UNIVERSITY OF BIRMINGHAM

University of Birmingham
Research at Birmingham

Equidistribution results for self-similar measures

Baker, Simon

[Link to publication on Research at Birmingham portal](#)

Equidistribution Results for Self-Similar Measures

Simon Baker[*]

School of Mathematics, University of Birmingham, Birmingham B15 2TT, UK

[*]*Correspondence to be sent to: e-mail: Email: simonbaker412@gmail.com*

A well-known theorem due to Koksma states that for Lebesgue almost every $x > 1$ the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one. In this paper, we give sufficient conditions for an analogue of this theorem to hold for a self-similar measure. Our approach applies more generally to sequences of the form $(f_n(x))_{n=1}^{\infty}$ where $(f_n)_{n=1}^{\infty}$ is a sequence of sufficiently smooth real-valued functions satisfying some nonlinearity conditions. As a corollary of our main result, we show that if C is equal to the middle 3rd Cantor set and $t \geq 1$, then with respect to the natural measure on $C + t$, for almost every x, the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one.

1 Introduction

A sequence $(x_n)_{n=1}^{\infty}$ of real numbers is said to be uniformly distributed modulo one if for every pair of real numbers u, v with $0 \leq u < v \leq 1$ we have

$$\lim_{N \to \infty} \frac{\#\{1 \leq n \leq N : x_n \bmod 1 \in [u, v]\}}{N} = v - u. \tag{1.1}$$

The study of uniformly distributed sequences has its origins in the pioneering work of Weyl [30] from the early 20th century. From these beginnings, this topic has developed into an important area of mathematics, with many deep connections to ergodic theory, number theory, and probability theory. Generally speaking, it is a challenging problem to determine whether a given sequence of real numbers is uniformly distributed modulo

one. Often, the sequences one considers are of dynamical or number theoretic origins. For an overview of this topic, we refer the reader to [6, 23] and the references therein.

In this paper, we are interested in the distribution of the sequence $(x^n)_{n=1}^{\infty}$ modulo one for $x > 1$. The study of these sequences dates back to the work of Hardy [15] and Pisot [24, 25]. It is a difficult problem to describe the distribution of $(x^n)_{n=1}^{\infty}$ modulo one for specific values of x. It is still unknown whether there exists a transcendental $x > 1$ such that $\lim_{n\to\infty} \inf_{m\in\mathbb{N}} |x^n - m| = 0$. For some further background and recent results on the distribution of the sequence $(x^n)_{n=1}^{\infty}$, we refer the reader to [1–4, 6–8, 12] and the references therein. The generic behaviour of the sequence $(x^n)_{n=1}^{\infty}$ modulo one for $x > 1$ is described by a well-known theorem due to Koksma [22]. This theorem states that for Lebesgue almost every $x > 1$ the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one. We are interested in determining whether analogues of Koksma's theorem hold for more general measures. More specifically, suppose μ is a Borel probability measure supported on $[1,\infty)$ that is defined "independently" from the family of maps $\{f_n(x) = x^n\}_{n=1}^{\infty}$, we are interested in determining whether for μ almost every x the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one. Of course, the important detail here is what exactly it means for a Borel probability measure to be independent from the family of maps $\{f_n(x) = x^n\}_{n=1}^{\infty}$. A natural family of measures to consider here is the self-similar measures generated by iterated function systems (defined in Section 3). For our purposes, an iterated function system will consist of a finite collection of contracting affine maps. Since for any $n \geq 2$ the map $f_n(x) = x^n$ is not affine, one could view the fact that self-similar measures are defined using affine maps as some sort of independence. As such the following conjecture seems plausible.

Conjecture 1.1.   Let μ be a non-atomic self-similar measure with support contained in $[1,\infty)$. Then for μ almost every x, the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one.

In this paper, we do not prove Conjecture 1.1. Our main contribution in this direction is Theorem 2.1, which lends significant weight to its validity. We conclude this introductory section by giving an overview of a number of related results that motivated the present work.

One of the most well-known results from uniform distribution theory states that for any integer $b \geq 2$, for Lebesgue almost every $x \in \mathbb{R}$ the sequence $(b^n x)_{n=1}^{\infty}$ is uniformly distributed modulo one (see [6, 23]). In what follows, we say that x is b-normal if $(b^n x)_{n=1}^{\infty}$ is uniformly distributed modulo one. For an arbitrary Borel probability

measure μ supported on $\mathbb{R}$, which is defined "independently" from the dynamical system $x \to bx$ mod 1, it is natural to wonder whether x is b-normal for μ almost every x. Just as above, the important detail here is what it means for a Borel probability measure to be independent from the dynamical system $x \to bx$ mod 1. The following metaconjecture encapsulates many important results in this direction.

Conjecture 1.2.   Suppose μ is a Borel probability measure that is "independent" from the dynamical system $x \to bx$ mod 1. Then μ almost every x is b-normal.

The 1st instances of this metaconjecture being verified are found in the papers of Cassels [9] and Schmidt [28]. These authors were motivated by a question of Steinhaus as to whether there exists an x that is b-normal for infinitely many b but not all b. They answered this question in the affirmative by proving that with respect to the natural measure on the middle 3rd Cantor set, almost every x is b-normal if b is not a power of three. The underlying independence here comes from the middle 3rd Cantor set being defined by similarities with contraction ratios equal to $1/3$ and b having a prime factor not equal to 3. The current state of the art in this area is the following two theorems due to Hochman and Shmerkin [17] and Dayan *et al.* [10].

Theorem 1.3.   [17, Theorem 1.4] Let $\{\varphi_i(x) = r_i x + t_i\}_{i \in \mathcal{A}}$ be an iterated function system satisfying the open set condition. Suppose $b \geq 2$ is such that $\frac{\log |r_i|}{\log b} \notin \mathbb{Q}$ for some $i \in \mathcal{A}$, then for every fully supported (we say that a self-similar measure is fully supported if the corresponding probability vector $(p_i)_{i \in \mathcal{A}}$ satisfies $p_i > 0$ for all $i \in \mathcal{A}$) non-atomic self-similar measure μ, μ almost every x is b-normal.

Theorem 1.4.   [10, Theorem 4] Let $\{\varphi_i(x) = \frac{x}{b} + t_i\}_{i \in \mathcal{A}}$ be an iterated function system. Suppose $t_i - t_j \notin \mathbb{Q}$ for some $i, j \in \mathcal{A}$, then for every fully supported non-atomic self-similar measure μ, μ almost every x is b-normal.

Some other important contributions in this area include the papers by Kaufman [20] and Queffélec and Ramaré [27], who constructed Borel probability measures supported on subsets of the badly approximable numbers whose Fourier transform converges to zero polynomially fast. Kaufman [21] has also shown that such measures exist for the α-well approximable numbers. The results of Kaufman [20] and Queffélec and Ramaré [27] were recently extended by Jordan and Sahlsten [19] to a more general class of measures. Importantly, if the Fourier transform of a Borel probability measure

converges to zero sufficiently fast (polynomial speed is sufficient), then it can be shown that almost every point with respect to this measure is b-normal for any $b \geq 2$. In fact, by a recent result of Pollington *et al.* [26], if the Fourier transform of a Borel probability measure converges to zero sufficiently fast, then for almost every x, (1.1) holds for the sequence $(b^n x)_{n=1}^{\infty}$ with an explicit error term.

Another related result was recently proved by Simmons and Weiss [29]. They proved that if $X \subset \mathbb{R}$ is a self-similar set satisfying the open set condition, then with respect to the natural measure on X, the orbit under the Gauss map ($x \to 1/x$ mod 1) of almost every x equidistributes with respect to the Gauss measure. Here the important point is that the natural measure on X is defined independently from the dynamics of the Gauss map.

One of the challenges faced when addressing Conjecture 1.1 is that, at least to the best of the author's knowledge, there is no dynamical system that effectively captures the distribution of $(x^n)_{n=1}^{\infty}$ modulo one. As such one cannot rely upon techniques from ergodic theory to prove this conjecture. Techniques from ergodic theory were previously applied with great success in the proofs of Theorems 1.3 and 1.4. Instead of using these techniques, our approach will exploit the fact that the maps $f_n(x) = x^n$ are not affine for $n \geq 2$ and the fact that self-similar measures are defined using affine maps.

2   Statement of Results

Our main contribution in the direction of Conjecture 1.1 is the following theorem.

Theorem 2.1.   Let $\{\varphi_i(x) = rx + t_i\}_{i \in \mathcal{A}}$ be an equicontractive iterated function system satisfying the convex strong separation condition with self-similar set X contained in $[1, \infty)$. Moreover, let $(p_i)_{i \in \mathcal{A}}$ be a probability vector satisfying

$$\frac{1}{2} < \frac{-\sum_{i \in \mathcal{A}} p_i \log p_i}{-\log |r|}.$$

Then with respect to the self-similar measure μ corresponding to $(p_i)_{i \in \mathcal{A}}$, for μ almost every x the sequence $(x^n)_{n=1}^{\infty}$ is uniformly distributed modulo one.

We define what we mean by iterated function system, self-similar set, and what it means for an iterated function system to be equicontractive and to satisfy the convex strong separation condition in Section 3. Importantly, both of these conditions are satisfied by the iterated function system $\{\phi_1(x) = \frac{x+2t}{3}, \phi_2(x) = \frac{x+2+2t}{3}\}$ for any $t \in \mathbb{R}$. The

self-similar set for this iterated function system is $C+t$ where C is the middle 3rd Cantor set. Using the fact that the restriction of the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on $C+t$ coincides with the self-similar measure corresponding to the probability vector $(p_i)_{i=1}^2 = (1/2, 1/2)$, we see that Theorem 2.1 immediately implies the following corollary. (The restriction of the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on $C+t$ is given by $\mu(A) = \mathcal{H}^{\frac{\log 2}{\log 3}}(A \cap (C+t))$. Here $\mathcal{H}^{\frac{\log 2}{\log 3}}$ is the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure. For more on Hausdorff measure, see [13]. The restriction of the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on $C+t$ can be thought of as the natural measure on $C+t$.)

Corollary 2.2. Let C be the middle 3rd Cantor set. Then for any $t \geq 1$, with respect to the restriction of the $\frac{\log 2}{\log 3}$-dimensional Hausdorff measure on $C+t$, for almost every x the sequence $(x^n)_{n=1}^\infty$ is uniformly distributed modulo one.

Theorem 2.1 is implied by the following more general theorem, which applies to a general class of functions.

Theorem 2.3. Let $\{\varphi_i(x) = rx + t_i\}_{i \in \mathcal{A}}$ be an equicontractive iterated function system satisfying the convex strong separation condition with self-similar set X contained in $[1, \infty)$. Let $(f_n)_{n=1}^\infty$ be a sequence of functions satisfying the following properties:

A. $f_n \in C^3(\text{conv}(X), \mathbb{R})$ for each n. (Here $\text{conv}(X)$ denotes the convex hull of X and $C^3(\text{conv}(X), \mathbb{R})$ denotes the set of three times differentiable functions from $\text{conv}(X)$ to $\mathbb{R}$.)

B. There exists $C_1, C_2 > 0$ such that for any m, n with $m < n$ we have the following:

$$|f_n'(x) - f_m'(x)| \leq C_1 n^{C_2} x^{n-1}$$

for all $x \in \text{conv}(X)$.

C. There exists $C_3 > 0$ such that for all n sufficiently large, for any $m < n$, we have the following:

$$|f_n''(x) - f_m''(x)| \geq C_3 x^{n-2}$$

for all $x \in \text{conv}(X)$.

D. For any m, n with $m < n$ we have either

$$f_n'''(x) - f_m'''(x) \geq 0$$

for all $x \in \text{conv}(X)$ or

$$f_n'''(x) - f_m'''(x) \le 0$$

for all $x \in \text{conv}(X)$.

Moreover, let $(p_i)_{i \in \mathcal{A}}$ be a probability vector satisfying

$$\frac{1}{2} < \frac{-\sum_{i \in \mathcal{A}} p_i \log p_i}{-\log |r|}.$$

Then with respect to the self-similar measure μ corresponding to $(p_i)_{i \in \mathcal{A}}$, for μ almost every x the sequence $(f_n(x))_{n=1}^{\infty}$ is uniformly distributed modulo one.

Remark 2.4.   To see how Theorem 2.1 follows from Theorem 2.3, let $f_n(x) = x^n$ for all $n \ge 1$. Then for any m, n with $m < n$ and $x \ge 1$ we have

$$|f_n'(x) - f_m'(x)| = nx^{n-1} - mx^{m-1} \le 2nx^{n-1}$$

and

$$f_n'''(x) - f_m'''(x) = n(n-1)(n-2)x^{n-3} - m(m-1)(m-2)x^{m-3} \ge 0.$$

Moreover, if n also satisfies $n \ge 2$, then

$$|f_n''(x) - f_m''(x)| = n(n-1)x^{n-2} - m(m-1)x^{m-1} \ge (n(n-1) - m(m-1))x^{n-2} \ge x^{n-2}.$$

Therefore, properties B, C, and D of Theorem 2.3 are satisfied by the sequence of functions $(f_n(x) = x^n)_{n=1}^{\infty}$. Property A of Theorem 2.3 is obviously satisfied by this sequence of functions. Therefore, Theorem 2.1 follows from Theorem 2.3.

Remark 2.5.   Note that we have deliberately phrased Theorem 2.3 with its application in the proof of Theorem 2.1 in mind. Theorem 2.3 still holds if the inequalities in property B and property C are replaced with the perhaps more natural inequalities:

$$|f_n'(x) - f_m'(x)| \le C_1 n^{C_2} x^n$$

and

$$|f_n''(x) - f_m''(x)| \geq C_3 x^n.$$

These inequalities can be shown to be equivalent to those stated in property B and property C by altering the constants C_1 and C_3 appropriately. In particular, because conv(X) is a compact subset of $[1, \infty)$, the extra powers of x can be reconciled by altering the leading constant term.

Remark 2.6.   The hypotheses of Theorem 2.3 are satisfied by many sequences of functions. For instance, we could take $f_n(x) = x^n + x^{n-1} + \cdots + x + 1$ for all n. Alternatively, we could fix a polynomial g with strictly positive coefficients and let $f_n(x) = g(x)x^n$ for all n, or $f_n(x) = g(n)x^n$ for all n. Each of these sequences of functions satisfies the hypotheses of Theorem 2.3.

We can build further examples by taking a sequence of functions $(f_n)_{n=1}^{\infty}$, which satisfies the hypotheses of Theorem 2.3, and a sequence of functions $(g_n)_{n=1}^{\infty}$ whose 1st and 2nd derivatives grow subexponentially in n and which also satisfies property D with the same sign as $(f_n)_{n=1}^{\infty}$. The sequence $(f_n + g_n)_{n=1}^{\infty}$ would then satisfy the hypotheses of Theorem 2.3. To be more precise, we could take $(f_n)_{n=1}^{\infty}$ to be any sequence of functions satisfying the hypotheses of Theorem 2.3 where property D is satisfied with positive sign; we then define a new sequence of functions $(h_n(x) = f_n(x) + n \log x)_{n=1}^{\infty}$. The sequence $(h_n)_{n=1}^{\infty}$ then satisfies the hypotheses of Theorem 2.3 if conv$(X) \subset (1, \infty)$.

The rest of this paper is organised as follows. In Section 3, we recall the necessary preliminaries from fractal geometry and the theory of uniform distribution. In Section 4, we prove Theorem 2.3.

3 Preliminaries

3.1 Fractal geometry

We call a map $\varphi : \mathbb{R} \to \mathbb{R}$ a similarity if it is of the form $\varphi(x) = rx + t$ for some $r \in (-1, 0) \cup (0, 1)$ and $t \in \mathbb{R}$. We call a finite set of similarities $\{\varphi_i\}_{i \in \mathcal{A}}$ an iterated function systems or IFS for short. Here and throughout, $\mathcal{A}$ denotes an arbitrary finite set. Given an IFS $\{\varphi_i(x) = r_i x + t_i\}_{i \in \mathcal{A}}$, we say that it is equicontractive if there exists $r \in (-1, 0) \cup (0, 1)$ such that $r_i = r$ for all $i \in \mathcal{A}$. Throughout this paper, we will assume that if $\{\varphi_i\}_{i \in \mathcal{A}}$ is an equicontractive IFS then $r \in (0, 1)$. For each of our theorems, there is no loss of generality in making this assumption. This is because if $\{\varphi_i\}_{i \in \mathcal{A}}$ is an equicontractive IFS satisfying

the convex strong separation condition, then $\{\varphi_i \circ \varphi_j\}_{(i,j) \in \mathcal{A}^2}$ is also an equicontractive IFS satisfying the convex strong separation condition and the contraction ratio is positive. Moreover, any self-similar measure for $\{\varphi_i\}_{i \in \mathcal{A}}$ can be realised as a self-similar measure for $\{\varphi_i \circ \varphi_j\}_{(i,j) \in \mathcal{A}^2}$.

An important result due to Hutchinson [18] states that for any IFS $\{\varphi_i\}_{i \in \mathcal{A}}$, there exists a unique non-empty compact set X satisfying

$$X = \bigcup_{i \in \mathcal{A}} \varphi_i(X).$$

X is called the self-similar set of $\{\varphi_i\}_{i \in \mathcal{A}}$. The middle 3rd Cantor set and the von-Koch curve are well-known examples of self-similar sets. Given a finite word $\mathbf{a} = (a_1, \ldots, a_M) \in \bigcup_{k=1}^{\infty} \mathcal{A}^k$, we let

$$\varphi_{\mathbf{a}} := \varphi_{a_1} \circ \cdots \circ \varphi_{a_M} \text{ and } X_{\mathbf{a}} := \varphi_{\mathbf{a}}(X).$$

For distinct $\mathbf{a}, \mathbf{b} \in \mathcal{A}^M$ we let

$$|\mathbf{a} \wedge \mathbf{b}| := \inf \left\{ 1 \leq k \leq M : a_k \neq b_k \right\}.$$

Given an IFS $\{\varphi_i\}_{i \in \mathcal{A}}$ and a probability vector $\mathbf{p} := (p_i)_{i \in \mathcal{A}}$, there exists a unique Borel probability measure $\mu_{\mathbf{p}}$ satisfying

$$\mu_{\mathbf{p}} = \sum_{i \in \mathcal{A}} p_i \cdot \mu_{\mathbf{p}} \circ \varphi_i^{-1}. \tag{3.1}$$

We call $\mu_{\mathbf{p}}$ the self-similar measure corresponding to $\{\varphi_i\}_{i \in \mathcal{A}}$ and $\mathbf{p}$. When the choice of $\mathbf{p}$ is implicit, we simply denote $\mu_{\mathbf{p}}$ by μ. For our purposes, it is important that the relation (3.1) can be iterated and for any $M \in \mathbb{N}$ the self-similar measure $\mu_{\mathbf{p}}$ satisfies

$$\mu_{\mathbf{p}} = \sum_{\mathbf{a} \in \mathcal{A}^M} p_{\mathbf{a}} \cdot \mu_{\mathbf{p}} \circ \varphi_{\mathbf{a}}^{-1}, \tag{3.2}$$

where $p_{\mathbf{a}} = \prod_{k=1}^{M} p_{a_k}$ for $\mathbf{a} = (a_1, \ldots, a_M)$. Given a probability vector $\mathbf{p}$, we define the entropy of $\mathbf{p}$ to equal

$$h(\mathbf{p}) := -\sum_{i \in \mathcal{A}} p_i \log p_i.$$

We emphasise that this quantity appears in the hypotheses of Theorems 2.1 and 2.3.

Many results in the study of self-similar sets require additional separation conditions on the IFS. Often, one restricts to the case when the IFS satisfies the strong separation condition or the open set condition (see [13, 14]). In this paper, we will require a slightly stronger separation condition that is still satisfied by many well-known self-similar sets. Given an IFS $\{\varphi_i\}_{i\in\mathcal{A}}$, we say that $\{\varphi_i\}_{i\in\mathcal{A}}$ satisfies the convex strong separation condition if the convex hull of X satisfies the following:

$$\varphi_i(\text{conv}(X)) \cap \varphi_j(\text{conv}(X)) = \emptyset \qquad \forall i \neq j.$$

Iterated function systems satisfying the convex strong separation condition were also studied by Boore and Falconer [5]. It is easy to construct iterated function systems satisfying the convex strong separation condition. For example, if we fix $r \in (0,1)$ and $\{t_i\}_{i=1}^n$ a finite set of real numbers satisfying $t_1 < t_2 < \cdots < t_n$, $1 - r \leq t_1$, $r < t_{i+1} - t_i$, and $t_n \leq t_1 + 1 - r$, then $\{\varphi_i(x) = rx + t_i\}_{i=1}^n$ is an IFS, which satisfies the convex strong separation condition and whose self-similar set is contained in $[1, \infty)$.

To help with our exposition, we state here an identity that will be used several times in our proof of Theorem 2.3. Suppose $\{\varphi_i\}_{i\in\mathcal{A}}$ is an equicontractive IFS and $f \in C^1(\text{conv}(I), \mathbb{R})$. Then for any $\mathbf{a} \in \mathcal{A}^M$, it follows from the chain rule that the following equality holds

$$(f \circ \varphi_{\mathbf{a}})'(x) = r^M f'(\varphi_{\mathbf{a}}(x)). \tag{3.3}$$

3.2 Uniform distribution

To prove Theorem 2.3, we will make use of a well-known criterion due to Weyl for uniform distribution in terms of exponential sums (see [6, Theorem 1.2] and [30]) and a result due to Davenport *et al.* (see [6, Lemma 1.8] and [11]). Combining these results, we may deduce the following statement.

Proposition 3.1. Let μ be a Borel probability measure on $\mathbb{R}$ and $(f_n)_{n=1}^\infty$ be a sequence of continuous real-valued functions. If for any $l \in \mathbb{Z} \setminus \{0\}$, the series

$$\sum_{N=1}^\infty \frac{1}{N} \int \left| \frac{1}{N} \sum_{n=1}^N e^{2\pi i l f_n(x)} \right|^2 \, d\mu$$

converges, then for μ almost every x the sequence $(f_n(x))_{n=1}^\infty$ is uniformly distributed modulo one.

Proposition 3.1 is the tool that enables us to prove Theorem 2.3. We will also rely on the following technical lemma due to van der Corput; for a proof of this lemma, see [23, Lemma 2.1].

Lemma 3.2 (van der Corput lemma).   Let $\phi : [a, b] \to \mathbb{R}$ be differentiable. Assume that $|\phi'(x)| \geq \gamma$ for all $x \in [a, b]$, and ϕ' is monotonic on $[a, b]$. Then

$$\left| \int_a^b e^{2\pi i \phi(x)}\, \mathrm{d}x \right| \leq \gamma^{-1}.$$

Notation. Throughout this paper, we will use $\exp(x)$ to denote $e^{2\pi i x}$. Given two complex valued functions f and g, we write $f = \mathcal{O}(g)$ if there exists $C > 0$ such that $|f(x)| \leq C|g(x)|$ for all x. If the underlying constant depends upon some parameter s, and we want to emphasise this dependence, we write $f = \mathcal{O}_s(g)$. Given an interval I, we let $|I|$ denote the Lebesgue measure of I.

4   Proof of Theorem 2.3

Let us now fix an IFS $\{\varphi_i\}_{i \in \mathcal{A}}$, a probability vector $\mathbf{p}$, and a sequence of functions $(f_n)_{n=1}^\infty$ so that the hypotheses of Theorem 2.3 are satisfied. We let μ denote the self-similar measure corresponding to $\mathbf{p}$. Recall that r denotes the contraction ratio of the elements of $\{\varphi_i\}_{i \in \mathcal{A}}$, and X denotes the corresponding self-similar set. In what follows, we let

$$I := \mathrm{conv}(X).$$

Moreover, given a word $\mathbf{a} \in \cup_{k=1}^\infty \mathcal{A}^k$, we let $I_\mathbf{a} := \varphi_\mathbf{a}(I)$.

Recall that $X \subset [1, \infty)$. For technical reasons, it is useful to restrict our arguments to subsets of X that are a uniform distance away from 1. With this in mind, we let the parameter $\kappa > 0$ denote any small real number such that $1 + \kappa \notin X$. It follows from the convex strong separation condition that κ exists and can be taken to be arbitrarily small. Given such a $\kappa > 0$, we fix $\delta_\kappa > 0$ to be any sufficiently small real number so that if we let

$$\Gamma_\kappa := \max \left\{ r^{\delta_\kappa}, \frac{1}{r^{2\delta_\kappa}} \left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r} \right)^{\frac{\log(1+\kappa)}{-2\log r}}, \frac{1+\delta_\kappa}{r^{3\delta_\kappa}} \left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r} \right)^{\frac{\log(1+\kappa)}{-2\log r}}, \right.$$

$$\left. \frac{1+\delta_\kappa}{r^{3\delta_\kappa}} \left(\frac{e^{-h(\mathbf{p})+\delta_\kappa}}{r^{\delta_\kappa}} \right)^{\frac{\log(1+\kappa)}{-2\log r}} \right\},$$

then

$$\Gamma_\kappa < 1.$$

Such a $\delta_\kappa > 0$ exists because of our underlying assumption

$$\frac{1}{2} < \frac{h(\mathbf{p})}{-\log r},$$

which is equivalent to

$$\frac{e^{-2h(\mathbf{p})}}{r} < 1.$$

Moreover, given such a κ, and δ_κ chosen to be sufficiently small so that the above is satisfied, we fix N_κ to be any sufficiently large natural number so that

$$\max_{\mathbf{a}\in\mathcal{A}^{N_\kappa}} \sup_{x,y\in I_\mathbf{a}} \frac{x}{y} < 1 + \delta_\kappa,$$

and for any $\mathbf{a} \in \mathcal{A}^{N_\kappa}$ we have either

$$\sup I_\mathbf{a} < 1 + \kappa \text{ or } \inf I_\mathbf{a} > 1 + \kappa.$$

Such an N_κ exists because $1 + \kappa \notin X$ and X is compact.

Given a word $\mathbf{c} \in \cup_{k=1}^\infty \mathcal{A}^k$, we let

$$\tilde{\mu}_\mathbf{c} := \frac{\mu|_{X_\mathbf{c}}}{\mu(X_\mathbf{c})}.$$

It is a consequence of the convex strong separation condition that $\tilde{\mu}_\mathbf{c} = \mu \circ \varphi_\mathbf{c}^{-1}$. We will use this equality during our proof of Theorem 2.3.

It is a consequence of the following proposition that we can use Proposition 3.1 to prove Theorem 2.3.

Proposition 4.1.   Assume that $\{\varphi_i\}_{i\in\mathcal{A}}$, $\mathbf{p}$, and $(f_n)_{n=1}^\infty$ satisfy the hypotheses of Theorem 2.3. Then for any $\kappa > 0$ such that $1 + \kappa \notin X$, there exists $\gamma := \gamma(\kappa, \mathbf{p}) \in (0, 1)$ such that for any $l \in \mathbb{Z} \setminus \{0\}$, $n > m$, and $\mathbf{c} \in \mathcal{A}^{N_\kappa}$ satisfying $\inf I_\mathbf{c} > 1 + \kappa$, we have

$$\int \exp(l(f_n(x) - f_m(x))) \, d\tilde{\mu}_\mathbf{c} = \mathcal{O}_{\kappa,l}(\gamma^n).$$

We now include the short argument explaining how Theorem 2.3 follows from Proposition 4.1.

Proof of Theorem 2.3.   It will be shown below that Proposition 4.1 implies that for any $\kappa > 0$ such that $1 + \kappa \notin X$, if $\mathbf{c} \in \mathcal{A}^{N_\kappa}$ is such that $\inf I_{\mathbf{c}} > 1 + \kappa$, then for $\tilde{\mu}_{\mathbf{c}}$ almost every x the sequence $(f_n(x))_{n=1}^\infty$ is uniformly distributed modulo one. It then follows from the definition of N_κ and the self-similarity of μ (i.e., (3.2)), that this statement implies that for μ almost every $x > 1 + \kappa$ the sequence $(f_n(x))_{n=1}^\infty$ is uniformly distributed modulo one. Since there exists arbitrarily small $\kappa > 0$ satisfying $1 + \kappa \notin X$, we may conclude that for μ almost every $x > 1$ the sequence $(f_n(x))_{n=1}^\infty$ is uniformly distributed modulo one. Since $\mu(\{1\}) = 0$, Theorem 2.3 follows. To complete our proof of Theorem 2.3, it suffices to show that our initial statement is true.

Let us now fix $\kappa > 0$ such that $1 + \kappa \notin X$ and $\mathbf{c} \in \mathcal{A}^{N_\kappa}$ such that $\inf I_{\mathbf{c}} > 1 + \kappa$. By Proposition 3.1, to prove that for $\tilde{\mu}_{\mathbf{c}}$ almost every x the sequence $(f_n(x))_{n=1}^\infty$ is uniformly distributed modulo one, it suffices to show that for any $l \in \mathbb{Z} \setminus \{0\}$ we have

$$\sum_{N=1}^\infty \frac{1}{N} \int \left| \frac{1}{N} \sum_{n=1}^N \exp(lf_n(x)) \right|^2 \, \mathrm{d}\tilde{\mu}_{\mathbf{c}}(x) < \infty. \tag{4.1}$$

Expanding this expression, we obtain

$$\sum_{N=1}^\infty \frac{1}{N} \int \left| \frac{1}{N} \sum_{n=1}^N \exp(lf_n(x)) \right|^2 \, \mathrm{d}\tilde{\mu}_{\mathbf{c}}(x)$$
$$= \sum_{N=1}^\infty \left(\frac{1}{N^2} + \frac{1}{N^3} \sum_{\substack{1 \leq n,m \leq N \\ n \neq m}} \int \exp(l(f_n(x) - f_m(x))) \, \mathrm{d}\tilde{\mu}_{\mathbf{c}} \right). \tag{4.2}$$

The $1/N^2$ term appearing in (4.2) does not affect the convergence properties of this series. As such it suffices to consider the remaining terms, which we can rewrite as

$$\sum_{N=1}^\infty \frac{1}{N^3} \sum_{\substack{1 \leq n,m \leq N \\ n \neq m}} \int \exp(l(f_n(x) - f_m(x))) \, \mathrm{d}\tilde{\mu}_{\mathbf{c}} = \sum_{N=1}^\infty \frac{1}{N^3} \sum_{n=2}^N \sum_{m=1}^{n-1} \int \exp(l(f_n(x) - f_m(x))) \, \mathrm{d}\tilde{\mu}_{\mathbf{c}}$$
$$\tag{4.3}$$
$$+ \overline{\sum_{N=1}^\infty \frac{1}{N^3} \sum_{n=2}^N \sum_{m=1}^{n-1} \int \exp(l(f_n(x) - f_m(x))) \, \mathrm{d}\tilde{\mu}_{\mathbf{c}}}.$$

Substituting the bound provided by Proposition 4.1 into (4.3), we obtain

$$\left| \sum_{N=1}^{\infty} \frac{1}{N^3} \sum_{\substack{1 \leq n, m \leq N \\ n \neq m}} \int \exp(l(f_n(x) - f_m(x))) \, \mathrm{d}\tilde{\mu}_{\mathbf{c}} \right| = \mathcal{O}_{\kappa, l}\left(\sum_{N=1}^{\infty} \frac{1}{N^3} \sum_{n=2}^{N} \sum_{m=1}^{n-1} \gamma^n \right)$$

$$= \mathcal{O}_{\kappa, l}\left(\sum_{N=1}^{\infty} \frac{1}{N^3} \sum_{n=2}^{N} n \gamma^n \right)$$

$$= \mathcal{O}_{\kappa, l}\left(\sum_{N=1}^{\infty} \frac{1}{N^3} \right)$$

$$< \infty.$$

In the penultimate line in the above, we have used the fact that $\sum_{n=2}^{N} n \gamma^n$ can be bounded above by a constant independent of N. We see that (4.1) now holds for any $l \in \mathbb{Z} \setminus \{0\}$ and our proof is complete. ∎

4.1 Proof of Proposition 4.1

Throughout the rest of this section, the parameter κ is fixed. We assume that δ_k and N_κ have been chosen so that the properties stated at the start of this section are satisfied. We also fix a word $\mathbf{c} \in \mathcal{A}^{N_\kappa}$ satisfying $\inf X_{\mathbf{c}} > 1 + \kappa$. We start our proof of Proposition 4.1 by defining several objects and collecting some useful estimates.

We let x_0 and x_1 be such that

$$I_{\mathbf{c}} = [x_0, x_1].$$

Recall that by the definition of N_κ we have

$$\frac{x_1}{x_0} < 1 + \delta_\kappa. \tag{4.4}$$

Given $l \in \mathbb{Z} \setminus \{0\}$ and $n \in \mathbb{N}$, we define

$$M = M(\mathbf{c}, l, \kappa, n) := \left\lfloor 1 + \frac{\log 2\pi C_1 |l| |I| + C_2 \log n + (n-1) \log x_1}{-2 \log r} \right\rfloor + \delta_k n.$$

Importantly, M has the property that

$$r^{\delta_\kappa n + N_\kappa + 2} \leq 2\pi C_1 |l| |I| n^{C_2} x_1^{n-1} r^{N_\kappa + 2M} \leq r^{\delta_\kappa n + N_\kappa}. \tag{4.5}$$

Given $k \in \mathbb{N}$, we let

$$B(k) := \left\{ \mathbf{a} \in \mathcal{A}^k : p_{\mathbf{a}} \geq e^{k(-h(\mathbf{p})+\delta_\kappa)} \right\}.$$

It follows from a well-known large deviation result due to Hoeffding [16] that for any $k \in \mathbb{N}$ there exists $\eta := \eta(\kappa, \mathbf{p}) > 0$ such that

$$\sum_{\mathbf{a} \in B(k)} p_{\mathbf{a}} \leq e^{-\eta k}. \tag{4.6}$$

For M as above, we define

$$G_M := \left\{ \mathbf{a} \in \mathcal{A}^M : (a_1, \ldots a_k) \notin B(k), \forall \lfloor \delta_\kappa M \rfloor \leq k \leq M \right\}.$$

It follows from (4.6) and properties of geometric series that

$$\sum_{\substack{\mathbf{a} \in \mathcal{A}^M \\ \mathbf{a} \notin G_M}} p_{\mathbf{a}} = \mathcal{O}_\kappa(e^{-\eta \delta_\kappa M}). \tag{4.7}$$

Given $m < n$, we define the function

$$W_M(x) := \sum_{\mathbf{a} \in G_M} p_{\mathbf{a}} \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)))).$$

The proof of the following lemma is inspired by the proof of [19, Lemma 6.1]. This lemma essentially allows us to bound from above the integral appearing in Proposition 4.1 by the L^2 norm of W_M multiplied by a term that grows exponentially with n.

Lemma 4.2.   Let $m < n$ and $l \in \mathbb{Z} \setminus \{0\}$. For M as defined above, we have

$$\left| \int \exp(l(f_n(x) - f_m(x))) \, d\tilde{\mu}_{\mathbf{c}} \right| \leq \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \int_I |W_M(x)|^2 \, dx + \mathcal{O}_\kappa(r^{\delta_\kappa n} + e^{-\eta \delta_\kappa M}).$$

Proof.   Using first of all the relation $\tilde{\mu}_{\mathbf{c}} = \mu \circ \varphi_{\mathbf{c}}^{-1}$, then (3.2), we can rewrite our integral as follows:

$$\int \exp(l(f_n(x) - f_m(x))) \, d\tilde{\mu}_{\mathbf{c}} = \int \exp(l(f_n(\varphi_{\mathbf{c}}(x)) - f_m(\varphi_{\mathbf{c}}(x)))) \, d\mu$$

$$= \int \sum_{\mathbf{a} \in \mathcal{A}^M} p_{\mathbf{a}} \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)))) \, d\mu.$$

Therefore, it suffices to show that the latter integral satisfies the required bounds. By (4.7), we see that

$$\int \sum_{\mathbf{a} \in \mathcal{A}^M} p_{\mathbf{a}} \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)))) \, \mathrm{d}\mu = \int W_M(x) \, \mathrm{d}\mu + \mathcal{O}_\kappa(e^{-\eta \delta_\kappa M}). \tag{4.8}$$

Let

$$R_M := \left\{ \mathbf{a} \in G_M : \sup_{x \in X_{\mathbf{a}}} |W_M(x)| \geq 2r^{\delta_\kappa n} \right\}.$$

If $\mathbf{a}' \in R_M$, then by the mean value theorem, (3.3), property B for the sequence of functions $(f_n)_{n=1}^\infty$, and (4.5), for all $x \in I_{\mathbf{a}'}$ we have the following:

$|W_M(x)|$

$$\overset{M.V.T.}{\geq} 2r^{\delta_\kappa n} - \sup_{y \in I_{\mathbf{a}'}} |W_M'(y)| \cdot |I_{\mathbf{a}'}|$$

$$\overset{(3.3)}{=} 2r^{\delta_\kappa n} - \sup_{y \in I_{\mathbf{a}'}} \left| \sum_{\mathbf{a} \in G_M} p_{\mathbf{a}} \cdot 2\pi i l r^{N_\kappa + M} (f_n'(\varphi_{\mathbf{ca}}(y)) - f_m'(\varphi_{\mathbf{ca}}(y))) \exp(l(f_n(\varphi_{\mathbf{ca}}(y)) - f_m(\varphi_{\mathbf{ca}}(y)))) \right|$$

$$\cdot r^M |I|$$

$$\overset{PropertyB}{\geq} 2r^{\delta_\kappa n} - \sup_{y \in I_{\mathbf{a}'}} \left(\sum_{\mathbf{a} \in G_M} p_{\mathbf{a}} \cdot 2\pi |l| r^{N_\kappa + M} C_1 n^{C_2} \varphi_{\mathbf{ca}}(y)^{n-1} \right) \cdot r^M |I|$$

$$\geq 2r^{\delta_\kappa n} - \left(\sum_{\mathbf{a} \in G_M} p_{\mathbf{a}} \cdot 2\pi |l| r^{N_\kappa + M} C_1 n^{C_2} x_1^{n-1} \right) \cdot r^M |I|$$

$$\geq 2r^{\delta_\kappa n} - 2\pi C_1 |l| |I| n^{C_2} x_1^{n-1} r^{N_\kappa + 2M}$$

$$\overset{(4.5)}{\geq} 2r^{\delta_\kappa n} - r^{\delta_\kappa n + N_\kappa}$$

$$\geq r^{\delta_\kappa n}.$$

We have shown that

$$|W_M(x)| \geq r^{\delta_\kappa n} \tag{4.9}$$

for all $x \in I_{\mathbf{a}'}$ for any $\mathbf{a}' \in R_M$. Now notice that for any $\mathbf{a} \in R_M$ we have

$$\int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu \leq p_{\mathbf{a}} \quad \text{and} \quad p_{\mathbf{a}} \leq e^{M(-h(\mathbf{p}) + \delta_\kappa)}.$$

It follows that

$$\sum_{\mathbf{a} \in R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu \leq \sum_{\mathbf{a} \in R_M} p_{\mathbf{a}} \leq \sum_{\mathbf{a} \in R_M} e^{M(-h(\mathbf{p})+\delta_\kappa)}.$$

Combining this upper bound with (4.9), we obtain

$$\sum_{\mathbf{a} \in R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu \leq \sum_{\mathbf{a} \in R_M} e^{M(-h(\mathbf{p})+\delta_\kappa)}$$

$$= \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \sum_{\mathbf{a} \in R_M} r^M |I| \cdot r^{2\delta_\kappa n}$$

$$\leq \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \sum_{\mathbf{a} \in R_M} \int_{I_{\mathbf{a}}} |W_M(x)|^2 \, \mathrm{d}x$$

$$\leq \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \int_I |W_M(x)|^2 \, \mathrm{d}x.$$

In the last line, we used that for distinct $\mathbf{a}, \mathbf{b} \in \mathcal{A}^M$ the intervals $I_{\mathbf{a}}$ and $I_{\mathbf{b}}$ are disjoint. Using this upper bound, together with (4.7) and the definition of R_M, we obtain

$$\left| \int W_M(x) \, \mathrm{d}\mu \right| \leq \int |W_M(x)| \, \mathrm{d}\mu = \sum_{\mathbf{a} \in G_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + \sum_{\substack{\mathbf{a} \in \mathcal{A}^M \\ \mathbf{a} \notin G_M}} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu$$

$$\leq \sum_{\mathbf{a} \in G_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + \sum_{\substack{\mathbf{a} \in \mathcal{A}^M \\ \mathbf{a} \notin G_M}} p_{\mathbf{a}}$$

$$\leq \sum_{\mathbf{a} \in R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + \sum_{\mathbf{a} \in G_M \setminus R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + \mathcal{O}_\kappa(e^{-\eta \delta_\kappa M})$$

$$\leq \sum_{\mathbf{a} \in R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + \sum_{\mathbf{a} \in G_M \setminus R_M} p_{\mathbf{a}} \cdot 2r^{\delta_\kappa n} + \mathcal{O}_\kappa(e^{-\eta \delta_\kappa M})$$

$$\leq \sum_{\mathbf{a} \in R_M} \int_{X_{\mathbf{a}}} |W_M(x)| \, \mathrm{d}\mu + 2r^{\delta_\kappa n} + \mathcal{O}_\kappa(e^{-\eta \delta_\kappa M})$$

$$\leq \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \int_I |W_M(x)|^2 \, \mathrm{d}x + \mathcal{O}_\kappa(r^{\delta_\kappa n} + e^{-\eta \delta_\kappa M}).$$

Substituting this bound into (4.8), we obtain

$$\left| \int \sum_{\mathbf{a} \in \mathcal{A}^M} p_{\mathbf{a}} \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)))) \, \mathrm{d}\mu \right| \leq \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{|I| \cdot r^{M+2\delta_\kappa n}} \int_I |W_M(x)|^2 \, \mathrm{d}x + \mathcal{O}_\kappa(r^{\delta_\kappa n} + e^{-\eta \delta_\kappa M})$$

as required. ∎

　　To complete our proof of Proposition 4.1, it is necessary to obtain good upper bounds for $\int_I |W_M(x)|^2 \, \mathrm{d}x$. These bounds are provided by the following lemma.

Lemma 4.3.   Let $m < n$ and $l \in \mathbb{Z} \setminus \{0\}$. For M as defined above, we have

$$\int_I |W_M(x)|^2 \, \mathrm{d}x = |I| \cdot e^{M(-h(\mathbf{p})+\delta_\kappa)} + \mathcal{O}_{\kappa,l} \left(\frac{1}{r^{M+\lfloor \delta_\kappa M \rfloor} x_0^n} + \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M} x_0^n} \right).$$

Proof.   We start by expanding $\int_I |W_M(x)|^2 \, \mathrm{d}x$:

$$\int_I |W_M(x)|^2 \, \mathrm{d}x$$
$$= |I| \sum_{\mathbf{a} \in G_M} p_{\mathbf{a}}^2 + \sum_{\substack{\mathbf{a},\mathbf{b} \in G_M \\ \mathbf{a} \neq \mathbf{b}}} p_{\mathbf{a}} \cdot p_{\mathbf{b}} \int_I \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)) - f_n(\varphi_{\mathbf{cb}}(x)) + f_m(\varphi_{\mathbf{cb}}(x)))) \, \mathrm{d}x$$
$$\leq |I| \cdot e^{M(-h(\mathbf{p})+\delta_\kappa)} + \sum_{\substack{\mathbf{a},\mathbf{b} \in G_M \\ \mathbf{a} \neq \mathbf{b}}} p_{\mathbf{a}} \cdot p_{\mathbf{b}} \int_I \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)) - f_n(\varphi_{\mathbf{cb}}(x)) + f_m(\varphi_{\mathbf{cb}}(x)))) \, \mathrm{d}x.$$

$$(4.10)$$

To bound the integral appearing in the summation in (4.10) we will use Lemma 3.2. Before doing this, we demonstrate below that the hypotheses of this lemma are satisfied.

Verifying the hypotheses of Lemma 3.2. Fix $\mathbf{a}, \mathbf{b} \in G_M$ such that $\mathbf{a} \neq \mathbf{b}$. Let

$$\phi(x) := l \left(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)) - f_n(\varphi_{\mathbf{cb}}(x)) + f_m(\varphi_{\mathbf{cb}}(x)) \right).$$

By (3.3), we have

$$\phi'(x) = r^{N_\kappa + M} l \left(f_n'(\varphi_{\mathbf{ca}}(x)) - f_m'(\varphi_{\mathbf{ca}}(x)) - f_n'(\varphi_{\mathbf{cb}}(x)) + f_m'(\varphi_{\mathbf{cb}}(x)) \right).$$

Define

$$h_{n,m}(x) := f_n'(x) - f_m'(x).$$

Then

$$\phi'(x) = r^{N_\kappa + M} l \left(h_{n,m}(\varphi_{\mathbf{ca}}(x)) - h_{n,m}(\varphi_{\mathbf{cb}}(x)) \right).$$

Applying the mean value theorem to the function $h_{n,m}$, we see that there exists $z \in I_{\mathbf{c}}$ such that

$$\phi'(x) = r^{N_\kappa + M} l \left(\varphi_{\mathbf{ca}}(x) - \varphi_{\mathbf{cb}}(x) \right) \left(f_n''(z) - f_m''(z) \right). \qquad (4.11)$$

It follows from the convex strong separation condition that there exists $c_0 > 0$ depending only on our underlying IFS such that

$$|\varphi_{\mathbf{ca}}(x) - \varphi_{\mathbf{cb}}(x)| \geq c_0 r^{N_\kappa + |\mathbf{a} \wedge \mathbf{b}|} \qquad (4.12)$$

for all $x \in I$. Using property C for our sequence of functions $(f_n)_{n=1}^\infty$, and the fact $z \in I_{\mathbf{c}}$ so $z \geq x_0$, it follows that

$$|f_n''(z) - f_m''(z)| \geq C_3 z^{n-2} \geq C_3 x_0^{n-2}. \qquad (4.13)$$

Substituting (4.12) and (4.13) into (4.11), we see that for all $x \in I$ we have

$$|\phi'(x)| \geq c_0 C_3 l r^{2N_\kappa + M + |\mathbf{a} \wedge \mathbf{b}|} x_0^{n-2}. \qquad (4.14)$$

The right-hand side of (4.14) is the value of γ we will use in our application of Lemma 3.2. It remains to check that ϕ' satisfies the monotonicity hypothesis of Lemma 3.2. Differentiating ϕ' and applying (3.3), we have

$$\phi''(x) = r^{2(N_\kappa + M)} l \left(f_n''(\varphi_{\mathbf{ca}}(x)) - f_m''(\varphi_{\mathbf{ca}}(x)) - f_n''(\varphi_{\mathbf{cb}}(x)) + f_m''(\varphi_{\mathbf{cb}}(x)) \right).$$

Applying the mean value theorem as above, this time to the function $f_n''(x) - f_m''(x)$, we may deduce that there exists $z \in I_{\mathbf{c}}$ such that

$$\phi''(x) = r^{2(N_\kappa + M)} l (\varphi_{\mathbf{ca}}(x) - \varphi_{\mathbf{cb}}(x))(f_n'''(z) - f_m'''(z)).$$

By property D, for our sequence of functions $(f_n)_{n=1}^\infty$, we know that $f_n'''(z) - f_m'''(z) \geq 0$ for all $z \in I_{\mathbf{c}}$ or $f_n'''(z) - f_m'''(z) \leq 0$ for all $z \in I_{\mathbf{c}}$. What is more, it follows from the convex strong separation condition that the sign of $\varphi_{\mathbf{ca}}(x) - \varphi_{\mathbf{cb}}(x)$ is independent of x and

depends solely upon $\mathbf{a}$ and $\mathbf{b}$. Therefore, we must have $\phi''(x) \leq 0$ for all $x \in I$ or $\phi'' \geq 0$ for all $x \in I$. In either case, ϕ' is monotonic, and we have shown that the monotonicity condition of Lemma 3.2 is satisfied.

Return to the proof of Lemma 4.3. Taking the right-hand side of (4.14) as our value of γ in Lemma 3.2, we obtain

$$\int_I \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)) - f_n(\varphi_{\mathbf{cb}}(x)) + f_m(\varphi_{\mathbf{cb}}(x)))) \, dx = \mathcal{O}_{\kappa,l}\left(\frac{1}{r^{M+|\mathbf{a}\wedge\mathbf{b}|}x_0^n}\right).$$
(4.15)

Substituting (4.15) into the summation appearing in (4.10), and using the definition of G_M, we see that the following holds:

$$\sum_{\substack{\mathbf{a},\mathbf{b}\in G_M \\ \mathbf{a}\neq\mathbf{b}}} p_{\mathbf{a}} \cdot p_{\mathbf{b}} \int_I \exp(l(f_n(\varphi_{\mathbf{ca}}(x)) - f_m(\varphi_{\mathbf{ca}}(x)) - f_n(\varphi_{\mathbf{cb}}(x)) + f_m(\varphi_{\mathbf{cb}}(x)))) \, dx$$

$$= \mathcal{O}_{\kappa,l}\left(\sum_{\mathbf{a}\in G_M} \sum_{\substack{\mathbf{b}\in G_M \\ \mathbf{a}\neq\mathbf{b}}} \frac{p_{\mathbf{a}} \cdot p_{\mathbf{b}}}{r^{M+|\mathbf{a}\wedge\mathbf{b}|}x_0^n}\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^M x_0^n} \sum_{\mathbf{a}\in G_M} p_{\mathbf{a}} \sum_{k=1}^{M} \sum_{\substack{\mathbf{b}\in G_M \\ |\mathbf{a}\wedge\mathbf{b}|=k}} \frac{p_{\mathbf{b}}}{r^k}\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^M x_0^n} \sum_{\mathbf{a}\in G_M} p_{\mathbf{a}} \sum_{k=1}^{M} \frac{\prod_{j=1}^{k} p_{\mathbf{a}_j}}{r^k}\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^M x_0^n} \sum_{\mathbf{a}\in G_M} p_{\mathbf{a}} \left(\sum_{k=1}^{\lfloor\delta_\kappa M\rfloor-1} \frac{\prod_{j=1}^{k} p_{\mathbf{a}_j}}{r^k} + \sum_{k=\lfloor\delta_\kappa M\rfloor}^{M} \frac{\prod_{j=1}^{k} p_{\mathbf{a}_j}}{r^k}\right)\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^M x_0^n} \sum_{\mathbf{a}\in G_M} p_{\mathbf{a}} \left(\sum_{k=1}^{\lfloor\delta_\kappa M\rfloor-1} \frac{1}{r^k} + \sum_{k=\lfloor\delta_\kappa M\rfloor}^{M} \frac{e^{k(-h(\mathbf{p})+\delta_\kappa)}}{r^k}\right)\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^M x_0^n} \sum_{\mathbf{a}\in G_M} p_{\mathbf{a}} \left(\frac{1}{r^{\lfloor\delta_\kappa M\rfloor}} + \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^M}\right)\right)$$

$$= \mathcal{O}_{\kappa,l}\left(\frac{1}{r^{M+\lfloor\delta_\kappa M\rfloor}x_0^n} + \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M}x_0^n}\right).$$

Substituting this bound into (4.10), we obtain

$$\int_I |W_M(x)|^2 \, dx = |I| \cdot e^{M(-h(\mathbf{p})+\delta_\kappa)} + \mathcal{O}_{\kappa,l}\left(\frac{1}{r^{M+\lfloor \delta_\kappa M \rfloor} x_0^n} + \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M} x_0^n} \right)$$

as required.   ∎

We are now in a position to prove Proposition 4.1 and in doing so complete our proof of Theorem 2.3.

Proof of Proposition 4.1.   Assume that $m < n$. Combining Lemmas 4.2 and 4.3, we obtain

$$\left| \int \exp(l(f_n(x) - f_m(x))) \, d\tilde{\mu}_{\mathbf{c}} \right|$$

$$\leq \underbrace{\frac{e^{2M(-h(\mathbf{p})+\delta_\kappa)}}{r^{M+2\delta_\kappa n}}}_{(1)} + \mathcal{O}_{\kappa,l}\left(\underbrace{\frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M+2\delta_\kappa n+\lfloor \delta_\kappa M \rfloor} x_0^n}}_{(2)} + \underbrace{\frac{e^{2M(-h(\mathbf{p})+\delta_\kappa)}}{r^{3M+2\delta_\kappa n} x_0^n}}_{(3)} + \underbrace{r^{\delta_\kappa n}}_{(4)} + \underbrace{e^{-\eta \delta_\kappa M}}_{(5)} \right). \qquad (4.16)$$

It remains to show that the terms (1)–(5) decay to zero exponentially fast with respect to n. To do this, it is useful to recall the definition of Γ_κ and recall that we chose δ_κ in such a way that $\Gamma_\kappa < 1$:

$$\Gamma_\kappa := \max\left\{ r^{\delta_\kappa}, \frac{1}{r^{2\delta_\kappa}} \left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r} \right)^{\frac{\log(1+\kappa)}{-2\log r}}, \frac{1+\delta_\kappa}{r^{3\delta_\kappa}} \left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r} \right)^{\frac{\log(1+\kappa)}{-2\log r}}, \right.$$

$$\left. \frac{1+\delta_\kappa}{r^{3\delta_\kappa}} \left(\frac{e^{-h(\mathbf{p})+\delta_\kappa}}{r^{\delta_\kappa}} \right)^{\frac{\log(1+\kappa)}{-2\log r}} \right\}.$$

As we will see, most of the terms in (4.16) can be bounded in terms of Γ_κ. To help with our exposition, we treat each of the five terms described above individually.

Bounding (1). A useful inequality that follows from the definition of M is that for n sufficiently large we have

$$M \geq n \cdot \frac{\log x_1}{-2\log r}. \qquad (4.17)$$

This inequality follows upon noticing that the floor term appearing in the definition of M can be bounded below by $\frac{(n-1)\log x_1}{-2\log r}$ for n sufficiently large and then using the additional $\delta_\kappa n$ term. Applying (4.17), the fact $x_1 \geq 1 + \kappa$, and the definition of Γ_κ, we see that the following holds for n sufficiently large:

$$
\frac{e^{2M(-h(\mathbf{p})+\delta_\kappa)}}{r^{M+2\delta_\kappa n}} = \frac{1}{r^{2\delta_\kappa n}}\left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r}\right)^M \overset{(4.17)}{\leq} \left(\frac{1}{r^{2\delta_\kappa}}\left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r}\right)^{\frac{\log x_1}{-2\log r}}\right)^n
$$

$$
\leq \left(\frac{1}{r^{2\delta_\kappa}}\left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r}\right)^{\frac{\log(1+\kappa)}{-2\log r}}\right)^n
$$

$$
\leq \Gamma_\kappa^n. \tag{4.18}
$$

Bounding (2). Applying (4.4), (4.5), and (4.17), we have

$$
\frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M+2\delta_\kappa n+\lfloor\delta_\kappa M\rfloor}x_0^n} = \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M+2\delta_\kappa n+\lfloor\delta_\kappa M\rfloor}x_1^n}\left(\frac{x_1}{x_0}\right)^n
$$

$$
\overset{(4.4)}{\leq} \frac{e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{2M+2\delta_\kappa n+\lfloor\delta_\kappa M\rfloor}x_1^n}(1+\delta_\kappa)^n
$$

$$
\overset{(4.5)}{=} \mathcal{O}_{\kappa,l}\left(\frac{n^{C_2}e^{M(-h(\mathbf{p})+\delta_\kappa)}}{r^{3\delta_\kappa n+\lfloor\delta_\kappa M\rfloor}}(1+\delta_\kappa)^n\right)
$$

$$
= \mathcal{O}_{\kappa,l}\left(n^{C_2}\left(\frac{1+\delta_\kappa}{r^{3\delta_\kappa}}\right)^n\left(\frac{e^{(-h(\mathbf{p})+\delta_\kappa)}}{r^{\delta_\kappa}}\right)^M\right)
$$

$$
\overset{(4.17)}{=} \mathcal{O}_{\kappa,l}\left(n^{C_2}\left(\frac{1+\delta_\kappa}{r^{3\delta_\kappa}}\left(\frac{e^{(-h(\mathbf{p})+\delta_\kappa)}}{r^{\delta_\kappa}}\right)^{\frac{\log x_1}{-2\log r}}\right)^n\right)
$$

$$
= \mathcal{O}_{\kappa,l}\left(n^{C_2}\left(\frac{1+\delta_\kappa}{r^{3\delta_\kappa}}\left(\frac{e^{(-h(\mathbf{p})+\delta_\kappa)}}{r^{\delta_\kappa}}\right)^{\frac{\log(1+\kappa)}{-2\log r}}\right)^n\right)
$$

$$
= \mathcal{O}_{\kappa,l}\left(n^{C_2}\Gamma_\kappa^n\right)
$$

$$
= \mathcal{O}_{\kappa,l}\left(\Gamma_\kappa^{n/2}\right). \tag{4.19}
$$

Bounding (3). Repeating the argument used to bound (2), one can show that

$$\frac{e^{2M(-h(\mathbf{p})+\delta_\kappa)}}{r^{3M+2\delta_\kappa n}x_0^n} = \mathcal{O}_{\kappa,l}\left(\Gamma_\kappa^{n/2}\right). \tag{4.20}$$

It is during this part of the proof that we use the fact that

$$\frac{1+\delta_\kappa}{r^{3\delta_\kappa}}\left(\frac{e^{2(-h(\mathbf{p})+\delta_\kappa)}}{r}\right)^{\frac{\log(1+\kappa)}{-2\log r}} \le \Gamma_\kappa.$$

Bounding (4). It is immediate from the definition of Γ_κ that we have

$$r^{\delta_\kappa n} \le \Gamma_\kappa^n. \tag{4.21}$$

Bounding (5). Applying (4.17) and the inequality $\log x_1 \ge \log(1+\kappa)$, we see that the following holds for n sufficiently large:

$$e^{-\eta\delta_\kappa M} \overset{(4.17)}{\le} e^{\frac{\eta\delta_\kappa \log x_1}{2\log r}\cdot n} \le e^{\frac{\eta\delta_\kappa \log(1+\kappa)}{2\log r}\cdot n} \tag{4.22}$$

We now let

$$\gamma = \max\{\Gamma_\kappa^{1/2}, e^{\frac{\eta\delta_\kappa \log(1+\kappa)}{2\log r}}\}.$$

Notice that $\gamma \in (0,1)$. Substituting (4.18), (4.19), (4.20), (4.21), and (4.22) into (4.16), we obtain

$$\left|\int \exp(l(f_n(x) - f_m(x)))\, d\tilde{\mu}_{\mathbf{c}}\right| = \mathcal{O}_{\kappa,l}\left(\gamma^n\right).$$

This completes our proof. ∎

Acknowledgments

References

[1] Aistleitner, C. "Quantitative uniform distribution results for geometric progressions." *Israel J. Math.* 204, no. 1 (2014): 155–97.

[2] Aistleitner, C. and S. Baker. "On the pair correlations of powers of real numbers." *Israel J. Math.* (forthcoming).

[3] Aistleitner, C., S. Baker, N. Technau, and N. Yesha. "Gap statistics and higher correlations for geometric progressions modulo one." arXiv:2010.10355 (forthcoming).

[4] Baker, S. "On the distribution of powers of real numbers modulo 1." *Unif. Distrib. Theory* 10, no. 2 (2015): 67–75.

[5] Boore, G. and K. Falconer. "Attractors of directed graph IFSs that are not standard IFS attractors and their Hausdorff measure." *Math. Proc. Cambridge Philos. Soc.* 154, no. 2 (2013): 325–49.

[6] Bugeaud, Y. *Distribution Modulo One and Diophantine Approximation*. Cambridge Tracts in Mathematics 193. Cambridge: Cambridge University Press, 2012.

[7] Bugeaud, Y., L. Liao, and M. Rams. "Metrical results on the distribution of fractional parts of powers of real numbers." *Proc. Edinburgh Math. Soc.* 62 (2019): 505–21.

[8] Bugeaud, Y. and N. Moshchevitin. "On fractional parts of powers of real numbers close to 1." *Math. Z.* 271, no. 3–4 (2012): 627–37.

[9] Cassels, J. W. S. "On a problem of Steinhaus about normal numbers." *Colloq. Math.* 7 (1959): 95–101.

[10] Dayan, Y., A. Ganguly, and B. Weiss. "Random walks on tori and normal numbers in self similar sets." arXiv:2002.00455 [math.DS] (forthcoming).

[11] Davenport, H., P. Erdoős, and W. J. LeVeque. "On Weyl's criterion for uniform distribution." *Michigan Math. J.* 10 (1963): 311–4.

[12] Dubickas, A. "On the powers of some transcendental numbers." *Bull. Austral. Math. Soc.* 76, no. 3 (2007): 433–40.

[13] Falconer, K. *Fractal Geometry. Mathematical Foundations and Applications*, 3rd ed., xxx+368. Chichester: John Wiley & Sons, Ltd., 2014.

[14] Falconer, K. *Techniques in Fractal Geometry*, xviii+256. Chichester: John Wiley & Sons, Ltd., 1997.

[15] Hardy, G. H. "A problem of Diophantine approximation." *J. Indian Math. Soc.* 11 (1919): 162–6.

[16] Hoeffding, W. "Probability inequalities for sums of bounded random variables." *J. Amer. Statist. Assoc.* 58 (1963): 13–30.

[17] Hochman, M. and P. Shmerkin. "Equidistribution from fractal measures." *Invent. Math.* 202, no. 1 (2015): 427–79.

[18] Hutchinson, J. "Fractals and self-similarity." *Indiana Univ. Math. J.* 30, no. 5 (1981): 713–47.

[19] Jordan, T. and T. Sahlsten. "Fourier transforms of Gibbs measures for the gauss map." *Math. Ann.* 364, no. 3–4 (2016): 983–1023.

[20] Kaufman, R. "Continued fractions and Fourier transforms." *Mathematika* 27, no. 2 (1980): 262–7.

[21] Kaufman, R. "On the theorem of Jarník and Besicovitch." *Acta Arith.* 39, no. 3 (1981): 265–7.

[22] Koksma, J. F. "Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins." *Compositio Math.* 2 (1935): 250–8.

[23] Kuipers, L. and H. Niederreiter. *Uniform Distribution of Sequences*. New York–London–Sydney: Wiley-Interscience, John Wiley & Sons, 1974.

[24] Pisot, C. "La répartition modulo 1 et les nombres algébriques." *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (2)* 7, no. 3–4 (1938): 205–48.

[25] Pisot, C. "Sur la répartition modulo 1 des puissances successives d'un même nombre." *C. R. Acad. Sci. Paris* 204 (1937): 312–4.

[26] Pollington, A., S. Velani, A. Zafeiropoulos, and E. Zorin. "Inhomogeneous Diophantine approximation on M_0-sets with restricted denominators." *Int. Math. Res. Not.* (2021). https://doi.org/10.1093/imrn/rnaa307.

[27] Queffélec, M. and O. Ramaré. "Analyse de Fourier des fractions continues à quotients restreints." *Enseign. Math.* 49, no. 3–4 (2003): 335–56.

[28] Schmidt, W. "On normal numbers." *Pacific J. Math.* 10 (1960): 661–72.

[29] Simmons, D. and B. Weiss. "Random walks on homogeneous spaces and Diophantine approximation on fractals." *Invent. Math.* 216, no. 2 (2019): 337–94.

[30] Weyl, H. "Über die Gleichverteilung von Zahlen mod. Eins." *Math. Ann.* 77, no. 3 (1916): 313–52.