathbf{v}}(0)$. If $\mathbf{w} \neq \mathbf{v}$, the disjointness of $T_i[0, 1)$ and the injectivity of the T_i imply inductively that one of these words is obtained from the other by concatenation with a word fixing the identity; i.e., without loss of generality, $\mathbf{w} = \mathbf{v} * \mathbf{j}$, where $\mathbf{j} \in W(\mathcal{I})$ satisfies $T_{\mathbf{j}}(0) = 0$. We now have the two cases:

If there is no symbol $i \in \mathcal{I}$ for which $T_i(0) = 0$, we must have $\mathbf{j} = \emptyset$; which implies the map $\mathbf{v} \mapsto T_{\mathbf{v}}(0)$ is a bijection $A_\lambda \rightarrow X_\lambda$.

For the other case, let $1 \in \mathcal{I}$ be such that $T_1(0) = 0$; it is unique by the disjointness of $T_i[0, 1)$. Then $\mathbf{j} \neq \emptyset$ only if $\mathbf{j} \in \{1\}^k \subset \mathcal{I}^k$ for some $k \in \mathbb{N}$, i.e., $\mathbf{j}$ is a tuple of 1's. It follows that, for each $y \in X_\lambda$, there is a unique word $\mathbf{v}_0(y) \in A_\lambda$ satisfying

- $T_{\mathbf{v}_0(y)}(0) = y$; and
- $T_{\mathbf{v}}(0) = y \implies \mathbf{v} = \mathbf{v}_0(y) * \mathbf{j}$, for some $\mathbf{j} \in \{\emptyset\} \cup \{1\}^k$, $k \in \mathbb{N}$.

In particular, $T_{\mathbf{v}}(0) = y$ implies $\alpha_{\mathbf{v}} = \alpha_{\mathbf{v}_0(y)}\alpha_1^k$, for some $k \in \mathbb{N}_0$. In other words, there is exactly one element of $A_\lambda \setminus A_{\lambda/\alpha_1}$, $\mathbf{v}_0(y)$, which gets mapped onto y . Therefore, $\mathbf{v} \mapsto T_{\mathbf{v}}(0)$ gives a bijection $A_\lambda \setminus A_{\lambda/\alpha_1} \rightarrow X_\lambda$. The right-hand side of (1) follows, completing the proof. ■

Now, to continue the proof of the theorem, we combine the previous lemma with the following, which expresses $\mu_\lambda(T_{\mathbf{v}}[0, 1))$ as a ratio involving $|X_\lambda|$.

LEMMA 2: For all $\mathbf{v} \in W(\mathcal{I})$ and $\lambda \in (0, 1]$,

$$(2) \quad |T_{\mathbf{v}}[0, 1) \cap X_{\lambda\alpha_{\mathbf{v}}}| = |X_\lambda|.$$

In particular, for all $\lambda \in (0, \alpha_{\mathbf{v}}]$,

$$(3) \quad \mu_\lambda(T_{\mathbf{v}}[0, 1)) = \frac{|X_{\lambda/\alpha_{\mathbf{v}}}|}{|X_\lambda|}.$$

Moreover, if $\mathbf{v} = \emptyset$, or $\mathbf{v} = \mathbf{w} * i$ with $T_i(0) \neq 0$, these hold for all $\lambda > 0$.

Proof of Lemma 2. Fix $\mathbf{v} \in W(\mathcal{I})$ and consider (2). First let $\lambda \leq 1$. By induction, using the properties of the T_i , if $T_{\mathbf{w}}(0) \in T_{\mathbf{v}}[0, 1)$, then, either $\mathbf{w} = \mathbf{v} * \mathbf{j}$ or $\mathbf{v} = \mathbf{w} * \mathbf{j}$, for some $\mathbf{j} \in W(\mathcal{I})$. Moreover, if $\mathbf{w} \in A_{\lambda\alpha_{\mathbf{v}}}$, the second option gives a contradiction: $\alpha_{\mathbf{w}} > \alpha_{\mathbf{v}} \geq \lambda\alpha_{\mathbf{v}} \geq \alpha_{\mathbf{w}}$. Thus, we have

$$T_{\mathbf{v}}[0, 1) \cap X_{\lambda\alpha_{\mathbf{v}}} = \{T_{\mathbf{v}*\mathbf{j}}(x) : \mathbf{j} \in A_\lambda\}.$$

By injectivity of $T_{\mathbf{v}}$, the right-hand side bijectively corresponds to X_λ , giving (2). Applying (2), with $\lambda/\alpha_{\mathbf{v}}$ in place of λ and dividing by $|X_\lambda|$, yields (3).

As for the final remark, now let $\lambda > 1$. This implies the second option holds, $\mathbf{v} = \mathbf{w} * \mathbf{j}$, since the first is satisfied only if $\lambda \leq 1$, by a similar contradiction argument. It follows that $T_{\mathbf{j}}(0) = 0$ and thus, as in the proof of Lemma 1, $\mathbf{j} \in \{1\}^k$ for some $k \in \mathbb{N}$. This contradicts the assumptions of the remark; thus both the left- and the right-hand side of (2) are empty when $\lambda > 1$. The lemma follows. ■

The significance of relating $|X_\lambda|$ to $|A_\lambda|$ will now become clear from the following renewal equation.

LEMMA 3: *The following holds for all $\lambda > 0$:*

$$(4) \quad |A_\lambda| = \sum_{i \in \mathcal{I}} |A_{\lambda/\alpha_i}| + \chi_{\{\lambda \leq 1\}},$$

where χ is the indicator function. Equivalently, the following renewal equation applies, for $Z(t) := e^{-t}|A_{e^{-t}}|$:

$$Z(t) = \sum_{i \in \mathcal{I}} \alpha_i Z(t - \log(\alpha_i^{-1})) + e^{-t} \chi_{\{t \geq 0\}}, \quad \forall t \in \mathbb{R}.$$

Proof of Lemma 3. The second equation is a restatement of the first, (4). One can obtain (4) by partitioning the non-empty words in A_λ according to their first symbol. That is, one can write the following disjoint union:

$$A_\lambda \setminus \{\emptyset\} = \bigsqcup_{i \in \mathcal{I}} \underbrace{\{i * \mathbf{v} \in W(\mathcal{I}) \mid \alpha_{\mathbf{v}} \leq \lambda/\alpha_i\}}_{\text{in bijection with } A_{\lambda/\alpha_i}}.$$

This gives rise to the sum in (4). Accounting for $\emptyset$, $\emptyset \in A_\lambda$ if and only if $\lambda \leq 1$, thus providing the indicator term in (4), and completing the proof. ■

To make use of this renewal equation, just as in [2, 8, 19] it is necessary to consider two cases which behave somewhat differently. These cases correspond to, for example, the α -Kakutani schemes for $\alpha = 1/3$ and $\alpha = 1/2$, as described in the introduction.

Definition (Rank): For $n \in \mathbb{N}$ we will say the collection $\{\alpha_i\}_{i \in \mathcal{I}}$ is **rank** n if the smallest additive subgroup of $\mathbb{R}$ containing the set $\{-\log(\alpha_i)\}_{i \in \mathcal{I}}$ is isomorphic to $\mathbb{Z}^n$. If $\{\alpha_i\}_i$ is not rank n for any $n \in \mathbb{N}$ we will say $\{\alpha_i\}_{i \in \mathcal{I}}$ is **infinite rank**. Also, whenever $\{\alpha_i\}_i$ is not rank one, we say it is **higher rank**.

Example 8: The following examples illustrate the different ranks:

- (1) $\{1/2^n\}_{n \in \mathbb{N}}$ is rank one;
- (2) $\{1/2\} \cup \{1/3^n\}_{n \in \mathbb{N}}$ is rank two;
- (3) $\{1/2\} \cup \{1/3\} \cup \{1/7^n\}_{n \in \mathbb{N}}$ is rank three; and
- (4) $\{1/n^s\}_{n=2}^\infty$ is infinite rank, where $s \approx 1.728$ satisfies $\zeta(s) = 2$ (here ζ denotes the Riemann zeta function).

3.1. UNIFORM DISTRIBUTION IN THE RANK ONE CASE. In this subsection, we concentrate on the rank one case, also known as the arithmetic, rationally-related or commensurable case (see [2, 8, 19]). The characteristic feature of this case is that the contraction ratios are all powers of some number x :

$$\{\alpha_i\}_{i \in \mathcal{I}} \subset \{x^n\}_{n \in \mathbb{N}}.$$

Thereby, fixing the minimal such $x > 0$ and writing

$$N_k = |\{i \in \mathcal{I} \mid \alpha_i = x^k\}|,$$

the renewal equation for the sequence $u_n := x^n |A_{x^n}|$ reads $u_0 = 1$ and for all $n \in \mathbb{N}$,

$$u_n = \sum_{k=1}^n N_k x^k u_{n-k} + x^n.$$

To this sequence, we apply the following renewal theorem, paraphrased from [13, Theorem 1, p. 330].

LEMMA 4 (Feller renewal theorem): *Suppose that $(\lambda_k)_{k=1}^\infty$ and $(\nu_k)_{k=1}^\infty$ satisfy $\lambda_k, \nu_k \geq 0$, $\sum_{k=1}^\infty \lambda_k = 1$, $\sum_{k=1}^\infty \nu_k < \infty$, and that the smallest subgroup of $\mathbb{Z}$ containing $\{k \in \mathbb{Z} \mid \lambda_k > 0\}$ is $\mathbb{Z}$ itself. Then, for the sequence $(u_n)_{n=0}^\infty$ defined by $u_0 = 1$ and*

$$u_n = \sum_{k=1}^n \lambda_k u_{n-k} + \nu_n,$$

we have the limit

$$(u_n)_{n=0}^\infty \rightarrow \frac{\sum_{n=1}^\infty \nu_n}{\sum_{k=1}^\infty k \lambda_k}.$$

(Convention: if the sum on the denominator is infinite, the limit is zero.)

This applies to give the following corollary.

COROLLARY: Suppose that $\{\alpha_i\}_{i \in \mathcal{I}}$ is rank one, that x is the minimal positive number for which $\{\alpha_i\}_{i \in \mathcal{I}} \subset \{x^n\}_{n \in \mathbb{N}}$, and that $H := -\sum_i \alpha_i \log(\alpha_i) < \infty$. Then

$$(5) \quad x^n |A_{x^n}| \rightarrow \frac{-\log(x)}{H(1-x)}$$

as $n \rightarrow \infty$. Consequently, for all $\mathbf{v} \in W(\mathcal{I})$,

$$(6) \quad \mu_{x^n}(T_{\mathbf{v}}[0, 1)) \rightarrow \alpha_{\mathbf{v}} = \|T_{\mathbf{v}}[0, 1)\|$$

as $n \rightarrow \infty$.

Proof of Corollary. We first deduce (5) from Lemma 4. To check some of the assumptions: Firstly, $\{k : N_k > 0\} \subset n\mathbb{Z}$ implies that every α_i is a power of x^n , contradicting minimality of x if $n \geq 2$. Moreover, $\sum_k k \lambda_k$ equals

$$\sum_{k=1}^{\infty} k N_k x^k = \sum_{i \in \mathcal{I}} \alpha_i \log_x(\alpha_i) = \frac{H}{-\log(x)}$$

and is finite by assumption, as is $\sum_n \nu_n = \sum_{n=1}^{\infty} x^n = (1-x)^{-1}$. Finally,

$$\sum_k N_k x^k = \sum_{i \in \mathcal{I}} \alpha_i = 1,$$

and (5) follows.

To show (6) follows from (5), fix $\mathbf{v} \in W(\mathcal{I})$ and let $\alpha_{\mathbf{v}} = x^m$. Also set $C = -\log(x)/H$ here for convenience. We have the following two cases, which complete the proof:

CASE 1: No T_i fixes 0, for any $i \in \mathcal{I}$.

Then, applying Lemma 1 followed by Lemma 2 gives, for all $n \in \mathbb{N}$,

$$\mu_{x^n}(T_{\mathbf{v}}[0, 1)) = \frac{|X_{x^{n-m}}|}{|X_{x^n}|} = \frac{|A_{x^{n-m}}|}{|A_{x^n}|} \sim \frac{Cx^{m-n}}{Cx^{-n}} = x^m = \alpha_{\mathbf{v}}$$

as $n \rightarrow \infty$.

CASE 2: $T_1(0) = 0$ for some $1 \in \mathcal{I}$.

Then, for $\alpha_1 = x^j$, the same lemmas give, for all $n \geq m$,

$$\begin{aligned} \mu_{x^n}(T_{\mathbf{v}}[0, 1)) &= \frac{|X_{x^{n-m}}|}{|X_{x^n}|} = \frac{|A_{x^{n-m}}| - |A_{x^{n-m-j}}|}{|A_{x^n}| - |A_{x^{n-j}}|} \\ &= \frac{x^n |A_{x^{n-m}}| - x^n |A_{x^{n-m-j}}|}{x^n |A_{x^n}| - x^n |A_{x^{n-j}}|} \rightarrow \frac{Cx^m - Cx^{m+j}}{C - Cx^j} = x^m = \alpha_{\mathbf{v}} \end{aligned}$$

as $n \rightarrow \infty$. ■

In summary, we have shown, under the assumptions of the theorem in the rank one case, that one has the required equidistribution on intervals of the form $T_{\mathbf{v}}[0, 1)$. It is then a simple matter to conclude equidistribution (in this case) on an arbitrary interval by packing it with subintervals of this form.

The following conclusion to the proof of Theorem 1 will be written in the continuous limit, $\lambda \rightarrow 0^+$, so as to align it with the higher rank case. In this case, since we naturally have

$$\{\alpha_{\mathbf{v}} \mid \mathbf{v} \in W(\mathcal{I})\} \subset \{x^n \mid x \in \mathbb{N}\},$$

$\lambda \mapsto |A_\lambda|$ is necessarily constant on intervals of the form $[x^{n+1}, x^n)$, and hence the discrete limits of the previous lemma ((5) and (6)) extend to continuous limits, e.g.,

$$\mu_\lambda(T_{\mathbf{v}}[0, 1)) \rightarrow \alpha_{\mathbf{v}}$$

as $\lambda \rightarrow 0^+$, so it makes no difference to this case.

Proof of Theorem 1. Let $I \subset [0, 1]$ be an interval, and let $\alpha_{\max} = \max_{i \in \mathcal{I}}(\alpha_i)$. Fixing $n \in \mathbb{N}$, let

$$\mathcal{U}_n := \{U = T_{\mathbf{v}}[0, 1) \mid \mathbf{v} \in \mathcal{I}^n, U \subset I\}.$$

Since the $\{T_i[0, 1)\}_{i \in \mathcal{I}}$ are pairwise disjoint, it follows inductively that the intervals in U are pairwise disjoint. Therefore,

$$\sum_{U \in \mathcal{U}_n} \mu_\lambda(U) \leq \mu_\lambda(I),$$

with a similar equality for Lebesgue measure. We now show that

$$(7) \quad \sum_{\mathbf{v} \in \mathcal{U}_n} \|T_{\mathbf{v}}[0, 1)\| \geq \|I\| - 2\alpha_{\max}^n.$$

Let $x \in I \setminus \bigcup_{U \in \mathcal{U}_n} U$. Then, one of the following two cases holds:

CASE 1: $x \in K_n := [0, 1] \setminus \bigcup_{\mathbf{v} \in \mathcal{I}^n} T_{\mathbf{v}}[0, 1)$.

A simple induction shows that $\text{Leb}(K_n) = 0$ for all n .

CASE 2: $x \in T_{\mathbf{v}}[0, 1)$ for some $\mathbf{v} \in \mathcal{I}^n$.

Then, since $T_{\mathbf{v}}[0, 1) \not\subset I$, it must meet an endpoint of I . Hence, there are at most two $\mathbf{v} \in \mathcal{I}^n$ with this property, and x is contained in the union of at most two intervals, each with length at most $\alpha_{\max}^n$. This proves (7).

Proceeding with the proof, since $\mu_\lambda(U) \rightarrow \|U\|$ for each $U \in \mathcal{U}_n$, we have

$$\liminf_{\lambda \rightarrow 0^+} \mu_\lambda(I) \geq \liminf_{\lambda \rightarrow 0^+} \sum_{U \in \mathcal{U}} \mu_\lambda(U) = \sum_{U \in \mathcal{U}} \|U\| \geq \|I\| - 2\alpha_{\max}^n.$$

Repeating this argument with an analogous collection of intervals contained wholly in $[0, 1] \setminus I$ gives the converse inequality,

$$\limsup_{\lambda \rightarrow 0^+} \mu_\lambda(I) \leq \|I\| + 2\alpha_{\max}^n,$$

and the proof is completed by taking $n \rightarrow \infty$. $\blacksquare$

3.2. UNIFORM DISTRIBUTION IN THE HIGHER RANK CASE. In the remaining, generic case, the proof is very similar to the above. It continues with the following lemma, a convenient application of the Blackwell renewal theorem (see [4] and [18]).

LEMMA 5 (Blackwell renewal theorem): *Suppose that $\{\alpha_i\}_{i \in \mathcal{I}}$ is not rank one and that $H = -\sum_i \alpha_i \log(\alpha_i) < \infty$. Then one has the continuous limit*

$$Z(t) \rightarrow \frac{1}{H} \quad \text{as } t \rightarrow \infty,$$

where $Z(\cdot)$ is defined in Lemma 3.

This gives the following corollary. The proof is similar to that of the preceding, with the convergence now taking place as $\lambda \rightarrow 0^+$.

COROLLARY: *For $\{\alpha_i\}_{i \in \mathcal{I}}$ as in the previous lemma, $|A_\lambda| \sim 1/\lambda H$ as $\lambda \rightarrow 0^+$, and consequently, all $\mathbf{v} \in W(\mathcal{I})$,*

$$\mu_\lambda(T_{\mathbf{v}}[0, 1]) \rightarrow \alpha_{\mathbf{v}} = \|T_{\mathbf{v}}[0, 1]\|.$$

The conclusion of Theorem 1 in this case is the same as given in the preceding case, see §3.1.

Remark: The same method of proof gives a more general result. Suppose we have a set $\mathbb{X}$ equipped with a probability measure ν , and there is a collection $\{T_i, \alpha_i\}_{i \in \mathcal{I}}$ such that

- (a) $T_i : \mathbb{X} \rightarrow \mathbb{X}$ are injective ν -measurable functions;
- (b) for all $i, j \in \mathcal{I}$ distinct, $T_i(\mathbb{X}) \cap T_j(\mathbb{X}) = \emptyset$;
- (c) for all $\mathbf{v} \in W(\mathcal{I})$, $\nu(T_{\mathbf{v}}(\mathbb{X})) = \alpha_{\mathbf{v}}$;
- (d) $\alpha_i > 0$ for all $i \in \mathcal{I}$ and $\sum_{i \in \mathcal{I}} \alpha_i = 1$; and
- (e) $-\sum_{i \in \mathcal{I}} \alpha_i \log(\alpha_i) < \infty$.

Then, for any $x \in \mathbb{X} \setminus \bigcup_{i \in \mathcal{I}} T_i(\mathbb{X})$, and for any ν -measurable set S which can be written as

$$S = \bigcup_{v \in V} T_v(\mathbb{X})$$

(where V is any subset of $w(\mathcal{I})$), we have, as $\lambda \rightarrow 0^+$,

$$\frac{|S \cap X_\lambda(x)|}{|X_\lambda(x)|} \rightarrow \nu(S),$$

where $X_\lambda(x) = \{T_v(x) : v \in W(\mathcal{I}), \alpha_v \geq \lambda\}$.

It is also possible to generalise the method of proof to include substitution schemes starting from arbitrary partitions, and one recovers results analogous to those of [2].

4. Discrepancy estimates

In [20], the author both generalised the method used by Adler and Flatto in [1] to general finite partitions, and posed questions which inspired various other papers. Of particular interest is the behaviour of the **discrepancy**. This corresponds to estimating the speed of convergence (where the I are intervals) of

$$\sup_{I \subset [0,1]} |\mu_\lambda(I) - \|I\|| \rightarrow 0 \quad \text{as } \lambda \rightarrow 0^+.$$

The Koksma inequality uses such estimates to give convergence rates for integrals of bounded-variation functions. Hence, fast-decaying discrepancies provide potential computationally-efficient numerical integration techniques. The theory of discrepancies (of equidistributing sets and sequences) has naturally received a lot of attention (see, e.g., [9] for an overview). There are also interesting open problems, for example on optimal discrepancy in higher dimensions.

Regarding interval substitution schemes, in the finite-partition case, discrepancy estimates are provided by Drmota–Infusino in [8], extending the application-focused work of Carbone in [5]. We in turn extend this to the context of infinite partitions. The results obtained here are different, depending on whether we are in the rank one case or the higher rank case.

4.1. DISCREPANCY ESTIMATES IN THE RANK ONE CASE. In this final section, we extend the analysis of the rank one case to estimate the discrepancy between the measure μ_λ and the Lebesgue measure. More precisely, we have the following result.

THEOREM 2: Suppose that

- (1) $\{\alpha_i\}_{i \in \mathcal{I}}$ is rank one;
- (2) $x > 0$ is the smallest number for which $\{\alpha_i\}_{i \in \mathcal{I}} \subset \{x^n\}_{n \in \mathbb{N}}$; and
- (3) there is some $\varepsilon > 0$ for which $\sum_{i \in \mathcal{I}} \alpha_i^{1-\varepsilon} < \infty$.

Then there is an $R^* \in (0, 1)$, made explicit in Lemma 6 below, such that, for all $\rho \in (x/R^*, 1)$, there is a constant $C > 0$ such that, for all $n \in \mathbb{N}$ and all intervals $I \subset [0, 1]$,

$$|\mu_{x^n}(I) - \|I\|| \leq C\rho^n.$$

The proof of Theorem 2 begins with the following simple lemma, defining R^* using a generating function for $|A_{x^n}|$.

LEMMA 6: Given $\{\alpha_i\}_{i \in \mathcal{I}}$ as in Theorem 2, the function formally defined by

$$g(z) = (z - x) \sum_{n=0}^{\infty} |A_{x^n}| z^n$$

has a holomorphic extension to the open disk of radius R^* about 0, where

$$R^* := \max \left(R \in (x, \min\{1, x^{1-\varepsilon}\}) \mid |z| = R \implies \sum_{j \in \mathcal{I}} z^{n_j} \neq 1 \right) > x$$

and where $n_i := \log_x(\alpha_i)$. Therefore, denoting by b_n the n th Taylor coefficient of g , given $R < R^*$, there exists $C > 0$ such that, for all $n \in \mathbb{N}$, $b_n \leq CR^{-n}$.

Proof of Lemma 6. From (4), the renewal equation of Lemma 3, one has, for $|z| \leq x^{1-\varepsilon}$ and $z \neq x$,

$$\begin{aligned} \frac{g(z)}{z-x} &= \sum_{n=0}^{\infty} |A_{x^n}| z^n = \sum_{n=0}^{\infty} \left(\sum_{j \in \mathcal{I}} |A_{x^{n-n_j}}| + 1 \right) z^n \\ &= \sum_{n=0}^{\infty} \sum_{j \in \mathcal{I}} |A_{x^{n-n_j}}| z^n + \frac{1}{1-z} \\ &= \sum_{j \in \mathcal{I}} z^{n_j} \sum_{n=0}^{\infty} |A_{x^{n-n_j}}| z^{n-n_j} + \frac{1}{1-z} \\ &= \sum_{j \in \mathcal{I}} z^{n_j} \frac{g(z)}{z-x} + \frac{1}{1-z}, \end{aligned}$$

which rearranges to

$$g(z) = \frac{z - x}{(z - 1)(\sum_{j \in \mathcal{I}} z^{n_j} - 1)}.$$

Therefore, g has a meromorphic expansion on the disk of convergence of

$$(8) \quad z \mapsto \sum_{j \in \mathcal{I}} z^{n_j}$$

which has radius at least $x^{1-\varepsilon}$, by assumption on the decay of $\{\alpha_i\}$.

An elementary argument (see [11, pp. 201–2]) shows that $z = x$ is not only a simple root of (8) with residue $1/H$ (see the statement of Lemmas 4 or 5), but it is also the only root of (8) in the closed disk $\{|z| \leq x\}$. Therefore, g is holomorphic on the open disk of radius $R^* > x$, where R^* is the absolute value of the next smallest root of (8), or equal to $x^{1-\varepsilon}$ if no other root exists. ■

Example 9: In certain nice cases, one can say more. For the simplest infinite example, $\alpha_n = 2^{-n}$ ($n \in \mathbb{N}$), $g(z) \equiv 1/2$ is constant.

The final stage of the proof of Theorem 2 is similar to that of Theorem 1, but the method of decomposing a general interval requires a little more care. We use the following concept of itinerary.

Definition (Itinerary): Given $b \in [0, 1)$, we say $\mathbf{v} = \text{It}_n(b)$ if $\mathbf{v} \in \mathcal{I}^n$ and $b \in T_{\mathbf{v}}[0, 1)$ (this is unique by the disjointness of $T_i[0, 1)$), and we let $\text{It}_0(b) = \emptyset$.

If $\text{It}_n(b)$ exists for all $n \in \mathbb{N}$, we say b has **infinite itinerary**, in which case (by disjointness again), there exists $(i_k)_{k=1}^\infty \in \mathcal{I}^\mathbb{N}$ such that

$$(9) \quad \text{It}_k(b) = (i_1, i_2, \dots, i_k)$$

for all $k \in \mathbb{N}$. Otherwise, we say b has **finite itinerary**, and for

$$n = \max\{k : \text{It}_k(b) \text{ exists}\}$$

and $\text{It}_n(b) = (i_1, \dots, i_n)$, (9) holds for $0 < k \leq n$.

Finally, for simplicity in later sums, we adopt the convention that $\alpha_{\text{It}_k(b)} = 0$ whenever $\text{It}_k(b)$ does not exist. In particular, recalling $\alpha_{\max} = \max_i(\alpha_i)$, we will use that $\alpha_{\text{It}_k(b)} \leq \alpha_{\max}^n$ for all $n \in \mathbb{N}$.

We are now ready to prove Theorem 2.

Proof of Theorem 2. Let $\{\alpha_i\}$, x and $\rho \in (x/R^*, 1)$ be as in the statement of the theorem. Fix $b \in (0, 1)$ and consider $I = [0, b)$ (so that $\|I\| = b$).

We consider a natural “greedy algorithm” for filling I with intervals of the form $T_v[0, 1)$. First, we fill I with as many intervals as possible from the “first generation”,

$$\{T_i[0, 1) \mid i \in \mathcal{I}\},$$

then, in the remaining gap, as many from the “second generation”,

$$\{T_{i_1}T_{i_2}[0, 1) \mid (i_1, i_2) \in \mathcal{I}^2\},$$

and so on. Accordingly, let $V_1 := \{i \in \mathcal{I} \mid T_i[0, 1) \subset I\}$ and for $k \geq 2$,

$$V_k := \{\mathbf{w} * i \in \mathcal{I}^k \mid (\mathbf{w}, i) \in \mathcal{I}^{k-1} \times \mathcal{I}, \underbrace{T_{\mathbf{w}}T_i[0, 1) \subset I}_{\textcircled{1}}, \underbrace{T_{\mathbf{w}}[0, 1) \not\subset I}_{\textcircled{2}}\}.$$

Notice that conditions $\textcircled{1}$ and $\textcircled{2}$ imply that $T_{\mathbf{w}}[0, 1)$ meets both I and $[0, 1] \setminus I$. It thus contains b , so $\mathbf{w} = \text{It}_{k-1}(b)$.

In this way, for $k \in \mathbb{N}$, $V_k = \emptyset$ if $\text{It}_{k-1}(b)$ does not exist, and otherwise,

$$V_k \subset \{\text{It}_{k-1}(b)\} \times \mathcal{I}.$$

In any case, for $V := \bigcup_{k \in \mathbb{N}} V_k$, the union $\mathcal{V} = \bigcup_{v \in V} T_v[0, 1)$ is disjoint and wholly contained in I , and hence has Lebesgue measure

$$\sum_{v \in V} \alpha_v \leq b.$$

To show the converse equality, we note that, for each $n \in \mathbb{N}$, $\mathcal{V}$ contains the union $\bigcup_{U \in \mathcal{U}_n} U$ from the proof of Theorem 1, which we recall has Lebesgue measure at least $b - 2\alpha_{\max}^n$, for $\alpha_{\max} = \max_i(\alpha_i)$. Hence

$$\sum_{v \in V} \alpha_v = b.$$

Unfortunately, the respective equality for μ_λ does not necessarily apply (specifically, in the case that no $i \in \mathcal{I}$ fixes 0, i.e., $0 \notin \bigcup_i T_i[0, 1)$), but nevertheless the difference decays in a favourable manner. Let

$$K := \{T_v(0) \in I \setminus \mathcal{V} \mid v \in W(\mathcal{I})\},$$

then

$$\mu_\lambda(I) = \mu_\lambda(\mathcal{V}) + \mu_\lambda(K)$$

for all $\lambda > 0$. Suppose that $\mathbf{v} \in \mathcal{I}^n$ has $T_{\mathbf{v}}(0) \in I \setminus \mathcal{V}$. Then $T_{\mathbf{v}}[0, 1) \not\subset I$ which, by the above reasoning, implies that $b \in T_{\mathbf{v}}[0, 1)$, i.e., $\mathbf{v} = \text{It}_n(b)$. Hence, by Lemma 2,

$$\begin{aligned}
 \mu_{x^n}(K) &\leq \mu_{x^n}\{T_{\text{It}_k(b)}(0) \mid k \in \mathbb{N}_0, \text{It}_k(b) \text{ exists}\} \\
 (10) \quad &= \frac{1}{|A_{x^n}|} |\{k \in \mathbb{N}_0 \mid \alpha_{\text{It}_k(b)} \geq x^n\}| \leq \frac{1}{|A_{x^n}|} |\{k \in \mathbb{N}_0 \mid \alpha_{\max}^k \geq x^n\}| \\
 &= \frac{1}{|A_{x^n}|} \left(\left\lfloor \frac{n}{\log_x(\alpha_{\max})} \right\rfloor + 1 \right),
 \end{aligned}$$

which is plainly $\mathcal{O}(nx^n)$ as $n \rightarrow \infty$.

Moving forwards, we write (by disjointness)

$$\mu_{x^n}(I) - b = \mu_{x^n}(K) + \mu_{x^n}(\mathcal{V}) - \sum_{\mathbf{v} \in V} \alpha_{\mathbf{v}} = \mu_{x^n}(K) + \underbrace{\sum_{\mathbf{v} \in V} \mu_{x^n}(T_{\mathbf{v}}[0, 1))}_{\textcircled{\Lambda}} - \alpha_{\mathbf{v}}.$$

We consider the decay of $\textcircled{\Lambda}$ in two cases.

CASE 1: No $i \in \mathcal{I}$ fixes 0.

In this case, by Lemma 1, $\mu_{x^n}(T_{\mathbf{v}}[0, 1)) = 0$ whenever $\alpha_{\mathbf{v}} > x^n$, and therefore

$$\textcircled{\Lambda} = \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} \geq x^n}} \frac{|X_{x^n/\alpha_{\mathbf{v}}} - \alpha_{\mathbf{v}}|X_{x^n}|}{|X_{x^n}|} - \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} < x^n}} \alpha_{\mathbf{v}}.$$

The second sum is simple to bound, using the fact that $V_k \subset \{\text{It}_{k-1}(b)\} \times \mathcal{I}$ if $\text{It}_{k-1}(b)$ exists, and is empty otherwise. This gives the inequality below, noting $\alpha_{\text{It}_n(b)} \leq \alpha_{\max}^n$:

$$\begin{aligned}
 \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} < x^n}} \alpha_{\mathbf{v}} &\leq x^{\varepsilon n} \sum_{\mathbf{v} \in V} \alpha_{\mathbf{v}}^{1-\varepsilon} = x^{\varepsilon n} \sum_{k=1}^{\infty} \sum_{\mathbf{v} \in V_k} \alpha_{\mathbf{v}}^{1-\varepsilon} \\
 &\leq x^{\varepsilon n} \sum_{k=1}^{\infty} \sum_{i \in \mathcal{I}} \alpha_{\text{It}_{k-1}(b)}^{1-\varepsilon} \alpha_i^{1-\varepsilon} \\
 &\leq x^{\varepsilon n} \sum_{k=1}^{\infty} \sum_{i \in \mathcal{I}} \alpha_{\max}^{(1-\varepsilon)(k-1)} \alpha_i^{1-\varepsilon} \\
 &= \frac{x^{n\varepsilon}}{1 - \alpha_{\max}^{1-\varepsilon}} \sum_{i \in \mathcal{I}} \alpha_i^{1-\varepsilon},
 \end{aligned}$$

which is a bounded multiple of $x^{n\varepsilon} \leq (x/R^*)^n < \rho^n$.

To bound the remaining first sum, we consider the generating function, for $m \in \mathbb{N}$, of $|A_{x^{n-m}}| - x^m |A_{x^n}|$, which relates to g from Lemma 6 in the following manner:

$$\begin{aligned} \sum_{n=0}^{\infty} (|A_{x^{n-m}}| - x^m |A_{x^n}|) z^n &= (z^m - x^m) \sum_{n=0}^{\infty} |A_{x^n}| z^n \\ &\leq \left(\frac{z^m - x^m}{z - x} \right) g(z) \\ &= (z^{m-1} + xz^{m-2} + \cdots + x^{m-1}) g(z). \end{aligned}$$

Writing $g(z) = \sum_{n \in \mathbb{Z}} b_n z^n$ (setting $b_n = 0$ for $n < 0$), equating coefficients gives, for all $n \in \mathbb{N}$,

$$|A_{x^{n-m}}| - x^m |A_{x^n}| = b_{n-m+1} + x b_{n-m+2} + \cdots + x^{m-1} b_n.$$

By Lemma 6, since $x/\rho < R^*$, there exists $C_1 > 0$ such that $|b_n| < C_1(\rho/x)^n$ for all $n \in \mathbb{N}_0$, which trivially applies for negative n also. Thus, we have

$$\begin{aligned} ||A_{x^{n-m}}| - x^m |A_{x^n}|| &\leq C_1 \frac{\rho^{n-m-1} + \rho^{n-m-2} + \cdots + \rho^n}{x^{n-m+1}} \\ &= C_1 \left(\frac{\rho}{x} \right)^n \frac{(x/\rho)^m - x^m}{x(1-\rho)} \\ &\leq C_2 \left(\frac{\rho}{x} \right)^n \left(\left(\frac{x}{\rho} \right)^m - x^m \right), \end{aligned}$$

where $C_1 = C_2 x(1-\rho)$. In particular, if $x^m = \alpha_v$, then

$$(x/\rho)^m \leq (R^*)^m < x^{m(1-\varepsilon)} \leq \alpha_v^{1-\varepsilon},$$

and, since $x^n |A^n|$ is bounded away from zero (since it is positive and convergent to a positive limit), the previous inequality gives, for $x^n \leq \alpha_v$,

$$\left| \frac{|A_{x^n/v}| - \alpha_v |A_{x^n}|}{|A_{x^n}|} \right| \leq \frac{C_2}{x^n |A_{x^n}|} \rho^n (\alpha_v^{1-\varepsilon} - \alpha_v) \leq C_3 \rho^n (\alpha_v^{1-\varepsilon} - \alpha_v),$$

for some constant $C_3 > 0$. Summing over v thus gives

$$\sum_{\substack{v \in V \\ \alpha_v \geq x^n}} \left| \frac{|A_{x^n/\alpha_v}| - \alpha_v |A_{x^n}|}{|A_{x^n}|} \right| \leq C_3 \rho^n \sum_{v \in V} (\alpha_v^{1-\varepsilon} - \alpha_v).$$

The right-hand side is a bounded multiple of ρ^n , completing the proof in this case.

CASE 2: There exists $1 \in \mathcal{I}$ such that $T_1(0) = 0$.

The proof here is similar to the previous case, but with some necessary modifications. Firstly, since it is now possible that $\mu_\lambda(T_{\mathbf{v}}[0, 1)) > 0$ if $\lambda > \alpha_{\mathbf{v}}$, we also need to bound

$$\sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} < \lambda}} \mu_\lambda(T_{\mathbf{v}}[0, 1)).$$

As in the proof of Lemma 2, if $\mu_\lambda(T_{\mathbf{v}}[0, 1)) > 0$ for $\lambda > \alpha_{\mathbf{v}}$, there exists $\mathbf{w} \in W(\mathcal{I})$ and $j \in \{1\}^k$ such that

$$T_{\mathbf{v}}[0, 1) = T_{\mathbf{w}}T_j[0, 1)$$

and

$$\mu_\lambda(T_{\mathbf{v}}[0, 1)) = \mu_\lambda\{T_{\mathbf{v}}(0)\} = \mu_\lambda\{T_{\mathbf{w}}(0)\}.$$

Moreover, by construction, $\mathbf{v} \in V$ again implies that $T_{\mathbf{w}}[0, 1) \not\subset I$ and hence that $\mathbf{w} = \text{It}_n(b)$ for some $n \in \mathbb{N}_0$ (i.e., $T_{\mathbf{v}}(0) = T_{\text{It}_n(b)}(0)$). Therefore, similarly to K considered above, one has

$$\sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} < \lambda}} \mu_{x^n}(T_{\mathbf{v}}[0, 1)) \leq \mu_{x^n}\{T_{\text{It}_n(b)}(0) \mid n \in \mathbb{N}\} = \mathcal{O}(nx^n)$$

for all $y > x$, as $n \rightarrow \infty$. Therefore, up to a quantity of this order, we may write $\mu_{x^n}(I) - b$ again as

$$\textcircled{\Lambda} = \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} \geq x^n}} \frac{|X_{x^n/\alpha_{\mathbf{v}}} - \alpha_{\mathbf{v}}|X_{x^n}|}{|X_{x^n}|} - \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} > x^n}} \alpha_{\mathbf{v}}.$$

The second sum is $\mathcal{O}(x^{n\varepsilon})$, by the argument in the first case. The first sum yields to a similar treatment as before. For $\alpha_{\mathbf{v}} = x^m$ and $\alpha_1 = x^j$, the numerator reads

$$|X_{x^{n-m}}| - x^m|X_{x^n}| = |A_{x^{n-m}}| - x^m|A_{x^n}| - |A_{x^{n-m-j}}| - x^m|A_{x^{n-j}}|,$$

so that

$$\sum_{k=0}^m |X_{x^{n-m}}| - x^m|X_{x^n}| = (z^m - x^m - z^{m+j} + x^m z^j) \frac{g(z)}{z - x}.$$

Equating coefficients as before and applying the inequalities at the end of the last case (in particular that $x/\rho \leq x^{1-\varepsilon}$) then gives

$$\begin{aligned} ||X_{x^{n-m}}| - x^m|X_{x^n}|| &\leq |(b_{n-m+1} + b_{n-m+2}x + \cdots + x^{m-1}b_n) \\ &\quad + (b_{n-m-j+1} + b_{n-m-j+2}x + \cdots + x^{m-1}b_{n-j})| \\ &\leq C_2 \left(\frac{\rho}{x}\right)^n \left(\left(\frac{x}{\rho}\right)^m - x^m + \left(\frac{x}{\rho}\right)^{m-j} - x^{m-j}\right) \\ &\leq C_2 \left(\frac{\rho}{x}\right)^n \left(\alpha_v^{1-\varepsilon} - \alpha_v + \left(\frac{\alpha_v}{\alpha_1}\right)^{1-\varepsilon} - \frac{\alpha_v}{\alpha_1}\right). \end{aligned}$$

Finally, as in the last case, dividing through by

$$|X_{x^n}| \sim (1 - \alpha_1^{-1})Hx^{-n}/\log(x),$$

and summing over $v \in V$ gives a bounded multiple of ρ^n , completing the proof in this case. ■

4.2. DISCREPANCY ESTIMATES IN THE HIGHER RANK CASE. When the collection $\{\alpha_j\}_{j \in \mathcal{I}}$ is not rank one, we require not only a strict decay property on the $\{\alpha_j\}$ —as in Theorem 2—, but also a kind of Diophantine condition. To proceed, we need the following definition.

Definition (R -badly approximable): For $R \in [2, \infty)$ we say a number $\gamma \in \mathbb{R}$ is R -badly approximable if there exists a $d > 0$ such that

$$\forall (l, k) \in \mathbb{Z}^2 \text{ s.t. } l \neq 0, \quad \left| \gamma - \frac{k}{l} \right| > \frac{d}{|l|^R}.$$

Remark: Larger values of R correspond to more easily approximable numbers:

- For any $R = 2$, the property is equivalent to γ having bounded continued fraction coefficients. (Such γ comprise a set of measure 0 containing all quadratic algebraic numbers.)
- For $R > 2$, the property holds Lebesgue almost-everywhere: by Jarník's theorem [12, Thm. 10.3], the Hausdorff dimension of the complementary set is $2/R < 1$.
- Certain transcendental numbers (e.g., Liouville numbers) do not satisfy the property for any R whatsoever.

THEOREM 3: *Suppose that $\{\alpha_i\}_{i \in \mathcal{I}}$ is not rank one, that there is some $\varepsilon > 0$ such that $\sum_i \alpha_i^{1-\varepsilon} < \infty$, and that there is a pair $\alpha_j, \alpha_k \in \{\alpha_i\}_{i \in \mathcal{I}}$ such that $\log(\alpha_j)/\log(\alpha_k)$ is $(2+r)$ -badly approximable, for some $r \in [0, 1/2)$. Then, for all $P \in (0, P^*)$, there exists a constant C such that, for all intervals $I \subset [0, 1]$,*

$$|\mu_\lambda(I) - \|I\|| \leq C(-\log(\lambda))^{-P},$$

where

$$P^* = \frac{1-2r}{8(1+r)}.$$

The proof of Theorem 3 requires us to consider the Mellin transform,

$$g(z) = \int_0^\infty t^{-z-1} |A_{1/t}| \, dt,$$

which has the following explicit form, courtesy of the renewal equation for $|A_\lambda|$. Let $\Re, \Im$ denote the real and imaginary parts of a complex number, respectively.

LEMMA 7: *For $\Re(z) > 1$, the Mellin transform $g(z)$ takes the form*

$$g(z) = \frac{1}{z(\sum_{j \in \mathcal{I}} \alpha_j^z - 1)}.$$

In particular, if $\sum_j \alpha_j^{1-\varepsilon} < \infty$, g has a meromorphic extension to the half-plane $\{\Re(z) > 1 - \varepsilon\}$.

Proof of Lemma 7. The formula follows from standard properties of the Mellin transform, namely $g(\alpha_j z) = \alpha_j^z g(z)$, and the renewal equation of Lemma 3. $\blacksquare$

As is well-known, one can obtain asymptotic information about a function from the distribution of poles of its Mellin transform.

In particular, this depends on the zeros of the almost-periodic function

$$f(z) = \sum_{j \in \mathcal{I}} \alpha_j^z - 1.$$

There is a lot we can say straight away. From our assumptions on $\{\alpha_j\}_j$, we have that $z = 1$ is a simple zero of f , and by the triangle inequality, there are no zeros of f for $\{\Re(z) > 1\}$. Moreover, if there were another zero of f on the line $\{\Re(z) = 1\}$, it would follow that $\{\alpha_j\}_j$ is rank one, a contradiction.

The role of the poles of the Mellin transform is illustrated in the proof of the following result.

PROPOSITION 1: For any given collection of positive numbers $\{\alpha_j\}_{j \in \mathcal{I}}$ which sum to 1 such that $H = -\sum_j \alpha_j \log(\alpha_j) < \infty$, there is no $\varepsilon > 0$ for which

$$(11) \quad |A_\lambda| = \frac{1}{H\lambda} + \mathcal{O}(\lambda^{1-\varepsilon})$$

as $\lambda \rightarrow 0^+$.

Proof of Proposition 1. Fix a collection $\{\alpha_j\}_{j \in \mathcal{I}}$ as above and assume for a contradiction that there is an $\varepsilon > 0$ for which (11) holds. Taking the Mellin transform of this equation yields that

$$\frac{1}{zf(z)} = \frac{1}{H(z-1)} + \int_1^\infty t^{-z-\varepsilon} \mathcal{O}(1) dt.$$

Since this second integral converges absolutely for all z with $\Re(z) > 1 - \varepsilon$, we see that the left-hand side has a meromorphic extension to the half plane $\{z \in \mathbb{C} : \Re(z) > 1 - \varepsilon\}$, with only one pole at $z = 1$. To obtain a contradiction, we provide a sequence of zeros $z_n = u_n + iv_n$ with $u_n \rightarrow 1$, $v_n \rightarrow \infty$, using the theory of almost-periodic functions. In particular, we follow the proof of the corollary to [7, Theorem 3.6].

It can quickly be seen that f is almost-periodic, by [7, Corollary to Thm. 3.12], and is bounded on $\{\Re(z) \geq 1 - \varepsilon/2\}$ since, in this half-plane,

$$|f(z)| \leq 1 + \sum_{i \in \mathcal{I}} \alpha_i^{1-\varepsilon/2}.$$

Also, since $f(1) = 0$, the definition of almost-periodicity provides a sequence of positive numbers $(y_n)_{n=1}^\infty$ for which $f(1 + iy_n) \rightarrow 0$ as $n \rightarrow \infty$. In other words, the holomorphic functions

$$f_n(z) = f(z + iy_n)$$

are bounded on the same half-plane, and $f_n(1) \rightarrow 0$.

Furthermore, fixing any index $j \in \mathcal{I}$, for $L := -2\pi/\log(\alpha_j)$, we see that

$$\sup_{|v| \leq L} |f_n(1 + iv)|$$

is bounded away from zero uniformly, since every interval of length L contains a v such that

$$\alpha_j^{1+iv} = -\alpha_j \quad \therefore \quad |f(1 + iv)| \geq 1 + \alpha_j - \sum_{k \in \mathcal{I} \setminus \{j\}} \alpha_k = 2\alpha_j.$$

Therefore, on the rectangle $(1-\varepsilon, 1+\varepsilon)+i(-l, l)$, an application of Montel's theorem shows that f_n uniformly converges (passing to a subsequence if necessary) to some analytic function f_∞ . From the last two considerations, $f_\infty(1) = 0$ and f_∞ is not identically zero.

Now, taking a circle about 1 small enough that f_∞ has no zeros on it, by Hurwitz's theorem, for all n sufficiently large, f_n has a zero $\hat{z}_n$ inside this circle, such that $\hat{z}_n \rightarrow 1$ as $n \rightarrow \infty$.

This thus provides us with a sequence of zeros of f accumulating on the line as required, contradicting the statement that f has only one zero in the above-mentioned half-plane. ■

Considering now the proof of Theorem 3, the following lemma uses the $(2+r)$ -badly approximable hypothesis, following [8]. For simplicity in the following two proofs, we write $\alpha = \max(\alpha_j, \alpha_k)$ and $\beta = \min(\alpha_j, \alpha_k)$.

LEMMA 8: *Suppose that $(\alpha_j)_{j \in \mathcal{I}}$ satisfies the conditions of Theorem 3. Then there exists $C > 0$ such that, whenever $z = 1 - u + iv \in \mathbb{C} \setminus \{1\}$ satisfies both $f(z) = 0$ and $u < \varepsilon$, then $u > 0$ and*

$$|v|^{2+2r} \geq \frac{C}{u}.$$

Proof of Lemma 8. The fact that $u > 0$ follows from the discussion preceding the previous proposition. We first show, if $f(z) = 0$, the argument of α^z in $(-\pi, \pi]$ is $\mathcal{O}(\sqrt{u})$ as $u \rightarrow 0^+$. That is, the quantity $\eta_\alpha \in (-\pi, \pi]$, satisfying

$$e^{i\eta_\alpha} = \frac{\alpha^z}{|\alpha^z|} = \frac{\alpha^z}{\alpha^{1-u}},$$

is $\mathcal{O}(\sqrt{u})$. This uses the triangle inequality and a small amount of trigonometry, as we now detail. We have that

$$|\alpha^z - 1| \leq 1 - \alpha + H(u),$$

where

$$H(u) := \sum_{n \in \mathcal{I}} \alpha_n^{1-u} - \alpha^{1-u} - 1 + \alpha.$$

In particular, $H(u) = \mathcal{O}(u)$ as $u \rightarrow 0^+$, by the mean value theorem. Therefore, for u sufficiently small, $H(u) < \alpha$, which gives rise to the picture in Figure 4(i).

Consequently, $|\eta_\alpha| < \theta$, where θ is as in Figure 4(ii) and satisfies the following equation:

$$\cos(\theta) = \frac{1 + \alpha^{2-2u} - (1 - \alpha + H(u))^2}{2\alpha^{1-u}} = 1 - \mathcal{O}(u) \quad (u \rightarrow 0),$$

where the constant now depends on the value of α . Using, for example, that

$$\lim_{y \rightarrow 0^+} \frac{\arccos(1-y)}{\sqrt{y}} = \sqrt{2}, \quad \arccos : [-1, 1] \rightarrow [0, \pi],$$

it is clear that $\theta = \mathcal{O}(\sqrt{u})$ as $u \rightarrow 0^+$, hence the same applies to η_α .

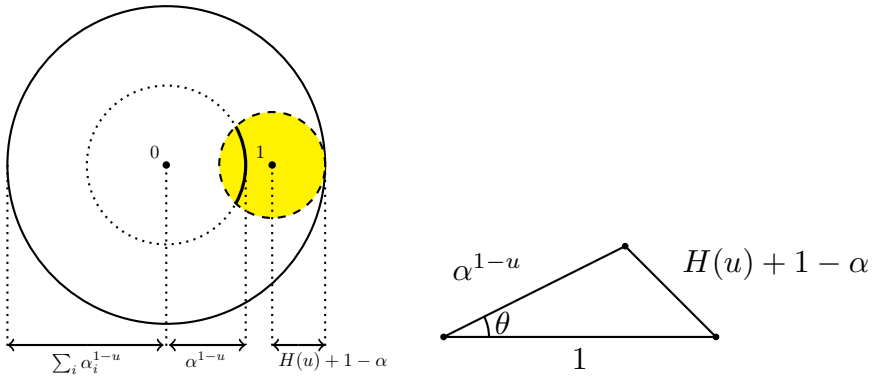


Figure 4. (i) The region in which α^z must lie for $f(z) = 0$ —the bold arc in the shaded circle; (ii) the triangle defining $\theta(u)$, the maximum possible value of $|\eta_\alpha|$.

We can repeat this argument with β in place of α to bound the analogously defined η_β —i.e., $\eta_\beta = \mathcal{O}(\sqrt{u})$.

Write $v \log(\alpha) = 2\pi k + \eta_\alpha$ and $v \log(\beta) = 2\pi l + \eta_\beta$, supposing $|v| \geq 2\pi / \log(\beta)$ so that k and l are non-zero. Substituting into the definition of $(2+r)$ -badly approximable gives

$$\frac{d}{|l|^{2+r}} \leq \left| \frac{\log(\alpha)}{\log(\beta)} - \frac{k}{l} \right| = \left| \frac{2\pi k + \eta_\alpha}{2\pi l + \eta_\beta} - \frac{k}{l} \right| = \left| \frac{\eta_\alpha}{2\pi l + \eta_\beta} - \eta_\beta \frac{2\pi k + \eta_\alpha}{(2\pi l + \lambda)^2} \right|,$$

where $\lambda \in \mathbb{R}$ is a constant, provided by the mean value theorem, satisfying

$$0 < |\lambda| < |\eta_\beta| \leq \pi.$$

Using the triangle inequality on the right-hand side, multiplying through by $(2\pi l + \lambda)^2 |l|^r$ and using that $|\eta_\alpha| \leq \pi$, one obtains

$$\begin{aligned} \pi^2 d &\leq 4\pi^2 \left(1 - \frac{\lambda}{l}\right)^2 d \leq 2\pi \left(|\eta_\alpha| \frac{(1 - \lambda/2\pi l)^2}{|1 + \eta_\beta/2\pi l|} + |\eta_\beta| \left| \frac{k}{l} - \frac{\eta_\alpha}{2\pi} \right| \right) |l|^{1+r} \\ &\leq 2\pi \left(9|\eta_\alpha| + |\eta_\beta| \left(\left| \frac{k}{l} \right| + \frac{1}{2} \right) \right) |l|^{1+r} \\ &\leq 2\pi \left(9|\eta_\alpha| + |\eta_\beta| \left(\frac{3}{2} + 2 \frac{\log(\alpha)}{\log(\beta)} \right) \right) |l|^{1+r}, \end{aligned}$$

where the last inequality uses that $|v| > -2\pi/\log(\beta)$. We can divide through by the large bracket on the right-hand side and, recalling the $\sqrt{u}$ asymptotic for η_α and η_β , obtain the required inequality for some constant C , for u sufficiently small and $|v|$ sufficiently large.

That the inequality holds in the whole of the specified region (with a possibly different C) follows simply from the fact that zeros of f can only accumulate on the vertical boundary $\{z = 1 - \varepsilon + iv : v \neq 0\}$, and so there is an open neighbourhood of $[1 - \varepsilon, 1]$ containing only one zero of f , at 1. Since there are finitely many zeros of f to cater for (at most), we can adapt C accordingly. ■

The next lemma is a variant on the last and allows us to estimate decay of the Mellin inverse integral inside the zero-free region.

LEMMA 9: Suppose $(\alpha_j)_{j \in \mathcal{I}}$ is a collection of positive numbers as given in Theorem 3. Then there exists $C > 0$ such that, whenever $z = 1 - u + iv \in \mathbb{C}$ with $u \geq 0$ and $\sigma > 0$ sufficiently small,

$$(12) \quad |f(z)| < \sigma$$

implies one of the following holds: either

$$|v| \leq 2\pi/\log(\beta)$$

or

$$|v|^{2+2r} > \frac{C}{\max(u, \sigma)}.$$

Proof of Lemma 9. The proof is an adaptation of that for Lemma 8. This time, the radius of the circle depicted in Figure 4 is $H(u) + 1 - \alpha + \sigma$ and correspondingly,

$$\cos(\theta) = \frac{1}{2} \frac{1 + \alpha^{2-2u} - (1 - \alpha + H(u) + \sigma)^2}{\alpha^{1-u}} = 1 - \mathcal{O}(\max(u, \sigma))$$

as $\max(u, \sigma) \rightarrow 0$, which gives, for $\max(u, \sigma)$ is sufficiently small,

$$|\theta| \leq \frac{\pi}{2} \sqrt{1 - \cos(\theta)} = \mathcal{O}(\sqrt{\max(u, \sigma)})$$

and the proof continues as before. $\blacksquare$

The next crucial lemma is the analogue to Lemma 6 in the higher rank case.

LEMMA 10: *Under the assumptions of Theorem 3,*

$$|A_\lambda| = \frac{1}{H\lambda} + \mathcal{O}(\lambda^{-1}(-\log(\lambda))^{-P}), \quad \lambda \in (0, 1],$$

where $P \in (0, P^*)$ is as given in Theorem 3.

Proof of Lemma 10. The proof uses simple complex analysis to estimate the integral

$$(13) \quad F(t) := \frac{1}{2\pi i} \int_{2-i\infty}^{2+i\infty} \frac{t^{z+3}}{z(z+1)(z+2)(z+3)f(z)} dz, \quad \text{for } t > 1.$$

We first relate $F(t)$ to $|A_{1/t}|$. On the line $\{\Re(z) = 2\}$, we see that f is uniformly bounded away from zero,

$$|f(z)| \geq 1 - \sum_{j \in \mathcal{I}} \alpha_j^2 > 0,$$

so $F(t)$ absolutely converges, for all t . Therefore, by the Mellin inversion theorem, the Mellin transform of $t \mapsto F(t)/t^3$ is the denominator of the integrand of F :

$$F^*(z) := \int_0^\infty t^{-z-1} F(t) dt = \frac{1}{z(z+1)(z+2)(z+3)f(z)}.$$

Therefore, using integration by parts, one has $F^{(3)}(t) = |A_{1/t}|$ Lebesgue almost-everywhere.

We now relate the integral in (13) to that over the contour Γ , parameterised by

$$\gamma : \mathbb{R} \rightarrow \mathbb{C}, \quad \gamma(v) = 1 + iv - D \min(1, |v|^{-2-r})$$

(see Figure 5), where $D > 0$ is chosen sufficiently small so that the previous two lemmas apply as follows: firstly, the only zero of f which lies on or to the right of Γ is at 1, and secondly, whenever z lies on or to the right of Γ and $|\Im(z)| \geq -2\pi(\log(\beta))^{-1}$, one has

$$(14) \quad |f(z)| \geq D|\Im(z)|^{-2-2r}.$$

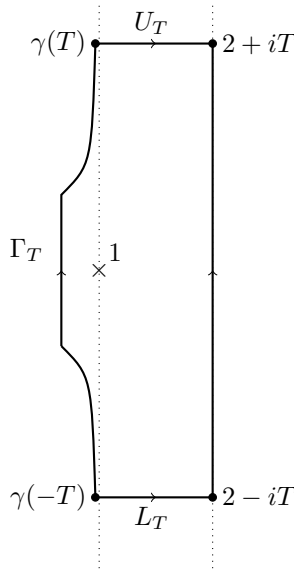


Figure 5. The contour Γ_t used in the proof of Lemma 10.

Consider, for $T \geq 1$, the contour Γ_T depicted in Figure 5. Cauchy's residue theorem gives

$$\int_{2-iT}^{2+iT} F^*(z) dz = 2\pi i \frac{t^4}{24H} + \int_{\Gamma_T} F^*(z) dz + \int_{U_T} F^*(z) dz + \int_{L_T} F^*(z) dz.$$

Since U_T and L_T have bounded length, a simple application of (14) shows that the corresponding last two integrals are $\mathcal{O}(|T|^{-2+2r})$ as $T \rightarrow \infty$; therefore, taking the limit, one has

$$(15) \quad F(t) = \frac{t^4}{24H} + \frac{1}{2\pi i} \int_{\Gamma} F^*(z) t^{-z-3} dz.$$

From this point, the proof follows along the lines of [10, Theorem 4.6, pp. 133–134]. Since $t \mapsto |A_{1/t}|$ is non-decreasing on the real line, so too are the functions F , F' and F'' . Using this, together with repeated applications of the mean value theorem, gives that, for any $t \in \mathbb{R}$ and $h > 0$, both

$$\frac{F(t-3h) - 3F(t-2h) + 3F(t-h) - F(t)}{-h^3} \leq |A_{1/t}|$$

and

$$|A_{1/t}| \leq \frac{F(t+3h) - 3F(t+2h) + 3F(t+h) - F(t)}{h^3}$$

hold. Substituting (15) into these expressions yields

$$\begin{aligned} & \frac{F(t \pm 3h) - 3F(t \pm 2h) + 3F(t \pm h) - F(t)}{\pm h^3} \\ &= t \pm \frac{3}{2}h + \frac{1}{2\pi i} \int_{\Gamma} \frac{(t \pm 3h)^{z+3} - 3(t \pm 2h)^{z+3} + 3(t \pm h)^{z+3} - t^{z+3}}{\pm h^3} F^*(z) dz. \end{aligned}$$

Thus, we have the estimate

$$\begin{aligned} |A_{1/t}| &= t + \mathcal{O}(h) \\ &+ \mathcal{O}\left(\int_{\Gamma} \left| \frac{(t \pm 3h)^{z+3} - 3(t \pm 2h)^{z+3} + 3(t \pm h)^{z+3} - t^{z+3}}{\pm h^3(z+1)(z+2)(z+3)} \right| \left| \frac{1}{zf(z)} \right| |dz| \right). \end{aligned}$$

From now on, let $h = h(t) \in (0, t)$ be a function of t to be determined later. To begin to estimate the integral, consider $|\Delta_{\pm}(t, h, z)|$ for $t > 1$ and $z \in \Gamma$, where

$$\Delta_{\pm}(t, h, z) := \frac{(t \pm 3h)^z - 3(t \pm 2h)^z + 3(t \pm h)^z - t^z}{\pm h^3(z+1)(z+2)(z+3)}.$$

We estimate $|\Delta_{\pm}|$ in two different ways. For the first, we express $|\Delta_{\pm}|$ as a series of nested integrals:

$$\begin{aligned} |\Delta_{\pm}(t, h, z)| &= \left| \frac{1}{\pm h} \int_t^{t \pm h} \frac{1}{\pm h} \int_{\hat{t}}^{\hat{t} \pm h} \frac{1}{\pm h} \int_{\hat{\hat{t}}}^{\hat{\hat{t}} \pm h} \hat{\hat{t}}^z d\hat{\hat{t}} d\hat{t} dt \right| \leq (t+3h)^{\Re(z)} \\ &\leq 4t^{\Re(z)}. \end{aligned}$$

For the first inequality, we have taken the absolute value signs inside the integral and applied the mean value theorem three times, noting that $t \mapsto |t^z| = t^{\Re(z)}$ is increasing on $[0, \infty)$, since $\Re(z) > 0$ for $z \in \Gamma$. (The second follows simply from $h < t$ and $\Re(z) < 1$.)

The triangle inequality gives us another estimate: for $t \geq 1$,

$$\begin{aligned} |\Delta_{\pm}(t, h, z)| &\leq \frac{(4t)^{\Re(z)+3} + 3(3t)^{\Re(z)+3} + 3(2t)^{\Re(z)} + t^{\Re(z)+3}}{h^3|z+1||z+2||z+3|} \\ &\leq \frac{548t^{\Re(z)+3}h^{-3}}{|z+1||z+2||z+3|}. \end{aligned}$$

From the Bernoulli inequality it follows that, for $z = \gamma(v) \in \Gamma$, the three quantities $|\gamma(v)+1|$, $|\gamma(v)+2|$, $|\gamma(v)+3|$ are all greater than or equal to $(1+|v|)/2$,

so altogether we have the following:

$$|\Delta_{\pm}(t, h, \gamma(v))| \leq \min \left(4t^{\Re(\gamma(v))}, \frac{4384t^{\Re(\gamma(v))+3}}{h^3(1+|v|)^3} \right).$$

From (14), one can easily deduce that $|\gamma(v)f(\gamma(v))|^{-1} = \mathcal{O}((1+|v|)^{1+2r})$, for all $v \in \mathbb{R}$. Combining the previous three inequalities and using the boundedness of $|\gamma'(v)|$ gives the following:

$$|A_{1/t}| = \frac{t}{H} + \mathcal{O}(h) + \mathcal{O} \left(\int_{-\infty}^{\infty} (1+|v|)^{1+2r} \min \left(t^{\Re(\gamma(v))}, \frac{t^{\Re(\gamma(v))+3}}{(1+|v|)^3} \right) dv \right).$$

Because the integral is symmetric in v , it suffices to estimate the integral from 0 to ∞ , as we now do. Writing $\Re(\gamma(v)) = 1 - \kappa(v)$, where

$$\kappa(v) = D \min(1, |v|^{-2-2r}),$$

the previous equation simplifies to the following:

$$\begin{aligned} \frac{|A_{1/t}|}{t} - \frac{1}{H} &= \mathcal{O}\left(\frac{h}{t}\right) + \mathcal{O} \left(\int_0^{\infty} (1+v)^{1+2r} t^{-\kappa(v)} \min \left(1, \frac{(t/h)^3}{(1+v)^3} \right) dv \right) \\ &= \mathcal{O}\left(\frac{h}{t}\right) + \mathcal{O} \left(\int_0^{\infty} (1+v)^{-2+2r} t^{-\kappa(v)} \min \left((1+v)^3, \left(\frac{t}{h}v^3\right) \right) dv \right) \\ &= \mathcal{O}\left(\frac{h}{t}\right) + \mathcal{O} \left(\int_1^{\infty} v^{-2+2r} t^{-\kappa(v-1)} \min \left(v^3, \left(\frac{t}{h}\right)^3 \right) dv \right). \end{aligned}$$

Now let $\delta \in (0, 1-2r)$. For $v, t \geq 1$, we have, since κ is decreasing on $[0, \infty)$,

$$\begin{aligned} v^{-2+2r} t^{-\kappa(v-1)} &\leq v^{-2+2r} t^{-\kappa(v)} \\ &= v^{-2+2r} \exp(-\kappa(v) \log(t)) \\ &= v^{-2+2r+\delta} \exp(-\kappa(v) \log(t) - \delta \log(v)) \\ &= v^{-2+2r+\delta} \exp(-Dv^{-2-2r} \log(t) - \delta \log(v)) \\ &= v^{-2+2r+\delta} \exp(-\xi_{\delta}(t)), \end{aligned}$$

where

$$\begin{aligned} \xi_{\delta}(t) &:= \inf_{v \geq 1} (Dv^{-2-2r} \log(t) + \delta \log(v)) \\ &= \frac{\delta}{2+2r} \left(1 + \log \left(\frac{D(2+2r) \log(t)}{\delta} \right) \right). \end{aligned}$$

This last equality holds for all t sufficiently large, by elementary calculus. Hence

$$\frac{|A_{1/t}|}{t} - \frac{1}{H} = \mathcal{O}\left(\frac{h}{t}\right) + \mathcal{O} \left(e^{-\xi_{\delta}(t)} \int_1^{\infty} v^{-2+2r+\delta} \min \left(v^3, \left(\frac{t}{h}\right)^3 \right) dv \right).$$

Now, writing $\omega = vh/t$ and substituting, the integral becomes

$$\left(\frac{t}{h}\right)^{2+2r+\delta} \int_{h/t}^{\infty} \omega^{-2+2r+\delta} \min(\omega^3, 1) d\omega,$$

and this can be split into two parts,

$$\int_{h/t}^1 \omega^{1+2r+\delta} d\omega + \int_1^{\infty} \omega^{-2+2r+\delta} d\omega,$$

both of which are finite, since $2r + \delta \in (0, 1)$. Therefore, for all $t \geq 1$,

$$\frac{|A_{1/t}|}{t} - \frac{1}{H} = \mathcal{O}\left(\frac{h}{t}\right) + \mathcal{O}\left(e^{-\xi_\delta(t)} \left(\frac{t}{h}\right)^{2+2r+\delta}\right).$$

Finally, choosing $h(t) = t \exp(\frac{-\xi_\delta(t)}{3+2r+\delta})$ ensures both terms have the same order of magnitude, and the previous equation simplifies to the required expression:

$$|A_{1/t}| = \frac{t}{H} + \mathcal{O}(t \log(t)^{-P}),$$

where

$$P = P(r, \delta) = \frac{\delta}{(2+2r)(3+2r+\delta)} \in \left(0, \frac{1-2r}{8(1+r)}\right). \quad \blacksquare$$

The concluding stages of the proof of Theorem 3 are similar to those of Theorem 2, but with some notable differences.

Proof of Theorem 3. For simplicity of writing, we consider only the case that none of the T_i fix 0. Let $\lambda \in (0, 1)$ and recall $I = [b, 1)$, V_k and V from the proof of Theorem 2.

Similarly to that proof, up to a quantity of the order $\mathcal{O}(-\lambda \log(\lambda))$, we may write the discrepancy in terms of three sums,

$$\begin{aligned} \mu_\lambda(I) - \|I\| &= \sum_{\mathbf{v} \in V} \mu_\lambda(T_{\mathbf{v}}[0, 1)) - \alpha_{\mathbf{v}} \\ &= \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} \geq \lambda/\alpha_{\max}}} \frac{|X_{\lambda/\alpha_{\mathbf{v}}}|}{|X_\lambda|} - \alpha_{\mathbf{v}} \\ &\quad + \sum_{\substack{\mathbf{v} \in V \\ \lambda \leq \alpha_{\mathbf{v}} < \lambda/\alpha_{\max}}} \frac{|X_{\lambda/\alpha_{\mathbf{v}}}|}{|X_\lambda|} - \sum_{\substack{\mathbf{v} \in V \\ \alpha_{\mathbf{v}} < \lambda/\alpha_{\max}}} \alpha_{\mathbf{v}}, \end{aligned} \tag{16}$$

where

$$\alpha_{\max} = \max_{i \in \mathcal{I}}(\alpha_i).$$

The second and third sum of (16) decay much faster than the first, which can be deduced from the summability of $\alpha_v^{1-\varepsilon}$ alone: For the third sum, the argument from Theorem 2 applies to bound it by a multiple of $\lambda^{1-\varepsilon}$. The second sum of (16) corresponds to newly split intervals: if $\alpha_v < \lambda/\alpha_{\max}$, then $\alpha_v \alpha_i < \lambda$ for any $i \in \mathcal{I}$, so no proper subinterval of $T_v[0, 1)$ has been split for this value of λ , and $|X_{\lambda/\alpha_v}| = 1$. Therefore, assuming b has an infinite itinerary,

$$\begin{aligned}
 \sum_{\substack{v \in V \\ \lambda \leq \alpha_v < \lambda/\alpha_{\max}}} |X_{\lambda/\alpha_v}| &= |\{\mathbf{v} \in V : \lambda \leq \alpha_v < \lambda/\alpha_{\max}\}| \\
 &= \sum_{n \in \mathbb{N}} |\{\mathbf{v} \in V_n : \lambda \leq \alpha_v < \lambda/\alpha_{\max}\}| \\
 &\leq \sum_{n \in \mathbb{N}} |\{i \in \mathcal{I} : \lambda \leq \alpha_{\text{It}_{n-1}(b)} \alpha_i < \lambda/\alpha_{\max}\}| \\
 &= \sum_{n \in \mathbb{N}} |\{i \in \mathcal{I} : \lambda/\alpha_{\text{It}_{n-1}(b)} \leq \alpha_i < \lambda/(\alpha_{\max} \alpha_{\text{It}_{n-1}(b)})\}| \\
 &\leq \sum_{n \in \mathbb{N}} |\{i \in \mathcal{I} : \lambda/\alpha_{\text{It}_{n-1}(b)} \leq \alpha_i < \lambda/\alpha_{\text{It}_n(b)}\}| \\
 &= |\{i \in \mathcal{I} : \lambda \leq \alpha_i\}| \\
 &\leq \sum_{i \in \mathcal{I}} \alpha_i^{1-\varepsilon} \lambda^{\varepsilon-1}.
 \end{aligned}$$

Thus, the second sum of (16) is bounded by a multiple of λ^ε . The finite itinerary case is similar, involving a finite sum in n .

It remains to bound the first sum of (16), using the asymptotics for $|A_{\lambda/\alpha_v}|$ provided by the previous lemma. One finds that there exists C, C' such that, for all $\mathbf{v} \in V$ with $\alpha_v \geq \lambda/\alpha_{\max}$,

$$\left| \frac{|X_{\lambda/\alpha_v}|}{|X_\lambda|} - \alpha_v \right| \leq C \alpha_v ((-\log(\lambda/\alpha_v))^{-P}) \leq C' \alpha_v ((-\log(\lambda))^{-P}),$$

the last inequality following from the fact that $x \mapsto \log(\lambda)/\log(\lambda/x)$ is uniformly bounded on $[\alpha_{\max}, \infty)$. Summing over $\mathbf{v}$ gives the required estimate. ■

OPEN ACCESS. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution and reproduction in any medium, provided the appropriate credit is given to the original authors and the source, and a link is provided to the Creative Commons license, indicating if changes were made (<https://creativecommons.org/licenses/by/4.0/>).

References

- [1] R. L. Adler and L. Flatto, *Uniform distribution of Kakutani's interval splitting procedure*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **38** (1977), 253–259.
- [2] C. Aistleitner and M. Hofer, *Uniform distribution of generalized Kakutani's sequences of partitions*, Annali di Matematica Pura ed Applicata **192** (2013), 529–538.
- [3] C. Aistleitner, M. Hofer and V. Ziegler, *On the uniform distribution modulo 1 of multi-dimensional LS-sequences*, Annali di Matematica Pura ed Applicata **193** (2014), 1329–1344.
- [4] D. Blackwell, *A renewal theorem*, Duke Mathematical Journal **15** (1948), 145–150.
- [5] I. Carbone, *Discrepancy of LS-sequences of partitions and points*, Annali di Matematica Pura ed Applicata **191** (2012), 819–844.
- [6] I. Carbone, M. R. Iacò and A. Volčič, *A dynamical system approach to the Kakutani–Fibonacci sequence*, Ergodic Theory and Dynamical Systems **34** (2014), 1794–1806.
- [7] C. Corduneanu, *Almost Periodic Functions*, Interscience Tracts in Pure and Applied Mathematics, Vol. 22, Interscience, London, 1968.
- [8] M. Drmota and M. Infusino, M. (2012). *On the discrepancy of some generalized Kakutani's sequences of partitions*, Uniform Distribution Theory **7** (2012), 75–104.
- [9] M. Drmota and R. Tichy, *Sequences, Discrepancies and Applications*, Lecture Notes in Mathematics, Vol. 1651, Springer, Berlin, 1997.
- [10] F. Ellison and W. Ellison, *Prime Numbers*, Interscience, London, 1985.
- [11] P. Erdős, W. Feller and H. Pollard, *A property of power series with positive coefficients*, Bulletin of the American Mathematical Society **55** (1949), 201–204.
- [12] K. Falconer, *Fractal Geometry*, John Wiley & Sons, Chichester, 2014.
- [13] W. Feller, *An Introduction to Probability Theory and Its Applications* Vol. I, John Wiley & Sons, London, 1968.
- [14] G. H. Hardy and J. E. Littlewood, *Some problems of diophantine approximation II*, Acta Mathematica **37** (1914), 193–239.
- [15] M. Hofer, M. R. Iacò and R. Tichy, *Ergodic properties of β -adic Halton sequences*, Ergodic Theory and Dynamical Systems **35** (2015), 895–909.
- [16] M. R. Iacò and V. Ziegler, *Discrepancy of generalized LS-sequences*, Uniform Distribution Theory **12** (2017), 37–63.
- [17] S. Kakutani, *A problem of equidistribution on the unit interval $[0, 1]$* , in *Measure Theory (Proc. Conf., Oberwolfach, 1975)*, Lecture Notes in Mathematics, Vol. 541, Springer, Berlin–New York, 1976, pp. 369–375.
- [18] B. Sewell, *Equidistribution of infinite interval substitution schemes, explicit resonances of Anosov toral maps, and the Hausdorff dimension of the Rauzy gasket*, Ph.D. thesis, University of Warwick, 2021,
<https://warwick.ac.uk/fac/sci/maths/people/staff/sewell>.
- [19] Y. Smilansky, *Uniform distribution of Kakutani partitions generated by substitution schemes*, Israel Journal of Mathematics **240** (2020), 667–710
- [20] A. Volčič, *A generalization of Kakutani's splitting procedure*, Annali di Matematica Pura ed Applicata **190** (2011), 45–54.

- [21] C. Weiss, *On the classification of LS-sequences*, Uniform Distribution Theory **13** (2018), 83–92.
- [22] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Mathematische Annalen **77** (1916), 313–352.