Hausdorff dimension of univoque sets and Devil's staircase

Vilmos Komornik [a], Derong Kong [b,*], Wenxia Li [c]

[a] *Département de mathématique, Université de Strasbourg, 7 rue René Descartes, 67084 Strasbourg Cedex, France*
[b] *School of Mathematical Science, Yangzhou University, Yangzhou, JiangSu 225002, People's Republic of China*
[c] *Department of Mathematics, Shanghai Key Laboratory of PMMP, East China Normal University, Shanghai 200062, People's Republic of China*

A R T I C L E   I N F O

A B S T R A C T

We fix a positive integer M, and we consider expansions in arbitrary real bases $q > 1$ over the alphabet $\{0, 1, \ldots, M\}$. We denote by $\mathcal{U}_q$ the set of real numbers having a unique expansion. Completing many former investigations, we give a formula for the Hausdorff dimension $D(q)$ of $\mathcal{U}_q$ for each $q \in (1, \infty)$. Furthermore, we prove that the dimension function $D : (1, \infty) \to [0, 1]$ is continuous, and has bounded variation. Moreover, it has a Devil's staircase behavior in (q', ∞), where q' denotes the Komornik–Loreti constant: although $D(q) > D(q')$ for all $q > q'$, we have $D' < 0$ a.e. in (q', ∞). During the proofs we improve and generalize a theorem of Erdős et al. on the existence of large blocks of zeros in β-expansions, and we determine for all M the Lebesgue measure and the Hausdorff dimension of the set $\mathcal{U}$ of bases in which $x = 1$ has a unique expansion.

© 2016 Elsevier Inc. All rights reserved.

* Corresponding author.
   E-mail addresses: komornik@math.unistra.fr (V. Komornik), derongkong@126.com (D. Kong), wxli@math.ecnu.edu.cn (W. Li).

1. Introduction

Fix a positive integer M and an *alphabet* $\{0, 1, \ldots, M\}$. By a *sequence* we mean an element $c = (c_i)$ of $\{0, 1, \ldots, M\}^{\infty}$.

Given a real *base* $q > 1$, by an *expansion* of a real number x we mean a sequence $c = (c_i)$ satisfying the equality

$$\pi_q(c) := \sum_{i=1}^{\infty} \frac{c_i}{q^i} = x.$$

Expansions of this type in *non-integer* bases have been extensively investigated since a pioneering paper of Rényi [29]. One of the striking features of such bases is that generically a number has a continuum of different expansions, a situation quite opposite to that of integer bases; see, e.g., [12] and Sidorov [31]. However, surprising unique expansions have also been discovered by Erdős et al. [13], and they have stimulated many works during the last 25 years.

We refer to the papers [23,6–9,3] and surveys [32,20,10] for more information.

Let us denote by $\mathcal{U}_q$ the set of numbers x having a unique expansion and by $\mathcal{U}'_q$ the set of the corresponding expansions. The topological and combinatorial structure of these sets have been described in [8]. The present paper is a natural sequel of this work, concerning the measure-theoretical aspects.

Daróczy and Kátai [5] have determined the Hausdorff dimension of $\mathcal{U}_q$ when $M = 1$ and q is a Parry number. Their results were extended by Kallós and Kátai [17–19], Glendinning and Sidorov [15], Kong et al. [25,24], and in [9,1].

We recall from [21] and [22] that there exists a smallest base $1 < q' < M+1$ (depending on M) in which $x = 1$ has a unique expansion: the so-called *Komornik–Loreti constant*.

We also recall two theorems on the *dimension function*

$$D(q) := \dim_H \mathcal{U}_q, \quad 1 < q < \infty,$$

obtained respectively in [15,25] and in [24]:

Theorem 1.1. *The function D vanishes in $(1, q']$, and $D > 0$ in (q', ∞). Its maximum $D(q) = 1$ is attained only in $q = M + 1$.*

It follows from this theorem that $\mathcal{U}_q$ is a (Lebesgue) null set for all $q \neq M + 1$, while $\mathcal{U}_{M+1} \subseteq [0, 1]$ has measure one because its complementary set is countable in $[0, 1]$. Since $\overline{\mathcal{U}_q} \setminus \mathcal{U}_q$ is countable for each q (see [8]), the same properties hold for $\overline{\mathcal{U}_q}$ as well.

Theorem 1.2. *For almost all $q > 1$, $\mathcal{U}'_q$ is a subshift, and*

$$D(q) = \frac{h(\mathcal{U}_q')}{\log q}, \tag{1.1}$$

where $h(\mathcal{U}_q')$ denotes the topological entropy of $\mathcal{U}_q'$.

Furthermore, the function D is differentiable almost everywhere.

We recall from Lind and Marcus [26] that

$$h(\mathcal{U}_q') = \lim_{n \to \infty} \frac{\log |B_n(\mathcal{U}_q')|}{n} = \inf_{n \geq 1} \frac{\log |B_n(\mathcal{U}_q')|}{n} \tag{1.2}$$

when $\mathcal{U}_q'$ is a subshift, where $B_n(\mathcal{U}_q')$ denotes the set of different initial words of length n occurring in the sequences $(c_i) \in \mathcal{U}_q'$, and $|B_n(\mathcal{U}_q')|$ means the cardinality of $B_n(\mathcal{U}_q')$. (Unless otherwise stated, in this paper we use base two logarithms.)

We will complete and improve Theorems 1.1 and 1.2 in Theorems 1.3, 1.4 and 1.7 below.

Theorem 1.3. *The formula* (1.1) *is valid for* all $q > 1$.

We recall from [8] that $\mathcal{U}_q'$ is not always a subshift. Theorem 1.3 states in particular that the limit in (1.2) exists even if $\mathcal{U}_q'$ is not a subshift, and it is equal to the infimum in (1.2).

Theorem 1.4. *The function D is continuous, and has bounded variation.*

Theorem 1.4 implies again that D is differentiable almost everywhere. In order to describe its derivative first we establish some results on general β-expansions and on univoque bases.

Following Rényi [29] we denote by $\beta(q) = (\beta_i(q))$ the lexicographically largest expansion of $x = 1$ in base q. It is also called the greedy or β-expansion of $x = 1$ in base q.

Theorem 1.5. *Fix $1 < r \leq M + 1$ arbitrarily. For almost all $q \in (1, r)$ there exist arbitrarily large integers m such that $\beta_1(q) \cdots \beta_m(q)$ ends with more than $\log_r m$ consecutive zero digits.*

This theorem improves and generalizes [12, Theorem 2] concerning the case $M = 1$. In particular, our result implies that $\beta(q)$ contains arbitrarily large blocks of consecutive zeros for almost all $q \in (1, M + 1]$. This was first established by Erdős and Joó [11] for $M = 1$, and their result was extended by Schmeling [30] for all M.

Next we denote by $\mathcal{U}$ the set of bases $q > 1$ in which $x = 1$ has a unique expansion, and by $\overline{\mathcal{U}}$ its closure. The elements of $\mathcal{U}$ are usually called *univoque bases*.
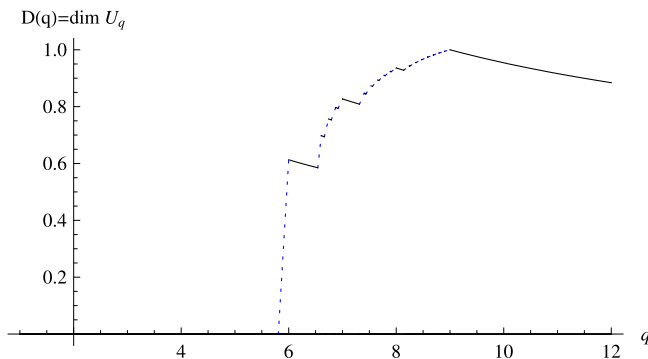
Fig. 1. Approximation plot for the graph of $D(q) = \dim_H \mathcal{U}_q$ with $M = 8$ and $q \in (1, 12]$. In this case $q' \approx 5.80676$, and $D(q) = 1$ only if $q = M + 1 = 9$.

Theorem 1.6.

  (i) $\mathcal{U}$ and $\overline{\mathcal{U}}$ are (Lebesgue) null sets.
 (ii) $\mathcal{U}$ and $\overline{\mathcal{U}}$ have Hausdorff dimension one.

This theorem was proved for $\mathcal{U}$ in case $M = 1$ by Erdős and Joó [11] and by Daróczy and Kátai [4], respectively. The case of $\overline{\mathcal{U}}$ hence follows because the set $\overline{\mathcal{U}} \setminus \mathcal{U}$ is countable (see [23]). Our proof of (ii) is shorter than the original one even for $M = 1$.

Finally, combining Theorems 1.1, 1.3, 1.4, 1.6 (i) and some topological results of [8] we prove that the dimension function is a natural variant of *Devil's staircase* (see Fig. 1):

Theorem 1.7.

  (i) D is continuous in $[q', \infty)$.
 (ii) $D' < 0$ almost everywhere in (q', ∞).
(iii) $D(q') < D(q)$ for all $q > q'$.

Remark. Compared to the classical Cantor function, we even have $D' < 0$ instead of $D' = 0$ almost everywhere.

The paper is organized as follows. In Section 2 we investigate the topological entropy of various subshifts that we need in the sequel. In Section 3 we prove Theorem 1.3 and we prepare the proof of Theorem 1.4. Theorem 1.4 is proved in Section 4, Theorems 1.5–1.6 in Sections 5–6, and Theorem 1.7 in Section 7. Sections 5–6 are independent of each other and of the other sections of the paper.

2. Topological entropies

We begin by proving that the topological entropy of $\mathcal{U}_q'$ is well defined even if $\mathcal{U}_q'$ is not a subshift:

Lemma 2.1. *The limit*

$$h(\mathcal{U}_q') := \lim_{n \to \infty} \frac{\log |B_n(\mathcal{U}_q')|}{n}$$

exists for each $q > 1$, and is equal to

$$\inf_{n \geq 1} \frac{\log |B_n(\mathcal{U}_q')|}{n}.$$

Proof. It suffices to show that the function $n \mapsto |B_n(\mathcal{U}_q')|$ is submultiplicative, i.e.,

$$|B_{m+n}(\mathcal{U}_q')| \leq |B_m(\mathcal{U}_q')| \cdot |B_n(\mathcal{U}_q')|$$

for all $m, n \geq 1$.

Denoting by $B_{k,\ell}(\mathcal{U}_q')$ the set of words $c_k \cdots c_\ell$ where (c_i) runs over $\mathcal{U}_q'$, we have clearly

$$|B_{m+n}(\mathcal{U}_q')| = |B_{1,m+n}(\mathcal{U}_q')| \leq |B_{1,m}(\mathcal{U}_q')| \cdot |B_{m+1,m+n}(\mathcal{U}_q')|.$$

Notice that $|B_{m+1,m+n}(\mathcal{U}_q')| \leq |B_n(\mathcal{U}_q')|$ because $(c_{m+i}) \in \mathcal{U}_q'$ for every $(c_i) \in \mathcal{U}_q'$. This completes the proof. $\square$

Lemma 2.2.

(i) *If $q \geq M + 1$, then $h\left(\mathcal{U}_q'\right) = \log(M + 1)$.*
(ii) *If $1 < q < q'$, then $h\left(\mathcal{U}_q'\right) = 0$.*

Proof. If $q > M + 1$, then $\mathcal{U}_q' = \{0, \ldots, M\}^\infty$ is the full shift. Therefore

$$h\left(\mathcal{U}_q'\right) = \lim_{n \to \infty} \frac{\log \left|B_n(\mathcal{U}_q')\right|}{n} = \lim_{n \to \infty} \frac{\log(M+1)^n}{n} = \log(M+1).$$

If $q = M + 1$, then the above equalities remain valid. Indeed, we still have $B_n(\mathcal{U}_q') = \{0, \ldots, M\}^n$ for all $n \geq 1$ because $c_1 \cdots c_n (0M)^\infty \in \mathcal{U}_q'$ for every word $c_1 \cdots c_n \in \{0, \ldots, M\}^n$.

The case $1 < q < q'$ follows from Theorem 1.2 because $\mathcal{U}_q'$ is countable by [15] (for $M = 1$) and [8,25,24] (for all $M \geq 1$) and therefore $D(q) = 0$. $\square$

Henceforth we assume that $q' \leq q \leq M + 1$. Then $x = 1$ has an expansion.

We start by recalling some properties of the greedy and quasi-greedy expansions. We denote by $\beta(q) = (\beta_i(q))$ the *greedy*, i.e., the lexicographically largest expansion of $x = 1$ in base q. Furthermore, we denote by $\alpha(q) = (\alpha_i(q))$ the *quasi-greedy*, i.e., the lexicographically largest *infinite* expansion of $x = 1$ in base q. Here and in the sequel an expansion is called *infinite* if it contains infinitely many non-zero digits.

Greedy expansions were introduced by Rényi [29], and they were characterized by Parry [28]. Quasi-greedy expansions were introduced by Daróczy and Kátai [4,5], in order to give an elegant Parry type characterization of unique expansions:

Lemma 2.3. *A sequence (c_i) belongs to $\mathcal{U}'_q$ if and only if the following two conditions are satisfied:*

$$(c_{n+i}) < \alpha(q) \quad \text{whenever} \quad c_1 \ldots c_n \neq M^n,$$

$$\overline{(c_{n+i})} < \alpha(q) \quad \text{whenever} \quad c_1 \ldots c_n \neq 0^n.$$

Here for a sequence $c = (c_i)$ we denote by $\overline{c} = (M - c_i)$, and for a word $c_1 \cdots c_k$ we write $\overline{c_1 \cdots c_k} = (M - c_1) \cdots (M - c_k)$.

We also recall some results on the relationship between greedy and quasi-greedy expansions, and on their continuity properties:

Lemma 2.4.

(i) *If $\beta(q)$ is infinite, then $\alpha(q) = \beta(q)$. Otherwise, $\beta(q)$ has a last non-zero digit $\beta_m(q)$, and $\alpha(q)$ is periodic with the period $\beta_1(q) \cdots \beta_{m-1}(q)(\beta_m(q) - 1)$.*
(ii) *If $q_n \nearrow q$, then $\alpha(q_n) \to \alpha(q)$ component-wise.*
(iii) *If $q_n \searrow q$, then $\beta(q_n) \to \beta(q)$ component-wise.*

See, e.g., [2,8] and [9] for proofs.

Instead of $\mathcal{U}'_q$ and $\mathcal{U}_q$ it will be easier to consider the slightly modified sets

$$\widetilde{\mathcal{U}}'_q := \left\{ (c_i) \; : \; \overline{\alpha(q)} < (c_{m+i}) < \alpha(q) \quad \text{for all} \quad m = 0, 1, \ldots \right\}$$

and

$$\widetilde{\mathcal{U}}_q := \pi_q(\widetilde{\mathcal{U}}'_q) = \left\{ \sum_{i=1}^{\infty} \frac{c_i}{q^i} \; : \; (c_i) \in \widetilde{\mathcal{U}}'_q \right\}.$$

Lemma 2.5.

(i) *$\mathcal{U}_q$ is the union of 0, $M/(q-1)$, and of countably many sets, each similar to $\widetilde{\mathcal{U}}_q$.*
(ii) *$\mathcal{U}'_q$ and $\widetilde{\mathcal{U}}'_q$ have the same topological entropy.*

Proof. (i) Let $(c_i) \in \mathcal{U}'_q$ be different from 0^∞ and M^∞. If $0 < c_1 < M$, then $(c_{1+i}) \in \widetilde{\mathcal{U}}'_q$ by Lemma 2.3.

If $c_1 = 0$, then there exists a smallest $m > 1$ such that $c_m > 0$, and $(c_{m+i}) \in \widetilde{\mathcal{U}}'_q$ by Lemma 2.3.

If $c_1 = M$, then there exists a smallest $m > 1$ such that $c_m < M$, and $(c_{m+i}) \in \widetilde{\mathcal{U}}'_q$ by Lemma 2.3.

It follows that $\mathcal{U}_q$ is the union of 0, $M/(q-1)$, and of the sets

$$\frac{c_1}{q} + \frac{1}{q}\widetilde{\mathcal{U}}_q, \quad c_1 = 1, \ldots, M - 1,$$

$$\frac{c_m}{q^m} + \frac{1}{q^m}\widetilde{\mathcal{U}}_q, \quad m = 2, 3, \ldots, \quad c_m = 1, \ldots, M,$$

$$\left(\sum_{i=1}^{m-1} \frac{M}{q^i}\right) + \frac{c_m}{q^m} + \frac{1}{q^m}\widetilde{\mathcal{U}}_q, \quad m = 2, 3, \ldots, \quad c_m = 0, \ldots, M - 1.$$

We conclude by observing that all these sets are similar to $\widetilde{\mathcal{U}}_q$.

(ii) The above reasoning shows also that each word of $B_n(\mathcal{U}'_q)$ has the form $0^k M^{m-k} w$ or $M^k 0^{m-k} w$ with some word $w \in B_{n-m}(\widetilde{\mathcal{U}}'_q)$ and some integers k, m satisfying $0 \le k \le m \le n$. Hence

$$\left|B_n(\mathcal{U}'_q)\right| \le \sum_{m=0}^{n} 2(m+1)\left|B_{n-m}(\widetilde{\mathcal{U}}'_q)\right| \le (n+1)(2n+2)\left|B_n(\widetilde{\mathcal{U}}'_q)\right|.$$

Since $\widetilde{\mathcal{U}}'_q \subseteq \mathcal{U}'_q$, it follows that

$$\lim_{n\to\infty} \frac{\log\left|B_n(\widetilde{\mathcal{U}}'_q)\right|}{n} \le \lim_{n\to\infty} \frac{\log\left|B_n(\mathcal{U}'_q)\right|}{n}$$

$$\le \lim_{n\to\infty} \frac{\log(2n+2)^2\left|B_n(\widetilde{\mathcal{U}}'_q)\right|}{n}$$

$$= \lim_{n\to\infty} \frac{\log\left|B_n(\widetilde{\mathcal{U}}'_q)\right|}{n} + \lim_{n\to\infty} \frac{2\log(2n+2)}{n}$$

$$= \lim_{n\to\infty} \frac{\log\left|B_n(\widetilde{\mathcal{U}}'_q)\right|}{n},$$

whence $h(\widetilde{\mathcal{U}}'_q) = h(\mathcal{U}'_q)$. $\square$

Remark. Since $\mathcal{U}'_q$ and $\widetilde{\mathcal{U}}'_q$ are not always subshifts, it will be useful to introduce the related set $\mathcal{V}'_q$ of the sequences (c_i) satisfying the inequalities in Lemma 2.3 with "$<$" replaced by "$\le$", and the set

$$\widetilde{\mathcal{V}}'_q := \left\{(c_i) \ : \ \overline{\alpha(q)} \le (c_{m+i}) \le \alpha(q) \quad \text{for all} \quad m = 0, 1, \ldots\right\}.$$

Furthermore, we define $\mathcal{V}_q := \pi_q(\mathcal{V}'_q)$ and $\widetilde{\mathcal{V}}_q := \pi_q(\widetilde{\mathcal{V}}'_q)$.

Lemma 2.5 and its proof remain valid if $\mathcal{U}_q$, $\widetilde{\mathcal{U}}_q$, $\mathcal{U}'_q$ and $\widetilde{\mathcal{U}}'_q$ are replaced by $\mathcal{V}_q$, $\widetilde{\mathcal{V}}_q$, $\mathcal{V}'_q$ and $\widetilde{\mathcal{V}}'_q$.

Lemma 2.6. $\widetilde{\mathcal{V}}_q'$ *is a subshift, and* $\widetilde{\mathcal{U}}_q' \subseteq \widetilde{\mathcal{V}}_q'$.

Proof. If $q = M + 1$, then $\alpha(q) = M^\infty$, so that $\widetilde{\mathcal{V}}_q' = \{0, 1, \cdots, M\}^\infty$ is the full shift.

Henceforth assume that $q < M+1$, and consider the set F of all finite blocks $d_1 \cdots d_n \in \{0, \ldots, M\}^n$ (of arbitrary length), satisfying one of the lexicographic inequalities

$$d_1 \cdots d_n < \overline{\alpha_1(q) \cdots \alpha_n(q)} \quad \text{and} \quad d_1 \cdots d_n > \alpha_1(q) \cdots \alpha_n(q).$$

By definition, none of these blocks appear in any $(c_i) \in \widetilde{\mathcal{V}}_q'$.

Conversely, if $(c_i) \in \{0, 1, \cdots, M\}^\infty \setminus \widetilde{\mathcal{V}}_q'$, then there is a positive integer m such that either

$$c_m c_{m+1} \cdots < \overline{\alpha(q)}$$

or

$$c_m c_{m+1} \cdots > \alpha(q),$$

and hence there is another positive integer n such that either

$$c_m \cdots c_{m+n} < \overline{\alpha_1(q) \cdots \alpha_n(q)}$$

or

$$c_m \cdots c_{m+n} > \alpha_1(q) \cdots \alpha_n(q).$$

Hence (c_i) contains at least one block from F.

The inclusion $\widetilde{\mathcal{U}}_q' \subseteq \widetilde{\mathcal{V}}_q'$ is obvious from the definition.   □

Since $\widetilde{\mathcal{U}}_q'$ is not always a subshift of finite type, we introduce for each positive integer n the set $\widetilde{\mathcal{U}}_{q,n}'$ of sequences (c_i) satisfying for all $m = 0, 1, \ldots$ the inequalities

$$\overline{\alpha_1(q) \cdots \alpha_n(q)} < c_{m+1} \cdots c_{m+n} < \alpha_1(q) \cdots \alpha_n(q).$$

Similarly, we define the sets $\widetilde{\mathcal{V}}_{q,n}'$ and $\widetilde{\mathcal{W}}_{q,n}'$ by replacing the above inequalities by

$$\overline{\alpha_1(q) \cdots \alpha_n(q)} \leq c_{m+1} \cdots c_{m+n} \leq \alpha_1(q) \cdots \alpha_n(q)$$

and

$$\overline{\beta_1(q) \cdots \beta_n(q)} \leq c_{m+1} \cdots c_{m+n} \leq \beta_1(q) \cdots \beta_n(q),$$

respectively.

Lemma 2.7. $\widetilde{\mathcal{U}}'_{q,n}$, $\widetilde{\mathcal{V}}'_{q,n}$ and $\widetilde{\mathcal{W}}'_{q,n}$ *are subshifts of finite type, and*

$$\widetilde{\mathcal{U}}'_{q,n} \subseteq \widetilde{\mathcal{U}}'_q \subseteq \widetilde{\mathcal{V}}'_q \subseteq \widetilde{\mathcal{V}}'_{q,n} \subseteq \widetilde{\mathcal{W}}'_{q,n} \tag{2.1}$$

for all n.

Furthermore, the sets $\widetilde{\mathcal{U}}'_{q,n}$ are increasing, while $\widetilde{\mathcal{V}}'_{q,n}$ and $\widetilde{\mathcal{W}}'_{q,n}$ are decreasing when n is increasing.

Proof. It is clear that $\widetilde{\mathcal{U}}'_{q,n}$ is characterized by the finite set of forbidden blocks $d_1 \cdots d_n \in \{0, \ldots, M\}^n$ satisfying the lexicographic inequalities

$$d_1 \cdots d_n \leq \overline{\alpha_1(q) \cdots \alpha_n(q)} \quad \text{or} \quad d_1 \cdots d_n \geq \alpha_1(q) \cdots \alpha_n(q).$$

Hence it is a subshift of finite type.

The proof for $\widetilde{\mathcal{V}}'_{q,n}$ and $\widetilde{\mathcal{W}}'_{q,n}$ is analogous.

The remaining assertions follow from the definition of lexicographic inequalities. $\quad\square$

We are going to show that these sets well approximate $\widetilde{\mathcal{U}}'_q$:

Proposition 2.8. *For* $q \in [q', M+1]$ *we have*

$$\lim_{n\to\infty} h(\widetilde{\mathcal{U}}'_{q,n}) = \lim_{n\to\infty} h(\widetilde{\mathcal{V}}'_{q,n}) = \lim_{n\to\infty} h(\widetilde{\mathcal{W}}'_{q,n}) = h(\widetilde{\mathcal{U}}'_q) = h(\widetilde{\mathcal{V}}'_q).$$

The proof of the proposition is divided into a series of lemmas.

Lemma 2.9. *Let* $q' \leq q < p \leq M+1$. *Then*

$$\widetilde{\mathcal{W}}'_{q,n} \subseteq \widetilde{\mathcal{U}}'_{p,n}$$

for all sufficiently large n.

Proof. Since there are only countably many finite greedy expansions, the set

$$\{r \in (1, M+1] \ : \ \beta(r) \neq \alpha(r)\}$$

is countable. There exists therefore $r \in (q, p)$ such that $\beta(r) = \alpha(r)$, and then

$$\beta(q) < \beta(r) = \alpha(r) < \alpha(p)$$

because the maps $r \mapsto \beta(r)$ and $r \mapsto \alpha(r)$ are strictly increasing by the definition of the greedy and quasi-greedy algorithms.

Fix a sufficiently large n such that

$$\alpha_1(p) \cdots \alpha_n(p) > \beta_1(q) \cdots \beta_n(q).$$

If $d = (d_i) \in \widetilde{\mathcal{W}}'_{q,n}$, then

$$d_{m+1} \cdots d_{m+n} \le \beta_1(q) \cdots \beta_n(q) < \alpha_1(p) \cdots \alpha_n(p)$$

and symmetrically

$$d_{m+1} \cdots d_{m+n} \ge \overline{\beta_1(q) \cdots \beta_n(q)} > \overline{\alpha_1(p) \cdots \alpha_n(p)}$$

for all $m \ge 0$, i.e., $d \in \widetilde{\mathcal{U}}'_{p,n}$.　$\square$

We recall that $\mathcal{U}$ is the set of bases $q > 1$ in which $x = 1$ has a unique expansion, and $\overline{\mathcal{U}}$ is its closure. Furthermore, we recall from [23] that $q \in \overline{\mathcal{U}}$ if and only if

$$\overline{\alpha_1(q)\alpha_2(q)\cdots} < \alpha_{k+1}(q)\alpha_{k+2}(q)\cdots \le \alpha_1(q)\alpha_2(q)\cdots \tag{2.2}$$

for all $k \ge 0$. Moreover, there exists infinitely many indices n such that

$$\overline{\alpha_1(q)\cdots\alpha_{n-k}(q)} < \alpha_{k+1}(q)\cdots\alpha_n(q) \le \alpha_1(q)\cdots\alpha_{n-k}(q) \tag{2.3}$$

for all $0 \le k \le n-1$. In particular, $\alpha_n(q) > 0$ for these indices.

Lemma 2.10. *Let $q \in \overline{\mathcal{U}}$ and $(\alpha_i) = \alpha(q)$.*

(i) *For each $n \ge 1$, $B_n(\widetilde{\mathcal{V}}'_q) = B_n(\widetilde{\mathcal{V}}'_{q,n})$ is the set of words $d_1 \cdots d_n$ satisfying*

$$\overline{\alpha_1 \cdots \alpha_{n-k}} \le d_{k+1} \cdots d_n \le \alpha_1 \cdots \alpha_{n-k} \tag{2.4}$$

for all $0 \le k \le n-1$.

(ii) *For each $n \ge 1$ satisfying (2.3), $B_n(\widetilde{\mathcal{U}}'_{q,n})$ is the set of words $d_1 \cdots d_n$ satisfying*

$$\overline{\alpha_1 \cdots \alpha_n} < d_1 \cdots d_n < \alpha_1 \cdots \alpha_n, \tag{2.5}$$

and relations (2.4) for all $1 \le k \le n-1$.

(iii) *If $n \ge 1$ satisfies (2.3), then*

$$B_n(\widetilde{\mathcal{V}}'_{q,n}) \setminus B_n(\widetilde{\mathcal{U}}'_{q,n}) = \left\{ \alpha_1(q) \ldots \alpha_n(q), \overline{\alpha_1(q) \ldots \alpha_n(q)} \right\}.$$

Proof. (i) Note that $B_n(\widetilde{\mathcal{V}}'_q) \subseteq B_n(\widetilde{\mathcal{V}}'_{q,n})$, and that each word of $B_n(\widetilde{\mathcal{V}}'_{q,n})$ satisfies the relations (2.4). It remains to prove that if a word $d_1 \cdots d_n$ satisfies the relations (2.4) for all $0 \le k \le n-1$, then it belongs to $B_n(\widetilde{\mathcal{V}}'_q)$.

Let $0 \le k_1 \le n$ be the first integer such that either

$$d_{k_1+1} \cdots d_n = \alpha_1 \cdots \alpha_{n-k_1}$$

or

$$d_{k_1+1} \cdots d_n = \overline{\alpha_1 \cdots \alpha_{n-k_1}}.$$

(We choose $k_1 = n$ if all inequalities in (2.4) are strict.) Assume by symmetry that

$$d_{k_1+1} \cdots d_n = \alpha_1 \cdots \alpha_{n-k_1}. \qquad (2.6)$$

The minimality of k_1 implies that

$$\overline{\alpha_1 \cdots \alpha_{n-k}} < d_{k+1} \cdots d_n < \alpha_1 \cdots \alpha_{n-k} \quad \text{for any } 0 \le k < k_1.$$

Combining this with (2.2) we conclude that

$$d_1 \cdots d_n \alpha_{n-k_1+1} \alpha_{n-k_1+2} \cdots = d_1 \cdots d_{k_1} \alpha_1 \alpha_2 \cdots \in \widetilde{\mathcal{V}}_q';$$

hence $d_1 \cdots d_n \in B_n(\widetilde{\mathcal{V}}_q')$.

(ii) Take n satisfying (2.3), and note that each word of $B_n(\widetilde{\mathcal{U}}_{q,n}')$ satisfies the above mentioned relations. It remains to prove that if a word $d_1 \cdots d_n$ satisfies (2.5), and relations (2.4) for all $1 \le k \le n-1$, then it belongs to $B_n(\widetilde{\mathcal{U}}_{q,n}')$.

Choosing k_1 as in (i), now we have $k_1 \ge 1$. We may assume (2.6) again. Using (2.3) it follows that $\alpha_n > 0$ and

$$\alpha_{k+1} \cdots \alpha_{n-1} \alpha_n^- \ge \overline{\alpha_1 \cdots \alpha_{n-k}} \quad \text{and} \quad \alpha_1 \cdots \alpha_k > \overline{\alpha_{n-k+1} \cdots \alpha_n}$$

for all $0 \le k \le n-1$ and $1 \le k \le n$, respectively, where we write $\alpha_n^- := \alpha_n - 1$. Hence

$$d_1 \cdots d_n (\alpha_{n-k_1+1} \cdots \alpha_{n-1} \alpha_n^- \alpha_1 \cdots \alpha_{n-k_1})^\infty = d_1 \cdots d_{k_1} (\alpha_1 \cdots \alpha_{n-1} \alpha_n^-)^\infty \in \widetilde{\mathcal{U}}_{q,n}',$$

and therefore $d_1 \cdots d_n \in B_n(\widetilde{\mathcal{U}}_{q,n}')$.

(iii) This follows from (i), (ii) and (2.2). $\quad\square$

The following lemma was essentially proved during the proof of [24, Theorem 2.6]. We briefly indicate an alternative proof based on some results of [8]. The set $\overline{\mathcal{U}}$ was defined in the introduction.

Lemma 2.11. *If p and q belong to the same connected component of $(1, \infty) \setminus \overline{\mathcal{U}}$, then $h(\mathcal{U}_p') = h(\mathcal{U}_q')$ and $h(\widetilde{\mathcal{U}}_p') = h(\widetilde{\mathcal{U}}_q')$.*

Proof. The two equalities being equivalent by Lemma 2.5, we only prove the first one.

Consider a connected component $I = (q_0, q_0^*)$. By [8, Theorems 1.6 and 1.7] there exists a sequence (q_n) satisfying $q_0 < q_1 < \cdots$ and converging to q_0^*, and such that $\mathcal{U}_q' = \mathcal{V}_{q_n}'$ for all $q \in (q_n, q_{n+1}]$, $n = 0, 1, \ldots$. (See the definition of $\mathcal{V}_q'$ in the remark

following Lemma 2.5.) This implies that $h(\mathcal{U}'_q)$ is constant in each interval $(q_n, q_{n+1}]$. It remains to show that $h(\mathcal{U}'_{q_n}) = h(\mathcal{U}'_{q_{n+1}})$, or equivalently that $h(\mathcal{U}'_{q_n}) = h(\mathcal{V}'_{q_n})$ for each $n \geq 1$.

Assume for the moment that $\widetilde{\mathcal{U}}'_{q_n} = \widetilde{\mathcal{U}}'_{q_n,k}$ and $\widetilde{\mathcal{V}}'_{q_n} = \widetilde{\mathcal{V}}'_{q_n,2k}$ for some positive integer k. Then by Lemma 2.7 they hold for all larger values of k as well. Applying Lemma 3.2 below (its proof will not depend on the present section) with $p = q_n$ and a sufficiently large k in place of $n(q_n)$, it follows that

$$\dim_H(\widetilde{\mathcal{U}}_{q_n}) = \frac{h(\widetilde{\mathcal{U}}'_{q_n})}{\log q_n} \quad \text{and} \quad \dim_H(\widetilde{\mathcal{V}}_{q_n}) = \frac{h(\widetilde{\mathcal{V}}'_{q_n})}{\log q_n}.$$

Applying Lemma 2.5 and the following remark they yield

$$\dim_H(\mathcal{U}_{q_n}) = \frac{h(\mathcal{U}'_{q_n})}{\log q_n} \quad \text{and} \quad \dim_H(\mathcal{V}_{q_n}) = \frac{h(\mathcal{V}'_{q_n})}{\log q_n}.$$

The equality $h(\mathcal{U}'_{q_n}) = h(\mathcal{V}'_{q_n})$ follows because $\mathcal{U}_{q_n}$ and $\mathcal{V}_{q_n}$ differ by a countable set by [8, Theorem 1.4 (ii)], and hence they have the same Hausdorff dimension.

Now we prove the equalities $\widetilde{\mathcal{U}}'_{q_n} = \widetilde{\mathcal{U}}'_{q_n,k}$ and $\widetilde{\mathcal{V}}'_{q_n} = \widetilde{\mathcal{V}}'_{q_n,2k}$. We have $\alpha(q_n) = (\alpha_1 \cdots \alpha_k \overline{\alpha_1 \cdots \alpha_k})^\infty$ with a minimal even period $2k$, and $\alpha_k > 0$. The proof of [8, Theorem 1.8] shows that $(c_i) \in \widetilde{\mathcal{U}}'_{q_n}$ is equivalent to

$$\overline{\alpha_1 \cdots \alpha_k} < c_{m+1} \cdots c_{m+k} < \alpha_1 \cdots \alpha_k \quad \text{for all} \quad m \geq 0,$$

i.e., to $(c_i) \in \widetilde{\mathcal{U}}'_{q_n,k}$.

Next we recall that writing $\alpha_1 \cdots \alpha_k^- := \alpha_1 \cdots \alpha_{k-1}(\alpha_k - 1)$ we have

$$\alpha(q_{n+1}) = (\alpha_1 \cdots \alpha_k \overline{\alpha_1 \cdots \alpha_k^-} \, \overline{\alpha_1 \cdots \alpha_k} \alpha_1 \cdots \alpha_k^-)^\infty$$

with a minimal even period $4k$. Since $\widetilde{\mathcal{V}}'_{q_n} = \widetilde{\mathcal{U}}'_{q_{n+1}}$, hence $(c_i) \in \widetilde{\mathcal{V}}'_{q_n}$ is equivalent to

$$\overline{\alpha_1 \cdots \alpha_k} \alpha_1 \cdots \alpha_k^- < c_{m+1} \cdots c_{m+2k} < \alpha_1 \cdots \alpha_k \overline{\alpha_1 \cdots \alpha_k^-} \quad \text{for all} \quad m \geq 0,$$

or equivalently to

$$\overline{\alpha_1 \cdots \alpha_k} \alpha_1 \cdots \alpha_k \leq c_{m+1} \cdots c_{m+2k} \leq \alpha_1 \cdots \alpha_k \overline{\alpha_1 \cdots \alpha_k} \quad \text{for all} \quad m \geq 0.$$

The last relation is the definition of $(c_i) \in \widetilde{\mathcal{V}}'_{q_n,2k}$. $\square$

Finally we recall the Perron–Frobenius Theorem (see [26, Theorem 4.4.4]):

Lemma 2.12. *Let $G(n)$ be an edge graph representation of $\widetilde{\mathcal{U}}'_{q,n}$, and λ_n its spectral radius. Then there exist positive constants c_1, c_2 such that*

$$c_1 \lambda_n^k \leq |B_k(\widetilde{\mathcal{U}}'_{q,n})| \leq c_2 k^s \lambda_n^k$$

for all $k \geq 1$, where s denotes the number of strongly connected components of $G(n)$.

If $G(n)$ is strongly connected, then the factor k^s may be omitted in the second inequality.

Proof of Proposition 2.8. All indicated topological entropies are well defined by Lemmas 2.6 and 2.7. Furthermore, the monotonicity of the set sequences $(\widetilde{\mathcal{U}}'_{q,n})$, $(\widetilde{\mathcal{V}}'_{q,n})$ and $(\widetilde{\mathcal{W}}'_{q,n})$ implies the existence of the indicated limits as $n \to \infty$.

If $q \in [q', M+1] \backslash \overline{\mathcal{U}}$, then $q \in (q', M+1)$ (because $q', M+1 \in \mathcal{U}$). Applying Lemma 2.11 we may choose a neighborhood (q_1, q_2) of q such that $h(\widetilde{\mathcal{U}}'_p) = h(\widetilde{\mathcal{U}}'_q)$ for all $p \in [q_1, q_2]$. Using Lemmas 2.7 and 2.9 we obtain that

$$\widetilde{\mathcal{U}}'_{q_1} \subseteq \widetilde{\mathcal{U}}'_{q,n} \subseteq \widetilde{\mathcal{W}}'_{q,n} \subseteq \widetilde{\mathcal{U}}'_{q_2}$$

for all sufficiently large indices n, and therefore

$$\lim_{n \to \infty} h(\widetilde{\mathcal{U}}'_{q,n}) = \lim_{n \to \infty} h(\widetilde{\mathcal{W}}'_{q,n}) = h(\widetilde{\mathcal{U}}'_q).$$

Henceforth we assume that $q \in \overline{\mathcal{U}}$. In view of the inclusions (2.1) it is sufficient to prove that

$$\lim_{n \to \infty} h(\widetilde{\mathcal{W}}'_{q,n}) \leq h(\widetilde{\mathcal{V}}'_q) \tag{2.7}$$

and

$$\lim_{n \to \infty} h(\widetilde{\mathcal{V}}'_{q,n}) \leq \lim_{n \to \infty} h(\widetilde{\mathcal{U}}'_{q,n}). \tag{2.8}$$

First we show that

$$|B_n(\widetilde{\mathcal{W}}'_{q,n})| \leq 2(n+1)^2 |B_n(\widetilde{\mathcal{V}}'_q)|$$

for all $n \geq 1$. If $\alpha(q) = \beta(q)$, then $\widetilde{\mathcal{W}}'_{q,n} = \widetilde{\mathcal{V}}'_{q,n}$ and therefore by Lemma 2.10 we have $B_n(\widetilde{\mathcal{W}}'_{q,n}) = B_n(\widetilde{\mathcal{V}}'_q)$ for all n.

If $\alpha(q) \neq \beta(q)$, then $\beta(q)$ has a last nonzero digit β_m, and by Lemma 2.4 $\alpha(q)$ is periodic with the period $\beta_1(q) \cdots \beta_{m-1}(q)\beta_m^-(q)$. In this case, if $d_1 \cdots d_n \in B_n(\widetilde{\mathcal{W}}'_{q,n}) \backslash B_n(\widetilde{\mathcal{V}}'_q)$, then for any $0 \leq k \leq n-1$

$$\overline{\beta_1(q) \cdots \beta_{n-k}(q)} \leq d_{k+1} \cdots d_n \leq \beta_1(q) \cdots \beta_{n-k}(q),$$

and by Lemma 2.10 it follows that there exists a least integer $0 \leq k \leq n-1$ such that either

$$d_{k+1}\cdots d_n < \overline{\alpha_1(q)\cdots\alpha_{n-k}(q)}$$

or

$$d_{k+1}\cdots d_n > \alpha_1(q)\cdots\alpha_{n-k}(q).$$

This implies that $d_{k+1}\cdots d_n$ or $\overline{d_{k+1}\cdots d_n}$ must be of the form

$$(\alpha_1(q)\cdots\alpha_m(q))^j\beta_1(q)\cdots\beta_{n-k-mj}(q), \quad j = 0, 1, \cdots, [(n-k)/m].$$

The number of these words can not exceed $2(n+1)$. Moreover, by the minimality of k and Lemmas 2.7, 2.10 it follows that

$$d_1\cdots d_k \in B_k(\widetilde{\mathcal{V}}'_{q,k}) = B_k(\widetilde{\mathcal{V}}'_q) = B_k(\widetilde{\mathcal{V}}'_{q,n}).$$

Hence

$$|B_n(\widetilde{\mathcal{W}}'_{q,n})| - |B_n(\widetilde{\mathcal{V}}'_{q,n})| = |B_n(\widetilde{\mathcal{W}}'_{q,n}) \setminus B_n(\widetilde{\mathcal{V}}'_{q,n})|$$

$$\leq 2(n+1)\sum_{k=0}^{n-1}|B_k(\widetilde{\mathcal{V}}'_{q,n})| \leq 2n(n+1)|B_n(\widetilde{\mathcal{V}}'_{q,n})|,$$

and the required estimate follows.

Using this estimate we have

$$h(\widetilde{\mathcal{W}}'_{q,n}) = \inf_{k\geq 1}\frac{\log|B_k(\widetilde{\mathcal{W}}'_{q,n})|}{k} \leq \frac{\log|B_n(\widetilde{\mathcal{W}}'_{q,n})|}{n}$$

$$\leq \frac{\log|B_n(\widetilde{\mathcal{V}}'_q)| + \log 2 + 2\log(n+1)}{n}.$$

Letting $n \to \infty$ the relation (2.7) follows.

Turning to the proof of the relation (2.8), first we consider the case $q = q'$. Using (2.1) and (2.7) it follows that

$$\lim_{n\to\infty}h(\widetilde{\mathcal{V}}'_{q,n}) = h(\widetilde{\mathcal{V}}'_q).$$

Furthermore, we also deduce from (2.1) and Lemma 2.10 that

$$|B_{n_k}(\widetilde{\mathcal{V}}'_q) \setminus B_{n_k}(\widetilde{\mathcal{U}}'_q)| \leq |B_{n_k}(\widetilde{\mathcal{V}}_{q,n_k}) \setminus B_{n_k}(\widetilde{\mathcal{U}}'_{q,n_k})| = 2,$$

where (n_k) is a sequence of indices satisfying (2.3). Hence

$$h(\widetilde{\mathcal{V}}_q') = \lim_{n \to \infty} \frac{\log |B_n(\widetilde{\mathcal{V}}_q')|}{n} = \lim_{k \to \infty} \frac{\log |B_{n_k}(\widetilde{\mathcal{V}}_q')|}{n_k}$$

$$\leq \lim_{k \to \infty} \frac{\log \left(|B_{n_k}(\widetilde{\mathcal{U}}_q')| + 2 \right)}{n_k} = h(\widetilde{\mathcal{U}}_q').$$

The existence of the last limit and the last equality follows from Lemma 2.1.

Since $h(\widetilde{\mathcal{U}}_q') = 0$ for $q = q'$ by Theorem 1.1, we conclude that

$$\lim_{n \to \infty} h(\widetilde{\mathcal{V}}_{q,n}') = 0.$$

Assume henceforth that $q > q'$, so that $h(\widetilde{\mathcal{U}}_q') > 0$. This was proved in [15] for $M = 1$, and the proof remains valid for all odd values of M, and in [25, Lemma 4.10] for $M = 2, 4, \ldots$. For each $n \geq N$ we have $h(\widetilde{\mathcal{U}}_{q,n}') = \log \lambda_n$ with the notations of Lemma 2.12, and

$$\lambda_n \geq \lambda_N > 1$$

by the increasing property of the set sequence $(\widetilde{\mathcal{U}}_{q,n}')$. We are going to estimate the size of $B_k(\widetilde{\mathcal{V}}_{q,n}') \setminus B_k(\widetilde{\mathcal{U}}_{q,n}')$ for each fixed $n \geq N$ satisfying (2.3) and $k \geq n$.

Let us denote by $G'(n)$ the edge graph representing $\widetilde{\mathcal{V}}_{q,n}'$, and set $u = \alpha_1(q) \cdots \alpha_n(q)$. Then $G(n)$ is a subgraph of $G'(n)$, and the words u and $\overline{u}$ are forbidden in $G(n)$. We seek an upper bound for $|B_k(\widetilde{\mathcal{V}}_{q,n}') \setminus B_k(\widetilde{\mathcal{U}}_{q,n}')|$.

Suppose that $d_1 \cdots d_k \in B_k(\widetilde{\mathcal{V}}_{q,n}') \setminus B_k(\widetilde{\mathcal{U}}_{q,n}')$. Then by Lemma 2.10 it follows that the word $d_1 \cdots d_k$ must contain at least once u or $\overline{u}$. If it contains exactly $r \geq 1$ times u or $\overline{u}$, then it has the form

$$d_1 \cdots d_k = \omega_0 \tau_1 \omega_1 \cdots \tau_r \omega_r$$

where each τ_j is equal to u or $\overline{u}$, and $k_0 + \cdots + k_r = k - rn$, where $k_j \geq 0$ denotes the length of ω_j.

Assuming first that the graph $G(n)$ is strongly connected, we may apply Lemma 2.12 without the factor k^s. Assuming without loss of generality that $c_1 \leq 1 \leq c_2$, we obtain the following estimate:

$$|B_k(\widetilde{\mathcal{V}}_{q,n}')| \leq |B_k(\widetilde{\mathcal{U}}_{q,n}')| + \sum_{r=1}^{[k/n]} \sum_{k_0 + \cdots + k_r = k - nr} 2^r \prod_{j=0}^{r} (c_2 \lambda_n^{k_j})$$

$$= |B_k(\widetilde{\mathcal{U}}_{q,n}')| + c_2 \lambda_n^k \sum_{r=1}^{[k/n]} \sum_{k_0 + \cdots + k_r = k - nr} (2 c_2 \lambda_n^{-n})^r$$

$$= |B_k(\widetilde{\mathcal{U}}_{q,n}')| + c_2 \lambda_n^k \sum_{r=1}^{[k/n]} \binom{k - r(n-1)}{r} (2 c_2 \lambda_n^{-n})^r$$

$$\leq |B_k(\widetilde{\mathcal{U}}'_{q,n})| + c_2 \lambda_n^k \sum_{r=1}^{k} \binom{k}{r} (2c_2 \lambda_n^{-n})^r$$

$$\leq |B_k(\widetilde{\mathcal{U}}'_{q,n})| \frac{c_2}{c_1} \sum_{r=0}^{k} \binom{k}{r} (2c_2 \lambda_n^{-n})^r$$

$$= |B_k(\widetilde{\mathcal{U}}'_{q,n})| \frac{c_2}{c_1} (1 + 2c_2 \lambda_n^{-n})^k$$

$$\leq |B_k(\widetilde{\mathcal{U}}'_{q,n})| \frac{c_2}{c_1} (1 + 2c_2 \lambda_N^{-n})^k.$$

If the graph $G(n)$ is not strongly connected, then we distinguish two cases:

- If u and $\overline{u}$ belong to the same strongly connected component of $G'(n)$, then we have to change $c_2 \lambda_n^{k_j}$ to $c_2 k_j^s \lambda_n^{k_j}$ in the above estimate for $j = 0$ and $j = r$.
- If u and $\overline{u}$ belong to different strongly connected components of $G'(n)$, then for each $d_1 \cdots d_k$ there is an index $0 \leq r' \leq r$ such that either $\tau_j = u \iff j \leq r'$ or $\tau_j = u \iff j > r'$. Then we may change the above factor 2^r to $r + 1$, and we have to change $c_2 \lambda_n^{k_j}$ to $c_2 k_j^s \lambda_n^{k_j}$ for $j = 0$, $j = r'$ and $j = r$.

Summarizing, we obtain in all cases the following estimate:

$$|B_k(\widetilde{\mathcal{V}}'_{q,n})| \leq |B_k(\widetilde{\mathcal{U}}'_{q,n})| \frac{c_2}{c_1} k^{3s} (1 + 2c_2 \lambda_N^{-n})^k.$$

It follows that

$$\frac{\log |B_k(\widetilde{\mathcal{V}}'_{q,n})|}{k} \leq \frac{\log |B_k(\widetilde{\mathcal{U}}'_{q,n})|}{k} + \frac{\log(c_2/c_1)}{k} + 3s \frac{\log k}{k} + \log(1 + 2c_2 \lambda_N^{-n})$$

for all $k \geq n$. Letting $k \to \infty$ we conclude that

$$h(\widetilde{\mathcal{V}}'_{q,n}) \leq h(\widetilde{\mathcal{U}}'_{q,n}) + \log(1 + 2c_2 \lambda_N^{-n})$$

for all $n \geq N$ satisfying (2.3). Since $\lambda_N > 1$, taking n satisfying (2.3) and letting $n \to \infty$ we get (2.8). $\quad \square$

3. Proof of Theorem 1.3

First we consider the cases $1 < q < q'$ and $q \geq M + 1$.

Lemma 3.1.

(i) *The formula* (1.1) *holds for* $1 < q < q'$ *with* $D(q) = h(\mathcal{U}'_q) = 0$.
(ii) *The formula* (1.1) *holds for all* $q \geq M + 1$ *with* $h(\mathcal{U}'_q) = \log(M + 1)$.

Proof. (i) We have shown in Lemma 2.2 that $h(\mathcal{U}'_q) = 0$. Since $\mathcal{U}_q$ is countable (see the proof of Lemma 2.2), we have also $D(q) = 0$.

(ii) We have shown in Lemma 2.2 that $h(\mathcal{U}'_q) = \log(M + 1)$.

Since $[0, 1] \setminus \mathcal{U}_{M+1}$ and $\{0, \ldots, M\}^{\infty} \setminus \mathcal{U}'_{M+1}$ are countable, we have $D(M + 1) = 1$ and $h\left(\mathcal{U}'_{M+1}\right) = \log(M + 1)$.

If $q > M + 1$, then $\mathcal{U}'_q = \{0, \ldots, M\}^{\infty}$, so that $h\left(\mathcal{U}'_q\right) = \log(M + 1)$, and $\mathcal{U}_q$ is a self-similar set satisfying the relation

$$\mathcal{U}_q = \bigcup_{j=0}^{M} \left(\frac{j}{q} + \frac{1}{q}\mathcal{U}_q \right).$$

The union is disjoint because each $x \in \mathcal{U}_q$ has a unique expansion.

Observe that $\mathcal{U}_q$ is a non-empty *compact* set. Indeed, it is bounded because $\mathcal{U}_q \subseteq [0, M/(q - 1)]$. It remains to show that it is closed, i.e., if $(x_k) \subset \mathcal{U}_q$ converges to some real number x, then $x \in \mathcal{U}_q$.

If two expansions (a_i) and (b_i) first differ at the mth position, then

$$\left| \sum_{i=1}^{\infty} \frac{a_i}{q^i} - \sum_{i=1}^{\infty} \frac{b_i}{q^i} \right| \geq \frac{1}{q^m} - \sum_{i=m+1}^{\infty} \frac{M}{q^i} = \frac{q - M - 1}{q^m(q - 1)} > 0.$$

Using this estimate we obtain that the expansion of x_k converges component-wise to some sequence (c_i), and that (c_i) is the (necessarily unique) expansion of x.

Applying [16] (see also [14, Proposition 9.7]) we conclude that $r := D(q)$ is the solution of the equation $(M + 1)q^{-r} = 1$, yielding

$$D(q) = \frac{\log(M + 1)}{\log q}. \quad \square$$

In view of Theorem 1.1 and Lemma 3.1 it remains to investigate the dimension function

$$D(q) = \dim_H \mathcal{U}_q = \dim_H \widetilde{\mathcal{U}}_q$$

for $q' \leq q \leq M + 1$. (The second equality follows from Lemma 3.1 (i).)

Lemma 3.2. *Let* $q \in [q', M + 1)$. *There exists a positive integer* $n(q)$ *and a real number* $\varepsilon(q) > 0$ *such that*

$$\dim_H \pi_p(\widetilde{\mathcal{U}}'_{q,n}) = \frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log p} \quad and \quad \dim_H \pi_p(\widetilde{\mathcal{V}}'_{q,n}) = \frac{h(\widetilde{\mathcal{V}}'_{q,n})}{\log p}$$

for all $n \geq n(q)$ *and* $p \in (q - \varepsilon(q), q]$.

Proof. The two cases being similar, we consider only that of $\widetilde{\mathcal{V}}'_{q,n}$.

Fix two positive integers $n > N$ satisfying $\alpha_N(q) < M$ and $q^{n-N}(q-1) > M$. Let $p \in (q', q]$ be sufficiently close to q such that

$$p^{n-N}(p-1) > M \quad \text{and} \quad \alpha_i(p) = \alpha_i(q), \quad i = 1, \ldots, n.$$

We already know that $\widetilde{\mathcal{V}}'_{q,n}$ is a subshift of finite type corresponding to the finite set F_n of forbidden blocks $d_1 \cdots d_n \in \{0, \ldots, M\}^n$ satisfying one of the lexicographic inequalities

$$d_1 \cdots d_n < \overline{\alpha_1(q) \cdots \alpha_n(q)} \quad \text{and} \quad d_1 \cdots d_n > \alpha_1(q) \cdots \alpha_n(q).$$

We are going to prove that $\pi_p(\widetilde{\mathcal{V}}'_{q,n})$ is a graph-directed set satisfying the strong separation condition: then we may conclude by using the results of Mauldin and Williams [27]. We argue similarly to [24, Lemma 6.4].

Let us denote by $\mathbf{G} = (G, V, E)$ the edge graph with the vertex set

$$V := B_{n-1}(\widetilde{\mathcal{V}}'_{q,n}) = \left\{ d_1 \cdots d_{n-1} \in \{0, \ldots, M\}^{n-1} : (d_i) \in \widetilde{\mathcal{V}}'_{q,n} \right\}.$$

For two vertices $\mathbf{u} = u_1 \cdots u_{n-1}$ and $\mathbf{v} = v_1 \cdots v_{n-1}$ we draw an edge $\mathbf{uv} \in E$ from $\mathbf{u}$ to $\mathbf{v}$ and label it $\ell_{\mathbf{uv}} = u_1$ if

$$u_2 \cdots u_{n-1} = v_1 \cdots v_{n-2} \quad \text{and} \quad u_1 \cdots u_{n-1} v_{n-1} \notin F_n.$$

Then the edge graph $\mathbf{G} = (G, V, E)$ is a representation of $\widetilde{\mathcal{V}}'_{q,n}$ (see [26]).

For $\mathbf{u} = u_1 \cdots u_{n-1} \in V$ we set

$$K_{\mathbf{u}} := \left\{ \sum_{i=1}^{\infty} \frac{d_i}{p^i} : d_1 \cdots d_{n-1} = u_1 \cdots u_{n-1}, \text{ and } d_{m+1} \cdots d_{m+n} \notin F_n \text{ for all } m \geq 0 \right\}.$$

For each edge $\mathbf{uv} \in E$ with vertices

$$\mathbf{u} = u_1 \cdots u_{n-1}, \quad \mathbf{v} = v_1 \cdots v_{n-1}$$

we define

$$f_{\mathbf{uv}}(x) := \frac{x + \ell_{uv}}{p} = \frac{x + u_1}{p}.$$

Then

$$\pi_p(\widetilde{\mathcal{V}}'_{q,n}) = \bigcup_{\mathbf{u} \in V} K_{\mathbf{u}} = \bigcup_{\mathbf{u} \in V} \bigcup_{\mathbf{uv} \in E} f_{\mathbf{uv}}(K_{\mathbf{v}}),$$

so that $\pi_p(\widetilde{\mathcal{V}}'_{q,n})$ is a graph-directed set (see [27]).

Denote by A the adjacency matrix corresponding to the edge graph $\mathbf{G} = (G, V, E)$. Applying [26, Theorem 4.4.4] we have

$$h(\widetilde{\mathcal{V}}'_{q,n}) = \log r(A), \tag{3.1}$$

where $r(A)$ is the spectral radius of A.

Assume for the moment that the graph-directed iterated function system $\{f_{\mathbf{uv}}(\cdot) : \mathbf{uv} \in E\}$ satisfies the strong separation condition. Then applying [27, Theorem 4] we obtain that $s = \dim_H \pi_p(\widetilde{\mathcal{V}}'_{q,n})$ satisfies the equality

$$p^{-s} r(A) = r(p^{-s} A) = 1.$$

Taking logarithm and using (3.1) we conclude that

$$\dim_H \pi_p(\widetilde{\mathcal{V}}'_{q,n}) = s = \frac{\log r(A)}{\log p} = \frac{h(\widetilde{\mathcal{V}}'_{q,n})}{\log p}.$$

It remains to verify the strong separation condition, i.e., the relation

$$f_{\mathbf{uv}}(K_{\mathbf{v}}) \cap f_{\mathbf{uv}'}(K_{\mathbf{v}'}) = \emptyset$$

for all $\mathbf{uv}, \mathbf{uv}' \in E$ with $\mathbf{v} \neq \mathbf{v}'$.

Consider two such edges, and write

$$\mathbf{u} = u_1 \cdots u_{n-1}, \quad \mathbf{v} = v_1 \cdots v_{n-1} \quad \text{and} \quad \mathbf{v}' = v'_1 \cdots v'_{n-1}.$$

Then

$$v_1 \cdots v_{n-2} = u_2 \cdots u_{n-1} = v'_1 \cdots v'_{n-2}.$$

Assuming by symmetry that $v_{n-1} < v'_{n-1}$, it suffices to show that $f_{\mathbf{uv}}(x) < f_{\mathbf{uv}'}(y)$ for all

$$x = \pi_p(v_1 \cdots v_{n-1} c_1 c_2 \cdots) \in K_{\mathbf{v}} \quad \text{and} \quad y = \pi_p(v'_1 \cdots v'_{n-1} d_1 d_2 \cdots) \in K_{\mathbf{v}'}.$$

This inequality is equivalent to

$$\sum_{i=1}^{n-1} \frac{u_i}{p^i} + \frac{v_{n-1}}{p^n} + \frac{1}{p^n} \sum_{i=1}^{\infty} \frac{c_i}{p^i} < \sum_{i=1}^{n-1} \frac{u_i}{p^i} + \frac{v'_{n-1}}{p^n} + \frac{1}{p^n} \sum_{i=1}^{\infty} \frac{d_i}{p^i},$$

and hence to

$$\pi_p(c) < v'_{n-1} - v_{n-1} + \pi_p(d).$$

The last inequality follows from our choice of N and p at the beginning of the proof and from the form of the forbidden blocks. Indeed, using the relations

$$\alpha_{k+1}(q)\cdots\alpha_{k+N}(q) \le M^{N-1}(M-1) \quad k = 0, 1, 2, \ldots$$

we have

$$
\begin{aligned}
\pi_p(c) &\le \pi_p\big((M^{N-1}(M-1))^\infty\big) = \frac{M}{p-1} - \frac{1}{p^N - 1} \\
&< \frac{M}{p-1} - \frac{M}{p^n(p-1)} = \pi_p(M^n 0^\infty) \\
&= \pi_p(\alpha_1(q)\cdots\alpha_n(q)\, 0^\infty) + \pi_p(\overline{\alpha_1(q)\cdots\alpha_n(q)}\, 0^\infty) \\
&< \pi_p(\alpha(p)) + \pi_p(d) = 1 + \pi_p(d) \\
&\le v'_{n-1} - v_{n-1} + \pi_p(d). \quad \square
\end{aligned}
$$

Lemma 3.3. *Let $q \in [q', M+1)$. There exists a positive integer $n(q)$ and a real number $\varepsilon(q) > 0$ such that*

$$
\dim_H \pi_p(\widetilde{\mathcal{U}}'_{q,n}) = \frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log p} \quad and \quad \dim_H \pi_p(\widetilde{\mathcal{W}}'_{q,n}) = \frac{h(\widetilde{\mathcal{W}}'_{q,n})}{\log p}
$$

for all $n \ge n(q)$ and $p \in [q, q + \varepsilon(q))$.

Proof. We only give the proof for $\widetilde{\mathcal{W}}'_{q,n}$.

Fix two positive integers $n > N$ satisfying $\beta_N(q) < M$ and $q^{n-N}(q-1) > M$. Let $p \in [q, M+1)$ be sufficiently close to q such that

$$\beta_i(p) = \beta_i(q), \quad i = 1, \ldots, n.$$

Since $p \ge q$, we have also $p^{n-N}(p-1) > M$.

Similarly to the proof of [Lemma 3.2](#) we construct an edge graph representing $\widetilde{\mathcal{W}}'_{q,n}$, so that $\pi_p(\widetilde{\mathcal{W}}'_{q,n})$ is a graph-directed set. Then it suffices to prove that the corresponding iterated function system satisfies the open set condition, i.e.,

$$\pi_p(c) < 1 + \pi_p(d)$$

for all $c, d \in \widetilde{\mathcal{W}}'_{q,n}$.

This follows again from our choice of N and p at the beginning of the proof. Indeed, using the relations

$$\beta_{k+1}(q)\cdots\beta_{k+N}(q) \le M^{N-1}(M-1) \quad k = 0, 1, 2, \ldots$$

we have

$$\pi_p(c) \le \pi_p\big((M^{N-1}(M-1))^\infty\big) = \frac{M}{p-1} - \frac{1}{p^N-1}$$

$$< \frac{M}{p-1} - \frac{M}{p^n(p-1)} = \pi_p(M^n 0^\infty)$$

$$= \pi_p(\beta_1(q) \cdots \beta_n(q) \, 0^\infty) + \pi_p(\overline{\beta_1(q) \cdots \beta_n(q)} \, 0^\infty)$$

$$< \pi_p(\beta(p)) + \pi_p(d) = 1 + \pi_p(d). \quad \square$$

We are ready to prove Theorem 1.3.

Proof of Theorem 1.3. In view of Lemma 3.1 we may assume that $q \in [q', M+1)$.

We apply the first relation of the preceding lemma with $p = q$. Letting $n \to \infty$ and using Lemma 2.7 and Proposition 2.8 we obtain that

$$\dim_H \widetilde{\mathcal{U}}_q = \frac{h(\widetilde{\mathcal{U}}_q')}{\log q}.$$

Since $\dim_H \widetilde{\mathcal{U}}_q = \dim_H \mathcal{U}_q$ and $h(\widetilde{\mathcal{U}}_q') = h(\mathcal{U}_q')$ by Lemma 2.5, the equality (1.1) follows. $\square$

4. Proof of Theorem 1.4

In view of Lemma 3.1 it suffices to prove the theorem for $q \in [q', M+1]$.

Lemma 4.1. *The function D is left continuous in every $q \in [q', M+1]$.*

Proof. Fix $q \in [q', M+1]$ and $\varepsilon > 0$ arbitrarily. We have to show that if $p \in (1, q)$ is sufficiently close to q, then $|D(p) - D(q)| < \varepsilon$. The proof will be split into the following two cases.

Case I: $q \in [q', M+1)$. Using Proposition 2.8 we fix a sufficiently large index n such that

$$h(\widetilde{\mathcal{V}}_{q,n}') - h(\widetilde{\mathcal{U}}_{q,n}') < \frac{\varepsilon \log q}{2}.$$

Next we fix $p_n \in (q', q)$ sufficiently close to q, such that

$$\alpha_i(p_n) = \alpha_i(q) \quad \text{for} \quad i = 1, \ldots, n.$$

If $p \in (p_n, q)$, then using the inclusions

$$\widetilde{\mathcal{U}}_{q,n}' \subseteq \widetilde{\mathcal{U}}_p' \subseteq \widetilde{\mathcal{U}}_q' \subseteq \widetilde{\mathcal{V}}_{q,n}'$$

and applying Lemma 3.2 we obtain

$$\frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log p} = \dim_H \pi_p(\widetilde{\mathcal{U}}'_{q,n}) \le \dim_H \widetilde{\mathcal{U}}_p \le \dim_H \pi_p(\widetilde{\mathcal{V}}'_{q,n}) = \frac{h(\widetilde{\mathcal{V}}'_{q,n})}{\log p}$$

and

$$\frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log q} = \dim_H \pi_q(\widetilde{\mathcal{U}}'_{q,n}) \le \dim_H \widetilde{\mathcal{U}}_q \le \dim_H \pi_q(\widetilde{\mathcal{V}}'_{q,n}) = \frac{h(\widetilde{\mathcal{V}}'_{q,n})}{\log q}.$$

It follows that

$$
\begin{aligned}
|D(p) - D(q)| &\le \frac{h(\widetilde{\mathcal{V}}'_{q,n})}{\log p} - \frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log q} \\
&= \frac{h(\widetilde{\mathcal{V}}'_{q,n}) - h(\widetilde{\mathcal{U}}'_{q,n})}{\log p} + h(\widetilde{\mathcal{U}}'_{q,n})\left(\frac{1}{\log p} - \frac{1}{\log q}\right) \\
&< \frac{\varepsilon \log q}{2 \log p} + h(\widetilde{\mathcal{U}}'_{q,n})\left(\frac{1}{\log p} - \frac{1}{\log q}\right).
\end{aligned}
$$

If $p \in (p_n, q)$ is close enough to q, then the right side is $< \varepsilon$.

Case II: $q = M + 1$. Since $D(q) = 1$ and $0 \le D(p) \le 1$ for all p, it is sufficient to show that $D(p) > 1 - \varepsilon$ for all $p \in (1, q)$, close enough to q.

Since $h(\widetilde{\mathcal{U}}'_q) = \log q = \log(M + 1) > 0$ by Lemma 3.1, applying Proposition 2.8 we may fix a large integer n such that

$$h(\widetilde{\mathcal{U}}'_{q,n}) > \left(1 - \frac{\varepsilon}{2}\right)\log q.$$

If $p \in (1, q)$ is close enough to q, then

$$\alpha_i(p) = \alpha_i(q) \quad \text{for} \quad i = 1, \dots, n,$$

whence $\widetilde{\mathcal{U}}'_{q,n} \subseteq \widetilde{\mathcal{U}}'_p$ by (2.1). It follows that

$$h(\widetilde{\mathcal{U}}'_p) > \left(1 - \frac{\varepsilon}{2}\right)\log q.$$

Dividing by $\log p$ and applying Lemma 3.2 we infer that

$$D(p) > \left(1 - \frac{\varepsilon}{2}\right)\frac{\log q}{\log p}.$$

We conclude by observing that the right side is $> 1 - \varepsilon$ if p is close enough to q. $\square$

We remark that for $M = 1$ a simple direct proof was given for the left continuity in $q = 2$ in [9, Proposition 4.1 (i)].

Lemma 4.2. *The function D is right continuous in $[q', M + 1)$.*

Proof. Fix $q \in [q', M+1)$ and $\varepsilon > 0$ arbitrarily. We have to show that if $p \in (q, M+1)$ is sufficiently close to q, then $|D(p) - D(q)| < \varepsilon$.

Using Proposition 2.8 we fix a sufficiently large index n such that

$$h(\widetilde{\mathcal{W}}'_{q,n}) - h(\widetilde{\mathcal{U}}'_{q,n}) < \frac{\varepsilon \log q}{2}.$$

Next we fix $p_n \in (q, M+1)$ sufficiently close to q, such that

$$\beta_i(p_n) = \beta_i(q) \quad \text{for} \quad i = 1, \dots, n.$$

If $p \in (q, p_n)$, then using the inclusions

$$\widetilde{\mathcal{U}}'_{q,n} \subseteq \widetilde{\mathcal{U}}'_q \subseteq \widetilde{\mathcal{U}}'_p \subseteq \widetilde{\mathcal{W}}'_{q,n}$$

and applying Lemma 3.3 we obtain that

$$\dim_H \pi_p(\widetilde{\mathcal{U}}'_{q,n}) = \frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log p} \quad \text{and} \quad \dim_H \pi_p(\widetilde{\mathcal{W}}'_{q,n}) = \frac{h(\widetilde{\mathcal{W}}'_{q,n})}{\log p}.$$

Repeating the proof of Lemma 4.1 with $\widetilde{\mathcal{V}}'_{q,n}$ changed to $\widetilde{\mathcal{W}}'_{q,n}$, now we obtain the estimate

$$|D(p) - D(q)| \leq \frac{h(\widetilde{\mathcal{W}}'_{q,n})}{\log q} - \frac{h(\widetilde{\mathcal{U}}'_{q,n})}{\log p},$$

and we may conclude as before. □

In the next result we take any $q \in (1, \infty)$.

Lemma 4.3. *D has bounded variation in $[q', M+1]$.*

Proof. We prove that for every finite subdivision

$$q_0 := q' < q_1 < \dots < q_n = M+1$$

the following inequality holds:

$$\sum_{i=1}^{n} |D(q_i) - D(q_{i-1})| \leq \frac{2 \log(M+1)}{\log q'} - 1.$$

We know that $q \to h(\mathcal{U}'_q)$ is non-decreasing in $[q_0, M+1]$ with $h(\mathcal{U}'_{q_0}) = 0$ and $h(\mathcal{U}'_{M+1}) = \log(M+1)$. Therefore we have the following elementary inequalities:

$$D(q_i) - D(q_{i-1}) = \frac{h(\mathcal{U}'_{q_i})}{\log q_i} - \frac{h(\mathcal{U}'_{q_{i-1}})}{\log q_{i-1}} \leq \frac{h(\mathcal{U}'_{q_i}) - h(\mathcal{U}'_{q_{i-1}})}{\log q_i} \leq \frac{h(\mathcal{U}'_{q_i}) - h(\mathcal{U}'_{q_{i-1}})}{\log q_0}$$

and

$$D(q_i) - D(q_{i-1}) \geq \frac{h(\mathcal{U}'_{q_{i-1}})}{\log q_i} - \frac{h(\mathcal{U}'_{q_{i-1}})}{\log q_{i-1}} \geq \frac{\log(M+1)}{\log q_i} - \frac{\log(M+1)}{\log q_{i-1}}.$$

It follows that

$$|D(q_i) - D(q_{i-1})| \leq \frac{h(\mathcal{U}'_{q_i}) - h(\mathcal{U}'_{q_{i-1}})}{\log q_0} + \left(\frac{\log(M+1)}{\log q_{i-1}} - \frac{\log(M+1)}{\log q_i} \right),$$

and hence

$$\sum_{i=1}^{n} |D(q_i) - D(q_{i-1})|$$

$$\leq \frac{h(\mathcal{U}'_{M+1}) - h(\mathcal{U}'_{q_0})}{\log q_0} + \frac{\log(M+1)}{\log q_0} - \frac{\log(M+1)}{\log(M+1)}$$

$$= \frac{2\log(M+1)}{\log q_0} - 1,$$

as stated. $\quad \square$

5. The Hausdorff dimension of $\mathcal{U}$

As usual, we denote by $\mathcal{U}$ the set of bases $q > 1$ in which $x = 1$ has a unique expansion, and by $\mathcal{U}'$ the set of corresponding expansions. We recall from [12] and [22] that a sequence $c = (c_i)$ belongs to $\mathcal{U}'$ if and only if the lexicographic inequalities

$$\overline{c_1 c_2 \cdots} < c_{k+1} c_{k+2} \cdots < c_1 c_2 \cdots \tag{5.1}$$

are satisfied for all $k \geq 1$.

Fix an integer $N \geq 2$ and, inspired by the proof of [9, Proposition 4.1 (i)], consider the set $\hat{\mathcal{U}}'_N$ of sequences $c = (c_i) \in \{0, \ldots, M\}^{\infty}$ satisfying the equality

$$c_1 \cdots c_{2N} = M^{2N-1} 0,$$

and the lexicographic inequalities

$$0^N < c_{kN+1} \cdots c_{kN+N} < M^N$$

for $k = 2, 3, \ldots$. All these sequences satisfy (5.1), so that $\hat{\mathcal{U}}'_N \subseteq \mathcal{U}'$ and $\hat{\mathcal{U}}_N \subseteq \mathcal{U}$, where we use the natural notation

$$\hat{\mathcal{U}}_N := \left\{ q \in (1, M+1] \ : \ \beta(q) \in \hat{\mathcal{U}}'_N \right\}.$$

(Here $\beta(q)$ denotes the unique and hence also greedy expansion of $x = 1$ in base q.)

It follows from the definition of $\hat{\mathcal{U}}'_N$ that

$$\left| B_{nN}(\hat{\mathcal{U}}'_N) \right| = \left((M+1)^N - 2 \right)^{n-2} \quad \text{for all} \quad n \geq 2 \tag{5.2}$$

and

$$\left| B_{kN+1,nN}(\hat{\mathcal{U}}'_N) \right| = \left((M+1)^N - 2 \right)^{n-k} \quad \text{for all} \quad n \geq k \geq 2. \tag{5.3}$$

Consider two elements $p < q$ of $\hat{\mathcal{U}}_N$, and let m be the smallest positive integer such that $\beta_m(p) \neq \beta_m(q)$. Then $\beta_m(p) < \beta_m(q)$, and we deduce from the definition of $\hat{\mathcal{U}}_N$ that

$$\left(\sum_{i=1}^m \frac{\beta_i(q)}{q^i} \right) + \frac{1}{q^{m+2N}} < 1 < \left(\sum_{i=1}^m \frac{\beta_i(p)}{p^i} \right) + \frac{1}{p^m} \leq \sum_{i=1}^m \frac{\beta_i(q)}{p^i}.$$

Hence

$$\frac{1}{q^{m+2N}} < \sum_{i=1}^m \beta_i(q) \left(p^{-i} - q^{-i} \right) < M \sum_{i=1}^\infty \left(p^{-i} - q^{-i} \right) = \frac{M(q-p)}{(p-1)(q-1)}$$

and therefore

$$\frac{1}{(M+1)^{m+2N}} < \frac{M(q-p)}{(q'-1)^2},$$

where q' denotes the Komornik–Loreti constant as usual.

Setting

$$c := \frac{(q'-1)^2}{M(M+1)^{2N}}$$

we conclude the following

Lemma 5.1. *If $p, q \in \hat{\mathcal{U}}_N$ and $0 < q - p \leq c(M+1)^{-m}$ for some positive integer m, then $\beta_i(p) = \beta_i(q)$ for all $i = 1, \ldots, m$.*

Now we are ready to compute the Hausdorff dimension of $\mathcal{U}$.

Proof of Theorem 1.6 (ii). Consider a finite cover $\cup I_j$ of $\hat{\mathcal{U}}_N$ by intervals I_j of length $|I_j| \leq c(M+1)^{-N}$. For each positive integer k we denote by J_k the set of indices j satisfying the inequalities

$$c(M+1)^{-(k+1)N} < |I_j| \leq c(M+1)^{-kN}.$$

We fix a large integer n satisfying $c(M+1)^{-nN} < |I_j|$ for all j; then $J_k = \varnothing$ for all $k \geq n$.

If $j \in J_k$ and $p, q \in \hat{\mathcal{U}}_N \cap I_j$, then the first kN digits of $\beta(p)$ and $\beta(q)$ coincide by the above lemma, so that at most $\left| B_{kN+1,nN}(\hat{\mathcal{U}}'_N) \right|$ elements of $B_{nN}(\hat{\mathcal{U}}'_N)$ may occur for the bases $q \in \hat{\mathcal{U}}_N \cap I_j$. Hence

$$\left| B_{nN}(\hat{\mathcal{U}}'_N) \right| \leq \sum_k \sum_{j \in J_k} \left| B_{kN+1,nN}(\hat{\mathcal{U}}'_N) \right|.$$

Using (5.2) and (5.3) this is equivalent to

$$\left((M+1)^N - 2 \right)^{-2} \leq \sum_k \sum_{j \in J_k} \left((M+1)^N - 2 \right)^{-k}.$$

Introducing the number $\sigma = \sigma(N) \in (0,1)$ by the equation

$$(M+1)^N - 2 = (M+1)^{\sigma N}, \tag{5.4}$$

we may rewrite the preceding inequality in the form

$$(M+1)^{-2\sigma N} \leq \sum_k \sum_{j \in J_k} (M+1)^{-\sigma Nk}.$$

Since

$$(M+1)^{-Nk} < c^{-1}(M+1)^N |I_j|$$

by the definition of J_k, it follows that

$$(M+1)^{-2\sigma N} \leq \sum_k \sum_{j \in J_k} c^{-\sigma}(M+1)^{\sigma N} |I_j|^\sigma$$

or equivalently

$$\sum_j |I_j|^\sigma \geq c^\sigma (M+1)^{-3\sigma N}.$$

Since the right side is positive and depends only on N, we conclude that $\dim_H \hat{\mathcal{U}}_N \geq \sigma(N)$.

It follows from the definition (5.4) that $\sigma(N) \to 1$ as $N \to \infty$. Since $\hat{\mathcal{U}}_N \subseteq \mathcal{U} \subseteq \mathbb{R}$ for all N, letting $N \to \infty$ we conclude that $\dim_H \mathcal{U} = 1$. $\quad\square$

6. Proof of Theorem 1.5 and the Lebesgue measure of $\mathcal{U}$

Set $\mathcal{B}' := \{\beta(q) \;:\; q \in (1, M+1]\}$ for brevity.

Our main tool is a generalization of a reasoning in [11]. Given two positive integers n, t and a word $\eta_1 \cdots \eta_n \in B_n(\mathcal{B}')$, the sets

$$\{q \in [1, M+1) \;:\; \beta_i(q) = \eta_i, \quad i = 1, \ldots, n\}$$

and

$$\left\{ q \in [1, M+1) \;:\; \beta_i(q) = \begin{cases} \eta_i, & i = 1, \ldots, n, \\ 0, & i = n+1, \ldots, n+t \end{cases} \right\}$$

are two intervals $[q_1, q_2)$ and $[q_1, q_3)$ satisfying $q_3 \le q_2$.

Lemma 6.1. *The following inequality holds:*

$$\frac{q_3 - q_1}{q_2 - q_1} \ge \frac{(q_1 - 1)^3}{M^2 q_2^{t+2}}.$$

We stress the fact that the right side does not depend on n.

Proof. It follows from the greedy algorithm that

$$\sum_{i=1}^{n} \frac{\eta_i}{q_1^i} = 1, \tag{6.1}$$

$$\sum_{i=1}^{n} \frac{\eta_i}{q_2^i} + \sum_{i=n+1}^{\infty} \frac{M}{q_2^i} \ge 1$$

and

$$\left(\sum_{i=1}^{n} \frac{\eta_i}{q_3^i} \right) + \frac{1}{q_3^{n+t}} = 1. \tag{6.2}$$

Using the first two relations and the relation $\eta_1 \ge 1$ we obtain that

$$\frac{M}{q_2^n (q_2 - 1)} \ge \sum_{i=1}^{n} \eta_i \left(q_1^{-i} - q_2^{-i} \right) \ge q_1^{-1} - q_2^{-1} = \frac{q_2 - q_1}{q_1 q_2}.$$

Hence

$$(0 <) q_2 - q_1 \le \frac{M q_1 q_2}{q_2^n (q_2 - 1)}. \tag{6.3}$$

Similarly, using (6.1) and (6.2) we obtain that

$$
\frac{1}{q_3^{n+t}} = \sum_{i=1}^{n} \eta_i \left(q_1^{-i} - q_3^{-i} \right) \leq M \sum_{i=1}^{\infty} \left(q_1^{-i} - q_3^{-i} \right)
$$

$$
= M \left(\frac{q_1^{-1}}{1 - q_1^{-1}} - \frac{q_3^{-1}}{1 - q_3^{-1}} \right) = \frac{M(q_3 - q_1)}{(q_1 - 1)(q_3 - 1)}.
$$

Hence

$$
q_3 - q_1 \geq \frac{(q_1 - 1)(q_3 - 1)}{M q_3^{n+t}}. \tag{6.4}
$$

Combining (6.3) and (6.4), and using the inequalities $q_1 \leq q_3 \leq q_2$ we conclude that

$$
\frac{q_3 - q_1}{q_2 - q_1} \geq \frac{(q_1 - 1)(q_3 - 1)}{M q_3^{n+t}} \cdot \frac{q_2^n (q_2 - 1)}{M q_1 q_2} \geq \frac{(q_1 - 1)^3}{M^2 q_2^{t+2}}. \quad \square
$$

In the next lemma λ denotes the usual Lebesgue measure.

Lemma 6.2. *The following inequality holds for all $1 < p < r \leq M+1$ and for all positive integers n and t:*

$$
\lambda \left(\{ q \in [p, r) \ : \ \beta_{n+1}(q) = \cdots = \beta_{n+t}(q) = 0 \} \right) \geq \frac{(p-1)^3}{M^2 r^{t+2}} (r - p).
$$

Before proving the lemma we recall that the bases q for which $\beta(q)$ is finite form a (countable) dense set in $[1, M+1]$. Indeed, if $\beta(q)$ is infinite for some q, then the truncated sequences $\beta_1(q) \cdots \beta_k(q) 0^\infty$ belong to $\mathcal{B}'$ for all $k = 1, 2, \ldots$ by an elementary reasoning given in [23, Lemma 3.1]. Therefore there exist bases $q_k \in [1, M+1]$ such that

$$
\beta(q_k) = \beta_1(q) \cdots \beta_k(q) 0^\infty,
$$

and then $q_k \to q$.

Proof. We use the notations of the preceding lemma.

We may assume by density that $\beta(p)$ and $\beta(r)$ are finite. Choose a sufficiently large integer n such that $\beta_i(p) = \beta_i(r) = 0$ for all $i > n$, and consider the intervals $[q_1, q_2)$ corresponding to n. Then some of these intervals form a finite partition of $[p, r)$. Since we have

$$
\frac{q_3 - q_1}{q_2 - q_1} \geq \frac{(q_1 - 1)^3}{M^2 q_2^{t+2}} \geq \frac{(p-1)^3}{M^2 r^{t+2}}
$$

for each of these intervals by the preceding lemma, the required inequality follows by summing the inequalities

$$q_3 - q_1 \geq \frac{(p-1)^3}{M^2 r^{t+2}}(q_2 - q_1). \quad \square$$

Lemma 6.3. *Given an arbitrary real number $s > 1$, there exists a sequence (n_k) of natural numbers satisfying the inequalities*

$$n_k > \log_s (n_1 + \cdots + n_k), \quad k = 1, 2, \ldots$$

and the divergence relation

$$\sum_{k=1}^{\infty} s^{-n_k} = \infty.$$

Proof. For $s = 2$ this was proved in [12, Lemma 6]. The proof remains valid for every $s > 1$. $\quad \square$

Now we are ready to prove Theorem 1.5:

Proof of Theorem 1.5. By density it suffices to show for any fixed $1 < p < r \leq M + 1$, the required property holds for almost all $q \in [p, r)$. For convenience we normalize λ and we use the equivalent probabilistic measure $\mu := \frac{\lambda}{r-p}$ on $[p, r)$. Then we may adapt the usual proof of the Borel–Cantelli lemma.

Choose a sequence (n_k) satisfying the conditions of the preceding lemma with $s := r$, and set

$$C_j := \left\{ q \in [p, r) \ : \ \sum_{i=n_1+\cdots+n_{j-1}+1}^{n_1+\cdots+n_j} \beta_i(q) > 0 \right\}, \quad j = 1, 2, \ldots.$$

It follows from Lemma 6.2 and 6.3 that

$$\mu \left(\cap_{j=k}^{\infty} C_j \right) \leq \prod_{j=k}^{\infty} \left(1 - \frac{(p-1)^3}{M^2 r^2} r^{-n_j} \right) = 0$$

for every $k = 1, 2, \ldots$.

Therefore $C := \cup_{k=1}^{\infty} \cap_{j=k}^{\infty} C_j$ has also zero Lebesgue measure. We complete the proof by observing that if $q \in [p, r) \setminus C$, then $\beta(q)$ has the required property for infinitely many $m = n_1 + \cdots + n_k$. $\quad \square$

Finally we compute the Lebesgue measure of $\mathcal{U}$:

Proof of Theorem 1.6 (i). Since $\overline{\mathcal{U}} \setminus \mathcal{U}$ is countable, it suffices to prove that $\mathcal{U}$ is a null set. Furthermore, it suffices to prove that $\mathcal{U} \cap [p, M+1)$ is a null set for each $p \in (1, M+1)$ such that $\beta(p)$ is finite.

It follows from the lexicographical characterization (5.1) of $\mathcal{U}$ that $\mathcal{U} \cap [p, M+1) \subseteq C$, where C is the null set in the proof of the above lemma, corresponding to the choice $[p, r) = [p, M+1)$. Hence $\mathcal{U} \cap [p, M+1)$ is a null set indeed. $\square$

7. Proof of Theorem 1.7

In view of Theorems 1.1 and 1.4 it suffices to prove that $D' < 0$ almost everywhere in (q', ∞). This was implicitly proved in [24, Theorems 2.5 and 2.6]. Here we give an alternative proof.

Since $\overline{\mathcal{U}}$ is a null set by Theorem 1.5 (i), it suffices to prove that $D' < 0$ everywhere in each connected component $I = (q_0, q_0^*)$ of $(q', \infty) \setminus \overline{\mathcal{U}}$. Fixing $p \in (q_0, q_0^*)$ arbitrarily, we deduce from Theorem 1.3 and Lemma 2.11 that

$$D(q) = \frac{h(\mathcal{U}_p')}{\log q}$$

for all $q \in I$, and therefore

$$D'(q) = -\frac{h(\mathcal{U}_p')}{q(\log q)^2}$$

for all $q \in I$. Since $p > q'$ and therefore $h(\mathcal{U}_p') > 0$ by Theorem 1.1, we have $D'(q) < 0$ for all $q \in I$ indeed.

Remark. Since q' and $M+1$ are the smallest and largest elements of $\overline{\mathcal{U}}$, the first and last connected components of $(1, \infty) \setminus \overline{\mathcal{U}}$ are $(1, q')$ and $(M+1, \infty)$.

We recall from [8] that the left and right endpoints of the remaining connected components $I = (q_0, q_0^*)$ run over $\overline{\mathcal{U}} \setminus \mathcal{U}$ and some proper subset $\mathcal{U}^*$ of $\mathcal{U}$, respectively.

It follows from some theorems of Parry [28] and Solomyak [33] that each element of $\overline{\mathcal{U}} \setminus \mathcal{U}$ is an algebraic integer, all of whose conjugates are smaller than the Golden Ratio in modulus.

On the other hand, it was proved in [24] that the points q_0^*, called *de Vries–Komornik numbers*, are transcendental. The smallest one is the *Komornik–Loreti constant* q'. Their expansions are closely related to the classical Thue–Morse sequence.

References

[1] M. Baatz, V. Komornik, Unique expansions in integer bases with extended alphabets, Publ. Math. Debrecen 79 (3–4) (2011) 251–267.

[2] C. Baiocchi, V. Komornik, Greedy and quasi-greedy expansions in non-integer bases, arXiv:0710.3001 [math.NT], October 16, 2007.

[3] S. Baker, Generalised golden ratios over integer alphabets, Integers 14 (2014) A15, 28 pp.

[4] Z. Daróczy, I. Kátai, Univoque sequences, Publ. Math. Debrecen 42 (3–4) (1993) 397–407.

[5] Z. Daróczy, I. Kátai, On the structure of univoque numbers, Publ. Math. Debrecen 46 (3–4) (1995) 385–408.

[6] M. de Vries, A property of algebraic univoque numbers, Acta Math. Hungar. 119 (1–2) (2008) 57–62.

[7] M. de Vries, On the number of unique expansions in noninteger bases, Topology Appl. 156 (3) (2009) 652–657.

[8] M. de Vries, V. Komornik, Unique expansions of real numbers, Adv. Math. 221 (2009) 390–427.

[9] M. de Vries, V. Komornik, A two-dimensional univoque set, Fund. Math. 212 (2011) 175–189.

[10] M. de Vries, V. Komornik, Expansions in non-integer bases, in: V. Berté, M. Rigo (Eds.), Combinatorics, Words and Symbolic Dynamics, 2016.

[11] P. Erdős, I. Joó, On the expansion $1 = \sum q^{-n_i}$, Period. Math. Hungar. 23 (1) (1991) 25–28.

[12] P. Erdős, I. Joó, V. Komornik, Characterization of the unique expansions $1 = \sum q^{-n_i}$ and related problems, Bull. Soc. Math. France 118 (1990) 377–390.

[13] P. Erdős, M. Horváth, I. Joó, On the uniqueness of the expansions $1 = \sum q^{-n_i}$, Acta Math. Hungar. 58 (1991) 333–342.

[14] K. Falconer, Fractal Geometry. Mathematical Foundations and Applications, second edition, John Wiley & Sons, Chichester, 2003.

[15] P. Glendinning, N. Sidorov, Unique representations of real numbers in non-integer bases, Math. Res. Lett. 8 (2001) 535–543.

[16] J. Hutchinson, Fractals and self-similarity, Indiana Univ. Math. J. 30 (1981) 713–747.

[17] G. Kallós, The structure of the univoque set in the small case, Publ. Math. Debrecen 54 (1–2) (1999) 153–164.

[18] G. Kallós, The structure of the univoque set in the big case, Publ. Math. Debrecen 59 (3–4) (2001) 471–489.

[19] I. Kátai, G. Kallós, On the set for which 1 is univoque, Publ. Math. Debrecen 58 (4) (2001) 743–750.

[20] V. Komornik, Expansions in noninteger bases, Integers 11B (2011) A9, 30 pp.

[21] V. Komornik, P. Loreti, Unique developments in noninteger bases, Amer. Math. Monthly 105 (1998) 636–639.

[22] V. Komornik, P. Loreti, Subexpansions, superexpansions and uniqueness properties in noninteger bases, Period. Math. Hungar. 44 (2) (2002) 195–216.

[23] V. Komornik, P. Loreti, On the topological structure of univoque sets, J. Number Theory 122 (2007) 157–183.

[24] D. Kong, W. Li, Hausdorff dimension of unique beta expansions, Nonlinearity 28 (2015) 187–209.

[25] D. Kong, W. Li, F.M. Dekking, Intersections of homogeneous Cantor sets and beta-expansions, Nonlinearity 23 (2010) 2815–2834.

[26] D. Lind, B. Marcus, An Introduction to Symbolic Dynamics and Coding, Cambridge University Press, 1996.

[27] D. Mauldin, C. Williams, Hausdorff dimension in graph directed constructions, Trans. Amer. Math. Soc. 309 (1988) 811–829.

[28] W. Parry, On the β-expansion of real numbers, Acta Math. Hungar. 11 (1960) 401–416.

[29] A. Rényi, Representations for real numbers and their ergodic properties, Acta Math. Hungar. 8 (1957) 477–493.

[30] J. Schmeling, Symbolic dynamics for β-shifts and self-normal numbers, Ergodic Theory Dynam. Systems 17 (1997) 675–694.

[31] N. Sidorov, Almost every number has a continuum of β-expansions, Amer. Math. Monthly 110 (9) (2003) 838–842.

[32] N. Sidorov, Arithmetic dynamics, in: Topics in Dynamics and Ergodic Theory, in: London Math. Soc. Lecture Note Ser., vol. 310, 2003, pp. 145–189.
[33] B. Solomyak, Conjugates of beta-numbers and the zero-free domain for a class of analytic functions, Proc. Lond. Math. Soc. (3) 68 (3) (1994) 477–498.