

## Unique expansions in integer bases with extended alphabets

By MARTINE BAATZ (Luxembourg) and VILMOS KOMORNIK (Strasbourg)

*Dedicated to Professors Kálmán Györy, Attila Pethő, János Pintz  
and András Sárközy on their birthdays*

**Abstract.** Since their introduction by Rényi more than fifty years ago, the investigation of expansions in noninteger bases led to a number of deep and unexpected results. Some of them led to the necessity to study expansions in integer bases on an enlarged alphabet containing the base itself as a possible digit. We show in the present paper how certain recent theorems change in this framework.

### 1. Introduction

Beginning with RÉNYI [20] many works have been devoted to expansions of the form

$$x = \sum_{i=1}^{\infty} \frac{c_i}{q^i} \quad (1.1)$$

in arbitrary *real* bases  $q > 1$  with integer *digits* satisfying  $0 \leq c_i < q$ . It is easy to see that a real number  $x$  has an expansion if and only if it belongs to the closed interval

$$J_q := \left[ 0, \frac{[q] - 1}{q - 1} \right]$$

---

*Mathematics Subject Classification:* Primary: 11A63; Secondary: 11B83.

*Key words and phrases:* greedy expansions, beta-expansions, univoque sequences, univoque sets, Cantor sets.

where  $\lceil q \rceil$  stands for the upper integer part of  $q$ .<sup>1</sup>

The case of integer bases is of course well-known:  $J_q = [0, 1]$  and the expansion of every  $x \in [0, 1]$  is unique, except if  $x = m/q^n$  for some positive integers  $m$  and  $n$ . These exceptional numbers  $x$  have exactly two expansions: a *finite* one ending with  $0^\infty$  and an *infinite* one ending with  $(q-1)^\infty$ . In particular, the set  $\mathcal{U}_q$  of numbers  $x$  having a unique expansion is a nonclosed set of Hausdorff dimension one, having full Lebesgue measure in  $J_q$ .

For noninteger bases the situation is radically different. The set  $\mathcal{U}_q$  has Hausdorff dimension strictly smaller than one and hence zero Lebesgue measure, but it is still an uncountable set. Furthermore,  $\mathcal{U}_q$  is closed for almost all values of  $q$ , the exceptional bases forming a Cantor set of Hausdorff dimension one but of zero Lebesgue measure. Moreover, the set  $B$  of bases for which  $\mathcal{U}_q$  is a Cantor set has both interior and exterior points, i.e., both  $B$  and  $(1, \infty) \setminus B$  contain nondegenerate intervals. In cases where  $\mathcal{U}_q$  is not closed,  $\overline{\mathcal{U}}_q \setminus \mathcal{U}_q$  is a countably infinite set, each  $x \in \overline{\mathcal{U}}_q \setminus \mathcal{U}_q$  has countably many expansions, all of which are explicitly known. We refer to [2], [3], [4], [10], [11], [12], [13], [14], [17] for details.

Sometimes more elegant results are obtained by considering expansions (1.1) with integer *digits* satisfying  $0 \leq c_i \leq q$  (the two definitions differ only for integer bases  $q$ ). Then  $J_q$  is replaced by

$$J_q^* := \left[ 0, \frac{\lfloor q \rfloor}{q-1} \right]$$

where  $\lfloor q \rfloor$  denotes the lower integer part of  $q$  as introduced above. The purpose of this work is to investigate the modified *univoque sets*  $\mathcal{U}_q^*$  for integer bases by using this extended alphabet. It turns out that their structure is different from the usual case.

As we will see, the cases  $q = 2$  and  $q > 2$  are quite different.

Throughout this paper, the index set for all sequences is the set of positive integers:  $(c_i) = (c_i)_{i=1}^\infty$ ,  $(\alpha_i) = (\alpha_i)_{i=1}^\infty$ , and so on. Hence we will often omit the indication of the index set.

## 2. Review of univoque bases

In this section we recall some results from [14]–[16]. Given a real number  $q > 1$  we consider expansions in base  $q$  on the *alphabet*  $\{0, \dots, \lceil q \rceil - 1\}$ , i.e.

---

<sup>1</sup>If  $q$  is not integer, then its lower and upper integer parts are by definition the consecutive integers satisfying the inequalities  $\lfloor q \rfloor < q < \lceil q \rceil$ . If  $q$  is integer, then we define  $\lfloor q \rfloor = \lceil q \rceil := q$ .

equalities of the form

$$x = \frac{c_1}{q} + \frac{c_2}{q^2} + \dots \quad (2.1)$$

with integer *digits* satisfying  $0 \leq c_i < q$ . A real number  $x$  has at least one expansion if and only if

$$0 \leq x \leq \frac{\lceil q \rceil - 1}{q - 1}.$$

We say that  $q$  is a *univoque base* if  $x = 1$  has only one expansion in base  $q$ . The integer bases are univoque, but there exist other univoque bases too. In order to characterize them, it is convenient to introduce a particular expansion  $(\alpha_i) = (\alpha_i)_{i=1}^\infty$  of  $x = 1$  in each fixed base  $q > 1$  by the following algorithm: if  $\alpha_n$  has already been defined for some  $n \geq 1$  (no hypothesis if  $n = 0$ ), then let  $\alpha_{n+1}$  be the biggest nonnegative integer satisfying

$$\frac{\alpha_1}{q} + \dots + \frac{\alpha_{n+1}}{q^{n+1}} < 1.$$

It is called the *quasi-greedy* expansion of 1 in base  $q$ .

*Remark 2.1.* The sequence  $(\alpha_i)$  always has the following two properties:

- (a)  $\alpha_1 = \lceil q \rceil - 1$  is the biggest element of the alphabet;
- (b) we have  $(\alpha_{n+i}) \leq (\alpha_i)$  for all  $n$  in the lexicographic sense.

In the following theorem and in the sequel we define the conjugate of a digit  $c_i$  by  $\overline{c}_i := \alpha_1 - c_i$ .

**Theorem 2.2** ([15, Theorem 3.1]). *A base  $q > 1$  is univoque if and only if the following lexicographic inequalities are satisfied:*

$$\begin{aligned} (\alpha_{n+i}) &< (\alpha_i) \quad \text{whenever} \quad \alpha_n < \alpha_1; \\ (\overline{\alpha_{n+i}}) &< (\alpha_i) \quad \text{whenever} \quad \alpha_n > 0. \end{aligned}$$

For example, the periodical sequence  $(c_i)_{i=1}^\infty := 1(10)^\infty$  is the unique expansion of  $x = 1$  in the base defined by the equality (2.1).

The sequence  $(\alpha_i)$  also allows us to characterize the closure  $\overline{\mathcal{U}}$  of the set  $\mathcal{U}$  of univoque bases:

**Theorem 2.3** ([16, Theorem 2.4]). *A base  $q > 1$  belongs to  $\overline{\mathcal{U}}$  if and only if*

$$(\overline{\alpha_{n+i}}) < (\alpha_i) \quad \text{whenever} \quad \alpha_n > 0.$$

Comparing the above two theorems it is natural to investigate also the set  $\mathcal{V}$  of bases  $q$  for which the second lexicographic inequality is satisfied only in the weaker sense:

$$(\overline{\alpha_{n+i}}) \leq (\alpha_i) \quad \text{whenever} \quad \alpha_n > 0.$$

The topological properties of these sets are summarized in the following result:

**Theorem 2.4** ([16, Theorem 2.5, 2.6]).

- (a) We have  $\mathcal{U} \subset \overline{\mathcal{U}} \subset \mathcal{V} = \overline{\mathcal{V}}$ . All these sets have zero Lebesgue measure and Hausdorff dimension one.
- (b)  $\overline{\mathcal{U}} \setminus \mathcal{U}$  is a countable dense set in  $\overline{\mathcal{U}}$  and therefore  $\overline{\mathcal{U}}$  is a Cantor set.
- (c)  $\mathcal{V}$  is a closed set and  $\mathcal{V} \setminus \overline{\mathcal{U}}$  is a discrete set, dense in  $\mathcal{V}$ .
- (d) The set of expansions of  $x = 1$  is countably infinite in each base  $q \in \mathcal{V} \setminus \mathcal{U}$ .

*Remarks 2.5.*

- (a) The smallest element of  $\mathcal{V}$  is the golden ratio.
- (b)  $\mathcal{U}$  has a smallest element: the now so-called *Komornik-Loreti constant* is a transcendental number, its approximate value is 1.787.

*Example 2.6.* If the base  $q$  is an integer, then the quasi-greedy expansion of  $x = 1$  is given by  $\alpha_i = q - 1$  for all  $i$ . It follows from Theorem 2.2 that  $q \in \mathcal{U}$ .

### 3. Review of unique expansions

In this section we recall some results from [4]. Given a real number  $q > 1$  we consider again expansions in base  $q$  on the *alphabet*  $\{0, \dots, [q] - 1\}$ , i.e. equalities of the form

$$x = \frac{c_1}{q} + \frac{c_2}{q^2} + \dots \quad (3.1)$$

with integer *digits* satisfying  $0 \leq c_i < q$ . We denote by  $\mathcal{U}_q$  the set of real numbers  $x$  which have only one expansion in base  $q$ . For example,  $1 \in \mathcal{U}_q$  if and only if  $q \in \mathcal{U}$ .

Using the sequence  $(\alpha_i)$  introduced in the preceding section, the following characterization of  $\mathcal{U}_q$  is an easy corollary of a classical theorem of Parry [18]:

**Theorem 3.1** ([4, Theorem 1.1 (ii)]). *Given a base  $q > 1$  and an expansion (3.1), we have  $x \in \mathcal{U}_q$  if and only if the following two lexicographic conditions are satisfied:*

$$(c_{n+i}) < (\alpha_i) \quad \text{whenever} \quad c_n < \alpha_1;$$

$$(\overline{c_{n+i}}) < (\alpha_i) \quad \text{whenever} \quad c_n > 0.$$

The set  $\mathcal{U}_q$  is closed for *almost every* base  $q$  with respect to the Lebesgue measure. More precisely and rather surprisingly,  $\mathcal{U}_q$  is closed if and only if  $q \notin \overline{\mathcal{U}}$ . In order to get a complete picture we define the *quasi-greedy* expansion of every real number

$$0 < x \leq \frac{\lceil q \rceil - 1}{q - 1}$$

in base  $q$  by the following algorithm: if  $a_n(x)$  has already been defined for some  $n \geq 1$  (no hypothesis if  $n = 0$ ), then let  $a_{n+1}(x)$  be the biggest nonnegative integer satisfying

$$\frac{a_1(x)}{q} + \dots + \frac{a_{n+1}(x)}{q^{n+1}} < x.$$

Furthermore, it is convenient to set  $(a_i(x)) := 0^\infty$  if  $x = 0$ .

*Remark 3.2.* All quasi-greedy expansions satisfy the condition

$$(a_{n+i}(x)) \leq (\alpha_i) \quad \text{whenever} \quad a_n(x) < \alpha_1.$$

Next, analogously to the preceding section, we write  $x \in \mathcal{V}_q$  if

$$(\overline{a_{n+i}(x)}) \leq (\alpha_i) \quad \text{whenever} \quad a_n(x) > 0.$$

Similarly to Theorem 2.4 (a) we have always  $\mathcal{U}_q \subset \overline{\mathcal{U}}_q \subset \mathcal{V}_q = \overline{\mathcal{V}}_q$ . However, the finer picture depends on the given value of  $q$ . The following results were given in [4, Theorems 1.3, 1.4, 1.5] and in the remarks following their statements.

**Theorem 3.3.**

- (a) If  $q \in \overline{\mathcal{U}}$ , then  $\overline{\mathcal{U}}_q = \mathcal{V}_q$  and  $\mathcal{V}_q \setminus \mathcal{U}_q$  is a countable dense set in  $\mathcal{V}_q$ .
- (b) If  $q \in \mathcal{V} \setminus \overline{\mathcal{U}}$ , then  $\mathcal{U}_q$  is closed:  $\overline{\mathcal{U}}_q = \mathcal{U}_q$ , and  $\mathcal{V}_q \setminus \mathcal{U}_q$  is a discrete set, dense in  $\mathcal{V}_q$ .
- (c) If  $q \in (1, \infty) \setminus \mathcal{V}$ , then the sets  $\mathcal{U}_q$  and  $\mathcal{V}_q$  are closed and equal:  $\mathcal{U}_q = \overline{\mathcal{U}}_q = \mathcal{V}_q = \overline{\mathcal{V}}_q$ .

We have, moreover, the following result concerning the number of expansions of any  $x \in \mathcal{V}_q$ :

**Theorem 3.4.**

- (a) If  $q \in \mathcal{U}$ , then each  $x \in \mathcal{V}_q \setminus \mathcal{U}_q$  has exactly two expansions.
- (b) If  $q \in \mathcal{V} \setminus \mathcal{U}$ , then the set of expansions of each  $x \in \mathcal{V}_q \setminus \mathcal{U}_q$  is countably infinite.

Since the integer bases are univoque, we have the following:

**Corollary 3.5.** *Let  $q \geq 2$  be an integer. Then*

- (a)  $\overline{\mathcal{U}}_q = \mathcal{V}_q = [0, 1]$ , and  $\mathcal{V}_q \setminus \mathcal{U}_q$  is a countable dense set in  $\mathcal{V}_q$ ;
- (b) each  $x \in \mathcal{V}_q \setminus \mathcal{U}_q$  has exactly two expansions.

The corollary follows at once from the well-known fact that all numbers  $x \in [0, 1]$  belong to  $\mathcal{U}_q$  except countably many rational numbers  $0 < x < 1$  of the form  $x = \frac{m}{q^n}$  with positive integers  $m$  and  $n$ , for which there are two expansions ending with  $0^\infty$  and  $(q-1)^\infty$ , respectively.

In the following two sections we investigate what happens with this corollary if we consider the expansions over the enlarged alphabet  $\{0, \dots, q\}$ .

#### 4. Unique expansions in integer bases with enlarged alphabets

Fix an integer  $q \geq 2$  and consider the expansions

$$x = \frac{c_1}{q} + \frac{c_2}{q^2} + \dots$$

on the alphabet  $\{0, 1, \dots, q\}$ . In order to have an expansion, now  $x$  has to belong to the interval

$$J_q^* := \left[0, \frac{q}{q-1}\right].$$

Conversely, each  $x \in J_q^*$  has at least one expansion, given for example the following modification of the quasi-greedy expansion  $(a_i(x))$  of the preceding section. The quasi-greedy expansion of  $x = 0$  is by definition  $0^\infty$ . If  $x > 0$  and if  $a_n(x)$  has already been defined for some  $n \geq 1$  (no hypothesis if  $n = 0$ ), then let  $a_{n+1}(x)$  be the biggest element of  $\{0, 1, \dots, q\}$  satisfying the inequality

$$\frac{a_1(x)}{q} + \dots + \frac{a_{n+1}(x)}{q^{n+1}} < x.$$

*Remark 4.1.* As a special case of a more general result in [1], all quasi-greedy expansions satisfy the condition

$$(a_{n+i}(x)) \leq (q-1)^\infty \quad \text{whenever} \quad a_n(x) < q,$$

and conversely, every infinite sequence  $(c_i)$  satisfying the condition

$$(c_{n+i}) \leq (q-1)^\infty \quad \text{whenever} \quad c_n < q$$

is the quasi-greedy expansion of a suitable real number  $x$ .

Let us denote by  $\mathcal{U}_q^*$  the set of numbers  $x \in J_q^*$  having a unique expansion in base  $q$  with digits belonging to the enlarged alphabet  $\{0, 1, \dots, q\}$ . Instead of Theorem 3.1 we have the following theorem (see also [19]):

**Theorem 4.2.** *We have  $x \in \mathcal{U}_q^*$  if and only if the following two lexicographic conditions are satisfied:*

$$\begin{aligned} (c_{n+i}) &< (q-1)^\infty \quad \text{whenever } c_n < q; \\ (q - c_{n+i}) &< (q-1)^\infty \quad \text{whenever } c_n > 0. \end{aligned}$$

PROOF. If  $(c_{n+i}) \geq (q-1)^\infty$  for some  $c_n < q$ , then another expansion of  $x$  is given by  $(d_i)$  where  $d_i = c_i$  for all  $i < n$ ,  $d_n = c_n + 1$ , and  $(d_{n+i})$  is an arbitrary expansion of

$$y := q^n \left( x - \frac{c_1}{q} - \dots - \frac{c_{n-1}}{q^{n-1}} - \frac{c_n + 1}{q^n} \right).$$

Such an expansion exists because

$$y = q^n \left( \left( \sum_{i=n+1}^{\infty} \frac{c_i}{q^i} \right) - \frac{1}{q^n} \right) = \left( \sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} \right) - 1 \in J_q^*;$$

the crucial inequality  $y \geq 0$  follows from the condition  $(c_{n+i}) \geq (q-1)^\infty$ . Indeed, if  $(c_{n+i}) = (q-1)^\infty$ , then we have

$$y = \frac{q-1}{q-1} - 1 = 0;$$

otherwise there is a first digit  $c_{n+m} = q$  and then

$$y \geq \left( \sum_{i=1}^m \frac{c_{n+i}}{q^i} \right) - 1 = 0.$$

Similarly, if  $(q - c_{n+i}) \geq (q-1)^\infty$  for some  $c_n > 0$ , then another expansion of  $x$  is given by  $(d_i)$  where  $d_i = c_i$  for all  $i < n$ ,  $d_n = c_n - 1$ , and  $(d_{n+i})$  is an arbitrary expansion of

$$z := q^n \left( x - \frac{c_1}{q} - \dots - \frac{c_{n-1}}{q^{n-1}} - \frac{c_n - 1}{q^n} \right).$$

Such an expansion exists because

$$z = q^n \left( \left( \sum_{i=n+1}^{\infty} \frac{c_i}{q^i} \right) + \frac{1}{q^n} \right) = \left( \sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} \right) + 1 \in J_q^*;$$

the crucial inequality  $z \leq \frac{q}{q-1}$  follows from the condition  $(q - c_{n+i}) \geq (q-1)^\infty$ . Indeed, if  $(q - c_{n+i}) = (q-1)^\infty$ , then we have

$$z = \frac{1}{q-1} + 1 = \frac{q}{q-1};$$

otherwise there is a first digit  $c_{n+m} = 0$  and then

$$z \leq \left( \sum_{i=1}^{m-1} \frac{1}{q^i} \right) + \left( \sum_{i=m+1}^{\infty} \frac{q}{q^i} \right) + 1 = \left( \sum_{i=1}^{\infty} \frac{1}{q^i} \right) + 1 = \frac{1}{q-1} + 1 = \frac{q}{q-1}.$$

Now assume that both lexicographic conditions are satisfied and let  $(d_i)$  be an arbitrary sequence on the alphabet  $\{0, \dots, q\}$ . We claim that if  $(d_i) \neq (c_i)$ , then

$$\sum_{i=1}^{\infty} \frac{d_i}{q^i} \neq \sum_{i=1}^{\infty} \frac{c_i}{q^i}.$$

To prove this we consider the first index  $n$  at which the sequences differ:  $d_i = c_i$  for all  $i < n$  but  $d_n \neq c_n$ .

If  $d_n > c_n$ , then  $c_n < q$ . It follows from our first lexicographic assumption that if  $c_n < q$  for some  $n$ , then  $c_i < q$  for all  $i > n$  too. Since, moreover, the equality  $(c_{n+i}) = (q-1)^\infty$  is also excluded, it follows that

$$\sum_{i=1}^{\infty} \frac{c_i}{q^i} < \left( \sum_{i=1}^n \frac{c_i}{q^i} \right) + \left( \sum_{i=n+1}^{\infty} \frac{q-1}{q^i} \right) = \frac{c_1}{q} + \dots + \frac{c_{n-1}}{q^{n-1}} + \frac{c_n+1}{q^n} \leq \sum_{i=1}^{\infty} \frac{d_i}{q^i}.$$

If  $d_n < c_n$ , then  $c_n > 0$ . It follows from our second lexicographic assumption that if  $c_n > 0$  for some  $n$ , then  $c_i > 0$  for all  $i > n$  too. Since, moreover, the equality  $(q - c_{n+i}) = (q-1)^\infty$ , i.e.,  $(c_{n+i}) = 1^\infty$  is also excluded, it follows that

$$\begin{aligned} \sum_{i=1}^{\infty} \frac{c_i}{q^i} &> \left( \sum_{i=1}^n \frac{c_i}{q^i} \right) + \left( \sum_{i=n+1}^{\infty} \frac{1}{q^i} \right) \\ &= \frac{c_1}{q} + \dots + \frac{c_{n-1}}{q^{n-1}} + \frac{c_n-1}{q^n} + \left( \sum_{i=n+1}^{\infty} \frac{q}{q^i} \right) \geq \sum_{i=1}^{\infty} \frac{d_i}{q^i}. \quad \square \end{aligned}$$

*Remarks 4.3.*

- (a) Since the sequence  $(q-1)^\infty$  does not satisfy the first condition (we have equalities instead of strict inequalities), the expansion of  $x = 1$  is not unique in base  $q$  any more if we use the enlarged alphabet  $\{0, 1, \dots, q\}$  instead of the earlier digit set  $\{0, 1, \dots, q-1\}$ . Indeed, it is straightforward to check that  $x = 1$  has for example the different expansions

$$\frac{q-1}{q} + \frac{q-1}{q^2} + \frac{q-1}{q^3} + \dots = \frac{q}{q} + \frac{0}{q^2} + \frac{0}{q^3} + \dots$$



(b) The list of all expansions of  $x = 1$  in integer bases for the enlarged alphabet has been determined in [16]:

– in base  $q = 2$  the expansions of  $x = 1$  are  $1^\infty$  and the sequences

$$1^n 20^\infty \quad \text{and} \quad 1^n 02^\infty, \quad n = 0, 1, \dots;$$

– in bases  $q = 3, 4, \dots$  the expansions of  $x = 1$  are  $(q - 1)^\infty$  and the sequences

$$(q - 1)^n q 0^\infty, \quad n = 0, 1, \dots$$

Now, analogously to the preceding section, we write  $x \in \mathcal{V}_q^*$  if, using the notation  $(a_i(x))$  for the quasi-greedy expansion of  $x$ ,

$$(q - a_{n+i}(x)) \leq (q - 1)^\infty \quad \text{whenever} \quad a_n(x) > 0$$

or equivalently

$$(a_{n+i}(x)) \geq 1^\infty \quad \text{whenever} \quad a_n(x) > 0.$$

In the next two sections we clarify the structure of these new sets  $\mathcal{U}_q^*$  and  $\mathcal{V}_q^*$ . As we will see, the cases  $q = 2$  and  $q \geq 3$  are rather different.

It will be convenient to use in the sequel the following notations: we denote by  $\mathcal{U}_q'^*$  (resp. by  $\mathcal{V}_q'^*$ ) the set of quasi-greedy expansions of the numbers of  $\mathcal{U}_q^*$  (resp. of  $\mathcal{V}_q^*$ ) on the alphabet  $\{0, \dots, q\}$ . It follows from Theorem 4.2 and from the definition of  $\mathcal{V}_q^*$  that a sequence  $(c_i)$  on the alphabet  $\{0, \dots, q\}$  belongs to  $\mathcal{U}_q'^*$  if and only if

$$(c_{n+i}) < (q - 1)^\infty \quad \text{whenever} \quad c_n < q \tag{4.1}$$

and

$$(c_{n+i}) > 1^\infty \quad \text{whenever} \quad c_n > 0, \tag{4.2}$$

and it belongs to  $\mathcal{V}_q'^*$  if and only if

$$(c_{n+i}) \leq (q - 1)^\infty \quad \text{whenever} \quad c_n < q \tag{4.3}$$

and

$$(c_{n+i}) \geq 1^\infty \quad \text{whenever} \quad c_n > 0. \tag{4.4}$$

## 5. Unique expansions in base $q = 2$ with digits 0, 1, 2

If we add the new digit 2 in base  $q = 2$ , then we have the following variant of Corollary 3.5:

**Proposition 5.1.** *We consider the expansions in base  $q = 2$  on the alphabet  $\{0, 1, 2\}$ .*

(a) We have

$$\mathcal{U}_2^* = \{0, 2\} \quad \text{and} \quad \mathcal{V}_2^* = \{0, 1, 2\} \cup \{2^{-N}, 2 - 2^{-N} : N = 1, 2, \dots\}.$$

Hence both  $\mathcal{U}_2^*$  and  $\mathcal{V}_2^*$  are countable closed sets, and  $\mathcal{V}_2^* \setminus \mathcal{U}_2^*$  is a countable discrete set, dense in  $\mathcal{V}_2^*$ .

(b) The set of expansions of each  $x \in \mathcal{V}_2^* \setminus \mathcal{U}_2^*$  is countably infinite.

The following proof will also provide the lists of all expansions in part (b).

PROOF.

(a) By definition (see (4.3)–(4.4))  $\mathcal{V}_2^*$  consists of the sequences  $(c_i)$  on the alphabet  $\{0, 1, 2\}$  satisfying

$$(c_{n+i}) \leq 1^\infty \quad \text{whenever} \quad c_n < 2$$

and

$$(c_{n+i}) \geq 1^\infty \quad \text{whenever} \quad c_n > 0.$$

Hence if  $(c_i) \in \mathcal{V}_2^*$ , then each 1 digit is followed by another 1 digit, and  $(c_i)$  cannot contain both 0 and 2 digits. This leaves us only the possibilities  $0^\infty$ ,  $1^\infty$ ,  $2^\infty$ ,  $0^N 1^\infty$  ( $N = 1, 2, \dots$ ) and  $2^N 1^\infty$  ( $N = 1, 2, \dots$ ), and a direct inspection shows that they belong to  $\mathcal{V}_2^*$  indeed. Hence

$$\mathcal{V}_2^* = \{0, 2\} \cup \{2^{-N}, 2 - 2^{-N} : N = 0, 1, \dots\}$$

which is equivalent to the first statement on  $\mathcal{V}_2^*$ .

It is clear that  $0, 2 \in \mathcal{U}_2^*$ . The other elements of  $\mathcal{V}_2^*$  are not univoque. Indeed, for any fixed  $N = 0, 1, \dots$ , two different expansions of  $2^{-N}$  and of  $2 - 2^{-N}$  are given by the equalities

$$\frac{1}{2^N} = \frac{1}{2^{N+1}} + \frac{1}{2^{N+2}} + \dots$$

and

$$\begin{aligned} \frac{2}{2} + \dots + \frac{2}{2^{N-1}} + \frac{1}{2^N} + \frac{2}{2^{N+1}} + \frac{2}{2^{N+2}} + \dots \\ = \frac{2}{2} + \dots + \frac{2}{2^{N-1}} + \frac{2}{2^N} + \frac{1}{2^{N+1}} + \frac{1}{2^{N+2}} + \dots \end{aligned}$$

Hence  $\mathcal{U}_2^* = \{0, 2\}$ .

The rest of part (a) readily follows from these explicit descriptions.

(b) We recall from Remark 4.3 (b) that the list of all expansions of  $x = 1$  is given by  $1^\infty$  and the sequences

$$1^n 20^\infty \quad \text{and} \quad 1^n 02^\infty, \quad n = 0, 1, \dots$$

Now fix a positive integer  $N$ . Every expansion  $(c_i)$  of  $x = 2^{-N}$  must begin with  $c_1 = \dots = c_{N-1} = 0$  and  $c_N \in \{0, 1\}$ . If  $c_N = 1$ , then we have necessarily  $c_i = 0$  for all  $i \neq N$ . Otherwise  $c_{N+1}, c_{N+2}, \dots$  is an expansion of  $x = 1$ . Using the list of expansions of  $x = 1$  we conclude that the list of all expansions of  $x = 2^{-N}$  is given by  $0^{N-1}10^\infty$ ,  $0^N1^\infty$  and the sequences

$$0^N1^n20^\infty \quad \text{and} \quad 0^N1^n02^\infty, \quad n = 0, 1, \dots$$

Furthermore,  $(c_i)$  is an expansion of  $2 - 2^{-N}$  if and only if  $(2 - c_i)$  is an expansion of  $2^{-N}$ . Using this correspondence we conclude from the preceding result that the list of all expansions of  $x = 2 - 2^{-N}$  is given by  $2^{N-1}12^\infty$ ,  $2^N1^\infty$  and the sequences

$$2^N1^n02^\infty \quad \text{and} \quad 2^N1^n20^\infty, \quad n = 0, 1, \dots$$

In particular, the set of expansions of each  $x \in \mathcal{V}_2^* \setminus \mathcal{U}_2^*$  is countably infinite.  $\square$

## 6. Unique expansions in bases $q = 3, 4, \dots$ with digits $0, 1, \dots, q$

For  $q = 3, 4, \dots$  we have the following variant of Corollary 3.5 for the enlarged digit set  $\{0, 1, \dots, q\}$ . We recall that a Cantor set is a nonempty set having neither interior, nor isolated points.

**Theorem 6.1.** *Let  $q \geq 3$  be an integer and consider the expansions in base  $q$  on the alphabet  $\{0, 1, \dots, q\}$ .*

- (a)  $\overline{\mathcal{U}}_q^* = \mathcal{V}_q^*$  is a Cantor set, and  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  is a countable dense set in  $\mathcal{V}_q^*$ .
- (b) The set of expansions of each  $x \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$  is countably infinite.
- (c) Both  $\mathcal{U}_q^*$  and  $\mathcal{V}_q^*$  have Hausdorff dimension  $\frac{\log(q-1)}{\log q}$  and hence zero Lebesgue measure.

*Remarks 6.2.*

- (a) The results differ very much from Corollary 3.5. Indeed, for the original digit set  $\{0, 1, \dots, q-1\}$ 
  - $\overline{\mathcal{U}}_q = \mathcal{V}_q = [0, 1]$  is not a Cantor set because it has interior points;
  - $\mathcal{U}_q$  and  $\mathcal{V}_q$  have full Lebesgue measure and Hausdorff dimension one;
  - each  $x \in \mathcal{V}_q \setminus \mathcal{U}_q$  has exactly two expansions.
- (b) The proof of the theorem will also provide the lists of all expansions in part (b) of the theorem.

**Lemma 6.3.**

- (a)  $\mathcal{V}_q^*$  consists of  $0^\infty$ ,  $q^\infty$  and of the sequences on the alphabet  $\{1, \dots, q-1\}$  preceded by  $0^n$  or  $q^n$  for some nonnegative integer  $n$ .
- (b)  $\mathcal{U}_q^*$  consists of the elements of  $\mathcal{V}_q^*$  which do not end with  $1^\infty$  or  $(q-1)^\infty$ .
- (c)  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  consists of the elements of  $\mathcal{V}_q^*$  which end with  $1^\infty$  or  $(q-1)^\infty$ .

PROOF. (a) If a sequence  $(c_i)$  on the alphabet  $\{0, 1, \dots, q\}$  belongs to  $\mathcal{V}_q^*$ , then it follows from (4.3)–(4.4) that

- after a nonzero digit we never have a zero digit;
- after a digit different from  $q$  we never have a digit  $q$ .

This leaves only the candidates  $0^\infty$ ,  $q^\infty$  and  $0^n c_{n+1} c_{n+2} \dots$  and  $q^n c_{n+1} c_{n+2} \dots$  with  $n = 0, 1, \dots$  and  $0 < c_i < q$  for all  $i > n$ . Conversely, all these sequences satisfy (4.3)–(4.4), so that they belong to  $\mathcal{V}_q^*$  indeed.

(b) The comparison of conditions (4.1)–(4.2) and (4.3)–(4.4) shows that  $\mathcal{U}_q^* \subset \mathcal{V}_q^*$ . The assertion now follows from part (a) and from conditions (4.1)–(4.2).

(c) This follows from parts (a) and (b).  $\square$

PROOF OF THEOREM 6.1. (a) In order to prove that  $\mathcal{V}_q^*$  is closed we show that  $J_q^* \setminus \mathcal{V}_q^*$  is open. Since  $\mathcal{V}_q^*$  is symmetric with respect to the midpoint of  $J_q^*$  by its definition, it is sufficient to find for each  $x \in J_q^* \setminus \mathcal{V}_q^*$  a point  $y < x$  such that  $(y, x] \cap \mathcal{V}_q^* = \emptyset$ . So fix  $x \in J_q^* \setminus \mathcal{V}_q^*$  arbitrarily and let  $(a_i)$  be its quasi-greedy expansion. Since  $x \notin \mathcal{V}_q^*$ , there exists  $n$  such that  $a_n > 0$  and  $(a_{n+i}) < 1^\infty$ . Fix  $m > n$  such that  $a_m = 0$  and set

$$y := \sum_{i=1}^m \frac{a_i}{q^i}.$$

Since  $(a_i)$  is an infinite sequence, we have  $y < x$ , and the quasi-greedy expansion of every  $z \in (y, x]$  begins with  $a_1 \dots a_m$ . Hence  $(a_{n+i}(z)) < 1^\infty$  and therefore  $z \notin \mathcal{V}_q^*$ .

It follows from Lemma 6.3 that  $\mathcal{U}_q^* \subset \mathcal{V}_q^*$  and that  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  is countably infinite.

Next we show that both  $\mathcal{U}_q^*$  and  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  are dense in  $\mathcal{V}_q^*$ . Given  $(c_i) \in \mathcal{V}_q^*$  and

$$x := \sum_{i=1}^{\infty} \frac{c_i}{q^i} \in \mathcal{V}_q^*$$

arbitrarily, it follows from Lemma 6.3 that the formulae

$$x_k := \frac{c_1}{q} + \dots + \frac{c_k}{q^k} + \frac{1}{q^{k+1}} + \frac{q-1}{q^{k+2}} + \frac{1}{q^{k+3}} + \frac{q-1}{q^{k+4}} + \dots,$$

$$y_k := \frac{c_1}{q} + \cdots + \frac{c_k}{q^k} + \frac{1}{q^{k+1}} + \frac{1}{q^{k+2}} + \frac{1}{q^{k+3}} + \frac{1}{q^{k+4}} + \cdots$$

and

$$z_k := \frac{c_1}{q} + \cdots + \frac{c_k}{q^k} + \frac{q-1}{q^{k+1}} + \frac{q-1}{q^{k+2}} + \frac{q-1}{q^{k+3}} + \frac{q-1}{q^{k+4}} + \cdots$$

define three sequences  $(x_k) \in \mathcal{U}_q^*$  and  $(y_k), (z_k) \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$  converging to  $x$ . This shows that both  $\mathcal{U}_q^*$  and  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  are dense in  $\mathcal{V}_q^*$ .

It follows also from this proof that  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  has no isolated points. Indeed, if  $x \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$ , then  $(c_i)$  ends either with  $1^\infty$  or  $(q-1)^\infty$ . In the first case  $z_k \neq x$  for all  $k$ , while in the second case  $y_k \neq x$  for all  $k$ . Hence its closure  $\mathcal{V}_q^*$  has no isolated points either.

For the Cantor property it remains to prove that  $\mathcal{V}_q^*$  has no interior points. Consider an element  $(c_i)$  of  $\mathcal{V}_q^*$  for which  $1 \leq c_i \leq q-1$  for all  $i > n$ . If we insert between  $c_k$  and  $c_{k+1}$  a zero digit, then for every  $k > n$  we obtain the quasi-greedy expansion of a number  $x_k$  (by Remark 4.1) which does not belong to  $\mathcal{V}_q^*$  by Lemma 6.3 (a). Since

$$x_k \rightarrow x := \frac{c_1}{q} + \frac{c_2}{q^2} + \cdots \in \mathcal{V}_q^*,$$

this shows that  $x$  is not an interior point of  $\mathcal{V}_q^*$ . We have thus shown that the interior of  $\mathcal{V}_q^*$  is a subset of  $\left\{0, \frac{q}{q-1}\right\}$ . If one of these two points would belong to the interior of  $\mathcal{V}_q^*$ , then it would be an isolated point of  $\mathcal{V}_q^*$ , which we have already excluded. Hence  $\mathcal{V}_q^*$  has no interior points.

(b) Let  $(c_i) \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$  be the quasi-greedy expansion of some  $x \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$ . Then by Lemma 6.3 (c) there exist two integers  $m \geq k \geq 0$  such that  $0 < c_i < q$  for  $i = k+1, \dots, m$ , and that the remaining digits satisfy one of the following four conditions:

$$c_i = 0 \quad \text{for all } i \leq k \quad \text{and} \quad c_i = q-1 \quad \text{for all } i > m;$$

$$c_i = q \quad \text{for all } i \leq k \quad \text{and} \quad c_i = q-1 \quad \text{for all } i > m;$$

$$c_i = 0 \quad \text{for all } i \leq k \quad \text{and} \quad c_i = 1 \quad \text{for all } i > m;$$

$$c_i = q \quad \text{for all } i \leq k \quad \text{and} \quad c_i = 1 \quad \text{for all } i > m.$$

By taking the minimal possible value of  $m$  we may also assume that in case  $m > 1$  we have  $c_m \neq c_{m+1}$ .

First we consider the cases where  $c_{m+1} = q-1$ . Let  $(d_i)$  be an arbitrary sequence on the alphabet  $\{0, 1, \dots, q\}$ . Using the equality

$$\frac{q-1}{q} + \frac{q-1}{q^2} + \cdots = 1$$

we see that

$$\frac{d_1}{q} + \frac{d_2}{q^2} + \dots = x$$

can only happen if  $d_i = c_i$  for all  $i < m$ , and  $d_m \in \{c_m, c_m + 1\}$ . Indeed:

- if  $d_i = c_i$  for all  $i < n$  and  $d_n < c_n$  for some  $n \leq m$ , then

$$\begin{aligned} \frac{d_1}{q} + \frac{d_2}{q^2} + \dots &\leq \frac{c_1}{q} + \dots + \frac{c_{n-1}}{q^{n-1}} + \frac{c_n - 1}{q^n} + \frac{q-1}{q^{n+1}} + \frac{q-1}{q^{n+2}} + \dots \\ &= \frac{c_1}{q} + \dots + \frac{c_{n-1}}{q^{n-1}} + \frac{c_n}{q^n} < x; \end{aligned}$$

- if  $d_i = c_i$  for all  $i < n$  and  $d_n > c_n$  for some  $n < m$ , then

$$\begin{aligned} \frac{d_1}{q} + \frac{d_2}{q^2} + \dots &\geq \frac{c_1}{q} + \dots + \frac{c_{n-1}}{q^{n-1}} + \frac{c_n + 1}{q^n} \\ &= \frac{c_1}{q} + \dots + \frac{c_n}{q^n} + \frac{q-1}{q^{n+1}} + \frac{q-1}{q^{n+2}} + \dots \\ &> \frac{c_1}{q} + \dots + \frac{c_m}{q^m} + \frac{q-1}{q^{m+1}} + \frac{q-1}{q^{m+2}} + \dots \\ &= x; \end{aligned}$$

- if  $d_i = c_i$  for all  $i < m$  and  $d_m \geq c_m + 1$ , then

$$\begin{aligned} \frac{d_1}{q} + \frac{d_2}{q^2} + \dots &\geq \frac{c_1}{q} + \dots + \frac{c_{m-1}}{q^{m-1}} + \frac{c_m + 1}{q^m} \\ &= \frac{c_1}{q} + \dots + \frac{c_m}{q^m} + \frac{q-1}{q^{m+1}} + \frac{q-1}{q^{m+2}} + \dots \\ &= x, \end{aligned}$$

with equality only if  $d_m = c_m + 1$  and  $d_i = 0$  for all  $i > m$ . This is only possible if  $c_m < q$ .

Apart from this last one, all the other expansions  $(d_i)$  of  $x$  start with  $c_1, \dots, c_m$ , so that

$$\frac{d_{m+1}}{q^{m+1}} + \frac{d_{m+2}}{q^{m+2}} + \dots = \frac{c_{m+1}}{q^{m+1}} + \frac{c_{m+2}}{q^{m+2}} + \dots$$

whence

$$\frac{d_{m+1}}{q} + \frac{d_{m+2}}{q^2} + \dots = \frac{q-1}{q} + \frac{q-1}{q^2} + \dots = 1.$$

Using Remark 4.3 (b) we conclude that the list of expansions of  $x$  is as follows:

- $c_1 \dots c_m (q-1)^\infty$ ;

- $c_1 \dots c_m (q-1)^n q 0^\infty$ ,  $n = 0, 1, \dots$ ;
- $c_1 \dots c_{m-1} (c_m + 1) 0^\infty$  if  $c_m < q$ .

Since  $\mathcal{V}_q^*$  is symmetric with respect to its midpoint and since  $(c_i)$  is an expansion of  $x$  if and only if  $(q - c_i)$  is an expansion of the reflection of  $x$ , the case  $c_{m+1} = 1$  follows from the preceding one. We conclude that the list of expansions of  $x$  is now as follows:

- $c_1 \dots c_m 1^\infty$ ;
- $c_1 \dots c_m 1^n 0 q^\infty$ ,  $n = 0, 1, \dots$ ;
- $c_1 \dots c_{m-1} (c_m - 1) q^\infty$  if  $c_m < q$ .

In particular, we have proved that the set of expansions of each  $x \in \mathcal{V}_q^* \setminus \mathcal{U}_q^*$  is countably infinite.

(c) Since  $\mathcal{V}_q^* \setminus \mathcal{U}_q^*$  is countable,  $\mathcal{U}_q^*$  and  $\mathcal{V}_q^*$  have the same Hausdorff dimension. Since  $\mathcal{V}_q^*$  is the union of the two points  $0$ ,  $\frac{q}{q-1}$  and of countably many sets, each of which is similar to the set

$$\mathcal{Z} := \left\{ \sum_{i=1}^{\infty} \frac{c_i}{q^i} : c_i \in \{1, \dots, q-1\}, i = 1, 2, \dots \right\},$$

we have  $\dim_{\text{H}} \mathcal{U}_q^* = \dim_{\text{H}} \mathcal{V}_q^* = \dim_{\text{H}} \mathcal{Z}$ .

Let us compute the similarity dimension  $s$  of  $\mathcal{Z}$ . Since  $\mathcal{Z}$  is the attractor of the iterated function system defined by

$$f_k(x) := \frac{k+x}{q}, \quad x \in \mathcal{Z}, \quad k = 1, \dots, q-1,$$

we have  $(q-1)q^{-s} = 1$  whence  $s = \frac{\log(q-1)}{\log q}$ .

The images  $f_k(\mathcal{Z})$  have disjoint closures. Indeed, if  $k < n$ , then

$$\sup f_k(\mathcal{Z}) = \frac{k}{q} + \frac{1}{q} \sum_{i=1}^{\infty} \frac{q-1}{q^i} = \frac{k+1}{q}$$

and

$$\inf f_n(\mathcal{Z}) = \frac{n}{q} + \frac{1}{q} \sum_{i=1}^{\infty} \frac{1}{q^i} = \frac{n}{q} + \frac{1}{q(q-1)},$$

so that

$$\inf f_n(\mathcal{Z}) - \sup f_k(\mathcal{Z}) = \frac{n-k-1}{q} + \frac{1}{q(q-1)} \geq \frac{1}{q(q-1)} > 0.$$

Since Moran's open set condition is thus satisfied, the Hausdorff dimension of  $\mathcal{Z}$  is equal to its similarity dimension, so that

$$\dim_{\text{H}} \mathcal{U}_q^* = \dim_{\text{H}} \mathcal{V}_q^* = \dim_{\text{H}} \mathcal{Z} = \frac{\log(q-1)}{\log q}.$$

As sets of Hausdorff dimension  $< 1$ , all these sets have zero Lebesgue measure.  $\square$

## References

- [1] C. BAIocchi and V. KOMORNIK, Greedy and quasi-greedy expansions in non-integer bases arXiv: 0710.3001 [math.], October 16, 2007.
- [2] Z. DARÓCZY and I. KÁTAI, Univoque sequences, *Publ. Math. Debrecen* **42** (1993), 397–407.
- [3] Z. DARÓCZY and I. KÁTAI, On the structure of univoque numbers, *Publ. Math. Debrecen* **46** (1995), 385–408.
- [4] M. DE VRIES and V. KOMORNIK, Unique expansions of real numbers, *Adv. Math.* **221** (2009), 390–427.
- [5] P. ERDŐS, M. HORVÁTH and I. JOÓ, On the uniqueness of the expansions  $1 = \sum q^{-n_i}$ , *Acta Math. Hungar.* **58** (1991), 333–342.
- [6] P. ERDŐS, I. JOÓ and V. KOMORNIK, Characterization of the unique expansions  $1 = \sum_{i=1}^{\infty} q^{-n_i}$  and related problems, *Bull. Soc. Math. France* **118**, no. 3 (1990), 377–390.
- [7] P. ERDŐS, I. JOÓ and V. KOMORNIK, On the number of  $q$ -expansions, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **37** (1994), 109–118.
- [8] P. ERDŐS, I. JOÓ and V. KOMORNIK, On the sequence of numbers of the form  $\varepsilon_0 + \varepsilon_1 q + \cdots + \varepsilon_n q^n$ ,  $\varepsilon_i \in \{0, 1\}$ , *Acta Arith.* **83**, no. 3 (1998), 201–210.
- [9] P. ERDŐS and V. KOMORNIK, Developments in non-integer bases, *Acta Math. Hungar.* **79**, no. 1–2 (1998), 57–83.
- [10] P. GLENDINNING and N. SIDOROV, Unique representations of real numbers in non-integer bases, *Math. Res. Lett.* **8**, no. 4 (2001), 535–543.
- [11] G. KALLÓS, The structure of the univoque set in the small case, *Publ. Math. Debrecen* **54** (1999), 153–164.
- [12] G. KALLÓS, The structure of the univoque set in the big case, *Publ. Math. Debrecen* **59** (2001), 471–489.
- [13] G. KALLÓS and I. KÁTAI, I., On the set for which 1 is univoque, *Publ. Math. Debrecen* **58** (2001), 743–750.
- [14] V. KOMORNIK and P. LORETI, Unique developments in non-integer bases, *Amer. Math. Monthly* **105**, no. 7 (1998), 636–639.
- [15] V. KOMORNIK and P. LORETI, Subexpansions, superexpansions and uniqueness properties in non-integer bases, *Period. Math. Hungar.* **44**, no. 2 (2002), 195–216.
- [16] V. KOMORNIK and P. LORETI, On the topological structure of univoque sets, *J. Number Theory* **122** (2007), 157–183.
- [17] V. KOMORNIK, P. LORETI and A. PETHŐ, The smallest univoque number is not isolated, *Publ. Math. Debrecen* **62** (2003), 429–435.
- [18] W. PARRY, On the  $\beta$ -expansions of real numbers, *Acta Math. Acad. Sci. Hungar.* **11** (1960), 401–416.
- [19] M. PEDICINI, Greedy expansions and sets with deleted digits, *Theoret. Comput. Sci.* **332**, no. 1–2 (2005), 313–336.



- [20] A. RÉNYI, Representations for real numbers and their ergodic properties, *Acta Math. Acad. Sci. Hungar.* **8** (1957), 477-493.

MARTINE BAATZ  
LYCÉE TECHNIQUE MICHEL LUCIUS  
157, AV. PASTEUR  
L-2311 LUXEMBOURG

*E-mail:* martine.batz@education.lu

VILMOS KOMORNIK  
DÉPARTEMENT DE MATHÉMATIQUE  
UNIVERSITÉ DE STRASBOURG  
7 RUE RENÉ DESCARTES  
67000 STRASBOURG  
FRANCE

*E-mail:* vilmos.komornik@math.unistra.fr

*(Received November 15, 2010; revised April 29, 2011)*