

## The structure of the univoque set in the big case

By GÁBOR KALLÓS (Győr)

**Abstract.** Let  $\beta > 1$ ,  $\Theta = 1/\beta$ ,  $\mathcal{D} = \{0, 1, \dots, [\beta]\}$ . In this paper we continue the investigation of the numbers which have only one expansion in the form  $\sum_{n=1}^{\infty} \varepsilon_n \Theta^n$  with  $\varepsilon \in \mathcal{D}^{\mathbb{N}}$ . We present a method for the determination of the Hausdorff dimension of the set of these numbers in the so-called big case, illustrated with interesting examples.

### 1. Introduction

Let  $\beta > 1$  be the base number of a (number) system,  $\Theta = \frac{1}{\beta}$ ,  $[\beta] = k$  and  $\mathcal{D} = \{0, 1, \dots, [\beta]\}$  the set of the digits. For an infinite sequence  $\varepsilon \in \mathcal{D}^{\mathbb{N}}$  let moreover  $\langle \varepsilon, \Theta \rangle = \sum_{n=1}^{\infty} \varepsilon_n \Theta^n$ . For every  $x \in [0, L]$ , where  $L = k\Theta + k\Theta^2 + \dots = \frac{k\Theta}{1-\Theta}$ , there exists at least one sequence  $\delta$  for which  $\langle \delta, \Theta \rangle = x$ . We can use e.g. the regular expansion  $x = \varepsilon_1 \Theta + \varepsilon_2 \Theta^2 + \dots$  with the iteration of the rule

$$x = \varepsilon_1(x)\Theta + \Theta x_1, \quad \varepsilon_1(x) = [\beta x], \quad x_1 = \{\beta x\},$$

here  $\varepsilon_i(x) \in \mathcal{D}$ . The quasiregular expansion of some  $x \in (0, L]$ ,  $x = \delta_1 \Theta + \delta_2 \Theta^2 + \dots = \delta_1 \Theta + \Theta x_1$  is defined as follows:  $\delta_1$  is the largest integer  $d \in \mathcal{D}$  for which  $x - d\Theta > 0$  and  $x_1 = \beta x - d$ .

For a set  $C$  and a word  $\beta = \beta_1 \dots \beta_p$  let  $\langle C, \Theta \rangle = \{\langle c, \Theta \rangle \mid c \in C\}$ , and

$$\langle \beta C, \Theta \rangle = \beta_1 \Theta + \dots + \beta_p \Theta^p + \Theta^p \langle C, \Theta \rangle.$$

For some fixed  $\Theta$  the number  $x \in [0, L]$  is said to be univoque, if  $x$  has a unique expansion in the form  $x = \langle \varepsilon, \Theta \rangle$ ,  $\varepsilon \in \mathcal{D}^{\mathbb{N}}$ , i.e. if  $x = \langle \varepsilon, \Theta \rangle = \langle \delta, \Theta \rangle$ , then  $\delta = \varepsilon$ . In this case the sequence  $\varepsilon$  is said to be univoque, too.

---

*Mathematics Subject Classification:* 11A67.

*Key words and phrases:* number theory, expansions of numbers, univoque sequences.

Let  $1 = t_1\Theta + t_2\Theta^2 + \dots$  be the quasiregular expansion of 1. Let  $T = t_1t_2\dots$  and  $T_p = t_1t_2\dots t_p$ . For a sequence  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots) \in \mathcal{D}^{\mathbb{N}}$  we use the usual cut and shift operations, respectively:  $F_n(\varepsilon) = \varepsilon_1 \dots \varepsilon_n$  and  $\sigma^n(\varepsilon) = (\varepsilon_{n+1}, \varepsilon_{n+2}, \dots)$ . We shall use the lexicographic ordering. We shall write  $\bar{a} = k - a$  for  $a \in \mathcal{D}$ , and  $\bar{\varepsilon} = (\bar{\varepsilon}_1, \bar{\varepsilon}_2, \dots)$  for  $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots)$ . Furthermore for some set  $C \subseteq \mathcal{D}^{\mathbb{N}}$  let  $\bar{C} = \{\bar{c} \mid c \in C\}$ . If  $\langle \varepsilon, \Theta \rangle = x$ , then  $\langle \bar{\varepsilon}, \Theta \rangle = L - x$ , and so the number  $x$  is univoque if and only if  $L - x$  is univoque.

The structure of the univoque numbers was investigated for  $k = 1$  by Z. DARÓCZY and I. KÁTAI in [1] and [2]. We shall examine the case  $k \geq 2$ . Let

$$\Theta_k = \frac{-k + \sqrt{k^2 + 4}}{2}$$

and  $\beta_k = k + \Theta_k$ . The univoque set is quite simple if  $k < \beta < \beta_k$  (“small case”). In [4] we have presented a method for the computation of the Hausdorff dimension of the univoque set in this simpler case. We call the case  $k + 1 > \beta > \beta_k$  as the “big case”, since the fractional part of  $\beta$  is larger than  $\Theta_k$ . In this case we do not have till yet general results, but we are able to determine the dimension of the univoque set in every specific system ([5]). Our purpose in this paper is to present a general method in the big case.

## 2. General results in the big case

Now  $k + \Theta < \beta$  (since  $\Theta_k > \Theta$ ), thus  $k\Theta + \Theta^2 < 1$  and  $L < \Theta + 1$ . Consequently, if we specify the univoque numbers in the interval  $I = (L - 1, 1)$ , we can cover the whole interval  $(0, 1)$  – and so specify the univoque numbers of  $(0, 1)$  – using that

$$(0, 1) \subseteq \bigcup_{n=0}^{\infty} \Theta^n(L - 1, 1).$$

Since  $\Theta^{i+1} > \Theta^i(L - 1)$ , the intervals  $(\Theta^{i+1}(L - 1), \Theta^{i+1})$  and  $(\Theta^i(L - 1), \Theta^i)$  have a nonempty intersection. However, it does not cause any problem: a univoque element has unique expansion, so we consider every univoque element only once. Thus, if we know the univoque numbers in the interval  $(L - 1, 1)$ , we can specify all univoque numbers in  $[0, L]$  with multiplication

by  $\Theta^i$  ( $i = 1, 2, \dots$ ) and – using the symmetry of the univoque set – with reflection.

It was already proved ([1], [4]), that

$$(1) \quad x = \langle \varepsilon, \Theta \rangle \text{ is univoque} \iff \varepsilon \text{ and } \bar{\varepsilon} \text{ are regular.}$$

To decide the regular property, we use a theorem of W. PARRY ([8], reformulated in [2] and [4]):

For a fixed  $\Theta$  the sequence  $\varepsilon \in \mathcal{D}^{\mathbb{N}}$  is the regular expansion of some  $x \in [0, 1)$  if and only if

$$(2) \quad \sigma^j(\varepsilon) < T \quad (j = 0, 1, 2, \dots).$$

He proved moreover, that  $T$  is a quasiregular expansion of 1 with  $[\beta] = k$  if and only if

$$(3) \quad \sigma^j(T) \leq T \quad (j = 1, 2, \dots), \quad T_1 = k.$$

In the non-periodic cases the equation in (3) is not allowed. According to condition (2), the set which contains exactly the univoque numbers in  $(L - 1, 1)$  is

$$(4) \quad K = \{\varepsilon \mid \sigma^s(\varepsilon) < T, \sigma^s(\bar{\varepsilon}) < T, s = 0, 1, \dots\}.$$

It is obvious that  $K = \overline{K}$ .

For easier treatment we would like to substitute the sequences with their finite parts. We can approach  $K$  with the following sets:

$$(5) \quad U_r = \{\varepsilon \mid F_r(\sigma^s(\varepsilon)) < T_r, F_r(\sigma^s(\bar{\varepsilon})) < T_r, s = 0, 1, \dots\},$$

$$(6) \quad V_r = \{\varepsilon \mid F_r(\sigma^s(\varepsilon)) \leq T_r, F_r(\sigma^s(\bar{\varepsilon})) \leq T_r, s = 0, 1, \dots\}.$$

The set  $U_{i+1}$  can not be smaller than  $U_i$ , since if  $F_i(\sigma^s(\varepsilon)) < T_i$  ( $s = 0, 1, \dots$ ), then  $F_{i+1}(\sigma^s(\varepsilon)) < T_{i+1}$  is true, and similarly  $F_i(\sigma^s(\bar{\varepsilon})) < T_i$  implies  $F_{i+1}(\sigma^s(\bar{\varepsilon})) < T_{i+1}$ . However, it can be larger, namely it can happen, that  $F_i(\sigma^s(\varepsilon)) \leq T_i$ , but  $F_{i+1}(\sigma^s(\varepsilon)) < T_{i+1}$  and the same holds for  $\bar{\varepsilon}$ , i.e.  $\varepsilon \notin U_i$ , but  $\varepsilon \in U_{i+1}$ . With similar arguments for  $V_i$ , we get eventually

$$U_1 \subseteq U_2 \subseteq \dots \subseteq K \subseteq \dots \subseteq V_2 \subseteq V_1.$$

We shall see, that we are able to apply this method in most of the cases. From conditions (1) and (3) follows

**Proposition 1.**  $1$  is univoque with respect to  $\Theta \iff \sigma^j(T) < T$  and  $\sigma^j(\overline{T}) < T$ , with  $j = 1, 2, \dots$

The expansion of  $1$  can not be periodic in this case. It was proved nowadays, that the Lebesgue measure of the set of these  $\Theta$ -s is  $0$  (I. KÁTAI and G. KALLÓS, [6]).

If  $1$  is not univoque, then we have  $t_i t_{i+1} \dots = T$  or  $t_i t_{i+1} \dots \leq \overline{T}$  for some index  $i \geq 2$ . The case  $t_i t_{i+1} \dots > T$  is not possible from condition (3). With the following theorems we investigate these cases.

**Theorem 1.** *If there exists an index  $u$  for which  $1 \leq u < p$  and  $t_{u+1} \dots t_p < \overline{T}_{p-u}$  then for all elements  $\varepsilon$  in  $K$*

$$\varepsilon_{k+1} \dots \varepsilon_{k+p} < T_p \text{ and } \overline{\varepsilon_{k+1} \dots \varepsilon_{k+p}} < T_p,$$

where  $k = 0, 1, \dots$  (i.e.  $K = U_p$ ).

PROOF. Assume the contrary, namely that there is an  $\varepsilon \in K$  for which one of the inequalities of the theorem fails to hold. Changing  $\varepsilon$  to  $\overline{\varepsilon}$ , if necessary, we may assume that the first inequality fails. From conditions in (4) it follows that  $\varepsilon_{k+1} \dots \varepsilon_{k+p} = T_p$ . Thus,  $\sigma^k(\varepsilon) \in K$  has prefix  $T_p$ , and  $\sigma^u(\overline{\sigma^k(\varepsilon)})$  has prefix  $\overline{t_{u+1} \dots t_p} > T_{p-u}$ , which is impossible.  $\square$

The proof of the following corollary is very similar to that of Theorem 1.

**Corollary 1.** *If there exists an  $\varepsilon \in V_p \setminus U_p$ , then for all  $u \geq 0$  we have  $t_{u+1} \dots t_p \geq \overline{T}_{p-u}$ .*

**Theorem 2.** *Let us assume that  $t_{u+1} \dots t_p \geq \overline{T}_{p-u}$  is satisfied for all  $0 \leq u < p$ . In this case either  $T$  is periodic, or  $\sigma^j(T) < T$  and  $\sigma^j(\overline{T}) < T$ , i.e.  $1$  is univoque.*

PROOF. a) From condition (3) we have  $\sigma^j(T) \leq T$ . If there exists an index  $j$ , for which  $\sigma^j(T) = T$ , then  $\sigma^{lj}(T) = T$ , ( $l = 1, 2, \dots$ ). Thus  $T = T_j T_j \dots$  holds.

b) Let us assume, that we have an index  $j$  for which  $\sigma^j(\overline{T}) = T$ . Then  $T = T_j \overline{T}$ , with another shift  $T = T_j \overline{T_j T}$ , i.e.  $T$  is periodic.

c) In the remaining cases  $\sigma^j(T) < T$  and  $\sigma^j(\overline{T}) < T$ .  $\square$

Thus, if  $1$  is not univoque (with respect to  $\Theta$ ), the following cases occur for some index  $i$ :

- a)  $t_i t_{i+1} \dots < \bar{T}$  – the case of Theorem 1.
- b)  $t_i t_{i+1} \dots = T$  – part a) in the proof of Theorem 2,  $T$  is periodic.
- c)  $t_i t_{i+1} \dots = \bar{T}$  – part b) in the proof of Theorem 2,  $T$  is periodic.

If  $T$  is periodic with period-length  $u$ , for sequences  $\varepsilon \in K$  clearly holds

$$\sigma^s(\varepsilon) < T \implies F_u(\sigma^s(\varepsilon)) \leq T_u,$$

so we have  $K = V_u$ , i.e. in cases b) and c) it is enough to investigate finite sequence parts, too. We call from now on the cases a), b) and c) together as the finite case. In this paper we work further with the *finite case* only. So there exist an index  $p$  for which  $K = U_p$  or  $K = V_p$ .

To give the structure of  $K$  when 1 is univoque is much harder, we hope to return to this question in another paper.

### 3. The structure of set $K$

*Reduction of  $T_p$ .* Our goal in this subsection is to gain a simpler  $T_p$ , and so simpler conditions in (5) and (6). Assume that for some  $\Theta$  we have  $K = U_p$ . It means that  $\varepsilon \in K$  if  $\varepsilon_{k+1} \dots \varepsilon_{k+p} < T_p$ ,  $\bar{\varepsilon}_{k+1} \dots \bar{\varepsilon}_{k+p} < T_p$  holds for every  $k = 0, 1, \dots$

*Case 1.* If  $t_p = 0$ , then these inequalities hold if and only if  $\varepsilon_{k+1} \dots \varepsilon_{k+p-1} < T_{p-1}$ ,  $\bar{\varepsilon}_{k+1} \dots \bar{\varepsilon}_{k+p-1} < T_{p-1}$ , and so  $K = U_{p-1}$ .

*Case 2.* Let  $t_p > 0$ ,  $T_p^* = t_1 \dots t_{p-1}(t_p - 1)$ . Then

$$(7) \quad K = \{\varepsilon \mid F_p(\sigma^s(\varepsilon)) \leq T_p^*, F_p(\sigma^s(\bar{\varepsilon})) \leq T_p^*\}.$$

There are two possibilities. Either for each  $u \in [0, p-1]$  holds  $t_{u+1} \dots t_{p-1} \times (t_p - 1) \geq \bar{T}_{p-u}$ , which implies that there exists at least one  $\varepsilon$  with prefix  $T_p^*$  (Corollary 1), or there is such an index  $u$  for which  $t_{u+1} \dots t_{p-1}(t_p - 1) < \bar{T}_{p-u}$ . In the former case we investigate  $K$  according to the inequalities in (7), while in the latter case we can change them to satisfy  $F_p(\sigma^s(\varepsilon)) < T_p^*$ ,  $F_p(\sigma^s(\bar{\varepsilon})) < T_p^*$ , and continue the reduction of the conditions given above. After finitely many steps we arrive to a non-reducible stage.

We keep  $T_p$  to denote the non-reducible case. Thus, after the procedure described above we have  $K = V_p$ , i.e.

$$(8) \quad (K(T_p) :=) K = \{\varepsilon \mid F_p(\sigma^s(\varepsilon)) \leq T_p, F_p(\sigma^s(\bar{\varepsilon})) \leq T_p\},$$

and

$$\bar{T}_{p-l} \leq t_{l+1} \dots t_p \leq T_{p-l} \quad (l = 0, \dots, p-1).$$

In some cases the investigation can be reduced further.

**Proposition 2.** If  $T_p = T_u T_v$  holds for some  $u, v \geq 1$ , then  $K(T_p) = K(T_u)$ , so  $T_p$  can be reduced to  $T_u$ .

PROOF. a)  $\subseteq$  – From conditions in (8) clearly follow  $F_u(\sigma^s(\varepsilon)) \leq T_u$ ,  $F_u(\sigma^s(\bar{\varepsilon})) \leq T_u$ , thus  $K(T_p) \subseteq K(T_u)$ .

b)  $\supseteq$  – *Case 1:*  $u \geq v$ . For some  $\varepsilon \in K(T_u)$  the inequalities

$$\varepsilon_{k+1} \dots \varepsilon_{k+u} \leq T_u, \quad \varepsilon_{k+u+1} \dots \varepsilon_{k+u+v} \leq T_v,$$

$$\overline{\varepsilon_{k+1}} \dots \overline{\varepsilon_{k+u}} \leq T_u, \quad \overline{\varepsilon_{k+u+1}} \dots \overline{\varepsilon_{k+u+v}} \leq T_v$$

obviously hold for every  $k$ , which implies that  $\varepsilon \in K(T_p)$ .

*Case 2:*  $u < v$ . Then

$$T_p = \underbrace{t_1 \dots t_u}_{T_u} \underbrace{t_1 \dots t_u t_{u+1} \dots t_v}_{T_v}.$$

Cutting the last  $u$  digits  $T_v = t_1 \dots t_u t_1 \dots t_{v-u}$ , i.e.  $T_p = T_u T_u T_{v-u} = T_{2u} T_{v-u}$ . If  $v - u \geq u$ , then similarly  $T_v = T_{2u} T_{v-2u}$ , i.e.  $T_p = T_{3u} T_{v-2u}$ . In general, if  $v = ru + s$ ,  $0 < s \leq u$ , then we write  $T_p = T_{ru} T_s$ , and – using case 1 – we obtain that  $K(T_p) = K(T_{ru})$ . Since  $T_{ru} = T_{(r-1)u} T_u$ , using case 1 we get  $K(T_{ru}) = K(T_{(r-1)u})$ . Continuing we obtain that  $K(T_p) = K(T_u)$ .  $\square$

**Proposition 3.** Assume that  $T_u \bar{T}_u$  is a prefix of  $T_p$ . Then exactly one element  $\xi$  of  $K = K(T_p)$  exists with prefix  $T_u$ , namely  $\xi = T_u \bar{T}_u T_u \bar{T}_u \dots$ . Thus, no more than countable many  $\varepsilon$  exists for which there exists at least one  $k$  with  $F_u(\sigma^k(\varepsilon)) = T_u$  or with  $F_u(\sigma^k(\bar{\varepsilon})) = T_u$ .

PROOF. Let  $\eta = T_u \eta_1 \in K$ . Since  $T_u \bar{T}_u = T_{2u}$ , therefore  $F_u(\eta_1) \leq \bar{T}_u$ , but  $F_u(\bar{\eta}_1) \leq T_u$ , which implies that  $\eta_1 = \bar{T}_u \eta_2$ . Since  $T_u$  is the prefix of  $\bar{\eta}_1$ , we obtain that  $\bar{\eta}_2 = \bar{T}_u \bar{\eta}_3$  i.e.  $\eta_2 = T_u \eta_3$ . Continuing, we obtain that  $\xi$  is the only element of  $K$  with prefix  $T_u$ . Similarly,  $\bar{\xi}$  is the only element of  $K$  with prefix  $\bar{T}_u$ .

To prove the second assertion, consider those sequences  $e = (e_1, e_2, \dots)$  for which  $e_{k+1} \dots e_{k+u} = T_u$ . Since  $\sigma^k(e) \in K$  and its prefix is  $T_u$ , therefore  $\sigma^k(e) = \xi$ . Similarly, if  $\sigma^k(\bar{e}) \in K$  and the prefix of  $\bar{e}$  is  $T_u$ , then  $\sigma^k(\bar{e}) = \bar{\xi}$ .

Let  $S$  be the set of the sequences  $\varepsilon$ , for which there exists a  $k$  with  $F_u(\sigma^k(\varepsilon)) = T_u$  or with  $F_u(\sigma^k(\bar{\varepsilon})) = T_u$ . Then set  $S$  can be estimated from above with

$$S \subseteq \sum_{r=0}^{\infty} a_1 a_2 \dots a_r \xi + a_1 a_2 \dots a_r \bar{\xi},$$

where  $a_i \in \{0, 1, \dots, k\}$ , but not every choices are good for  $a_1 a_2 \dots a_r$ . So  $S$  is a countable set.  $\square$

*Remark.* If  $T_{2u} = T_u \bar{T}_u$ ,  $2u \leq p$ , then the Hausdorff dimension of  $S = K(T_p) \setminus U_u$  is zero. Thus we can reduce  $T_p$  into  $T_u$ .

From now on we assume, that none of  $T_p = T_u T_v$ ,  $T_{2u} = T_u \bar{T}_u$  occurs.

*A partition of set  $K$ .* In the sequel we build a graph, which represents the structure of set  $K$ . We shall see later, that using this graph in most of the cases we are able to determine the dimension of the univoque set.

To achieve this, first we have to find for set  $K$  such a partition, the components of which are characterized by the prefixes of their elements. For a finite word  $\alpha$  let  $K_\alpha = \{\alpha\eta \mid \eta, \alpha\eta \in K\}$ . Let us introduce sets  $\mathcal{A}_i$  – containing finite words – in the following way:

$$\begin{aligned} \mathcal{A}_1 &= \{i \mid i = 1, \dots, k-1\}, \\ \mathcal{A}_2 &= \{t_1 i \mid 0 \leq i < t_2, K_{t_1 i} \neq \emptyset\}, \\ &\dots \\ (9) \quad \mathcal{A}_j &= \{T_{j-1} i \mid 0 \leq i < t_j, K_{T_{j-1} i} \neq \emptyset\}, \quad \text{where } j = 1, \dots, p \\ &\dots \\ \mathcal{A}_{p+1} &= \{T_p\}, \end{aligned}$$

moreover let

$$(10) \quad \mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \overline{\mathcal{A}_2} \cup \dots \cup \mathcal{A}_{p+1} \cup \overline{\mathcal{A}_{p+1}}.$$

*Remarks.* 1. Since  $\mathcal{A}_1 = \overline{\mathcal{A}_1}$ , it is enough to indicate exactly one of them in (10).

2. It is possible, that for some  $\alpha \in \mathcal{A}_i$  the set  $K_\alpha$  is empty. We removed these  $\alpha$ -s from the set  $\mathcal{A}_i$  with part  $K_{T_{j-1} i} \neq \emptyset$  in the definition.

The proof of the following two assertions is easy, it is left to the reader.

**Lemma 2.** *Let  $\alpha$  be an arbitrary word. Then  $\overline{K_\alpha} = K_{\overline{\alpha}}$  is satisfied.*

**Lemma 3.** *For words  $\alpha, \beta \in \mathcal{A}$ , if  $\alpha \neq \beta$  then  $K_\alpha \cap K_\beta = \emptyset$  holds.*

**Proposition 4.**  $K = \bigcup_{\alpha \in \mathcal{A}} K_\alpha$ .

PROOF. a)  $\supset$  – This part directly follows from the definition of  $K_\alpha$ .

b)  $\subset$  – Let now  $\varepsilon \in K$ .

(i) If  $\varepsilon_1 = i \neq 0, k$  then the sequence  $\varepsilon$  can be written in the form  $i\eta$ , where  $i \in \mathcal{A}_1$ . Then  $\varepsilon \in K_i$ .

(ii) If  $\varepsilon_1 = k$ , then either  $\varepsilon_1 \dots \varepsilon_p = T_p$  and then  $\varepsilon \in K_{T_p}$  ( $T_p \in \mathcal{A}_{p+1}$ ), or  $\varepsilon_1 \dots \varepsilon_p \neq T_p$ . In the latter case let  $s < p$  be the greatest index, for which  $\varepsilon_1 \dots \varepsilon_s = T_s$ . Since  $F_p(\sigma^j(\varepsilon)) \leq T_p$  for arbitrary index  $j$ , so  $\varepsilon_{s+1} < t_{s+1}$  ( $\varepsilon_{s+1} > t_{s+1}$  is not possible, and  $\varepsilon_{s+1} \neq t_{s+1}$ ). Thus  $\varepsilon_1 \dots \varepsilon_s \varepsilon_{s+1} \in \overline{\mathcal{A}_{s+1}}$ .

(iii) If  $\varepsilon_1 = 0$ , then similarly either  $\varepsilon \in K_{\overline{T_p}}$  or  $\varepsilon_1 \dots \varepsilon_{s+1} \in \overline{\mathcal{A}_{s+1}}$ .  $\square$

*Construction of the graph describing the univoque set.* Let  $\mathcal{G}(\mathcal{A})$  be the following directed graph. The nodes of the graph are the elements of  $\mathcal{A}$ .

a) Let  $i \in \mathcal{A}_1$ . We lead edges from  $i$  to every elements of  $\mathcal{A}$ . We label these edges by  $i$ .

b) Let  $\alpha = T_{j-1}i \in \mathcal{A}_j$ . Let  $u$  be the smallest integer for which  $\alpha = T_u \overline{T_{j-u}}$ , if such an index exists. Then we lead edges from  $\alpha$  to the elements of  $\bigcup_{l=1}^{p-(j-u)+1} \overline{\mathcal{A}_{(j-u)+l}}$ , each of which is labeled with  $T_u$ . If no such  $u$  exists then we lead edges to each elements of  $\mathcal{A}$ , and label them with  $\alpha$ .

c) Let  $\alpha = T_p$ . Let  $u$  be the smallest integer for which  $\alpha = T_u \overline{T_{p-u}}$ , if any. Let us continue, as in part b).

d) If there is an edge leading from  $\beta$  to  $\gamma$  labeled with  $\delta$ , then lead an edge from  $\overline{\beta}$  to  $\overline{\gamma}$  and label this with  $\overline{\delta}$ .

We constructed this graph so, that the following assertion holds.

**Proposition 5.** *The elements of  $K_\alpha$  can be given by walking on the graph starting from  $\alpha$  and concatenated the sequence of the labels of the edges.*

PROOF. a) For  $i = 1, \dots, k-1$  we have  $K_i = iK$ , where  $iK = i \sum_{\alpha \in \mathcal{A}} K_\alpha$ .

b) See the proof of Proposition 6, and the investigation after this proof.

c) As part b). We used, that  $T_p$  cannot be factorized as  $T_p = T_u T_v$ .

d) This part is obvious.

*Remark.* We can rewrite the construction rule of the graph as follows. From node  $u = u_1 \dots u_m$  there is an edge  $e$  to node  $v = v_1 \dots v_n$  if  $u_1 \dots u_m = u_1 \dots u_k v_1 \dots v_l$ , with some  $1 \leq k \leq m$  and  $0 \leq l \leq n$ , and  $u_1 \dots u_m v_{l+1} \dots v_n$  is an allowed sequence part. We label  $e$  with  $u_1 \dots u_k$ . In the case  $k = m, l = 0$  the part  $u_1 \dots u_m v_{l+1} \dots v_n$  can occur in sequences (see examples below).

*Investigation of the graph.* In the sequel we examine, how we can specify the dimension of the graph. Let  $\alpha$  be an arbitrary word from  $\mathcal{A}$ . Let  $H_\alpha = \langle K_\alpha, \Theta \rangle = \{ \sum \varepsilon_n \Theta^n \mid \varepsilon \in K_\alpha \}$ . If  $\langle \varepsilon, \Theta \rangle = x$ , then  $\langle \bar{\varepsilon}, \Theta \rangle = L - x$ , so  $H_{\bar{\alpha}} = L - H_\alpha$ . Let  $\beta_1, \dots, \beta_n$  be the endpoints of the edges coming from  $\alpha$ , and  $\delta_1, \dots, \delta_n$  be the corresponding labels. Let  $\delta_s = u_1^{(s)} \dots u_{r_s}^{(s)}$ . Then  $K_\alpha = \sum u_1^{(s)} \dots u_{r_s}^{(s)} K_{\beta_s}$ , and

$$(11) \quad H_\alpha = \sum (u_1^{(s)} \Theta + \dots + u_{r_s}^{(s)} \Theta^{r_s} + \Theta^{r_s} H_{\beta_s}).$$

From the construction it follows that sets  $H_\alpha$  can be covered by open intervals  $I_\alpha$ , such that  $I_\alpha \cap I_\beta = \emptyset$  for each  $\alpha, \beta \in \mathcal{A}$ ,  $\alpha \neq \beta$ . Thus, the so called open set criterion fulfils (detailed proof in [5], Theorem 3.). From fractal geometry it is known, that the dimension of a strongly connected graph is the same, as the Hausdorff dimension of the set – now on the number line – which realizes this graph, if the open set criterion is satisfied (see e.g. [3], p. 170–173).

If from every  $\alpha \in \mathcal{A}$  leads a path to a word of length one, then  $\mathcal{G}(\mathcal{A})$  is strongly connected. It means, that the Hausdorff dimension of the components  $H_\alpha$  are equal (this is the Hausdorff dimension of the univoque set), and it is the same as the graph dimension of  $\mathcal{G}(\mathcal{A})$ .

#### 4. Not strongly connected graph in the finite case

However, sometimes  $\mathcal{G}(\mathcal{A})$  is strongly connected. Then there is some  $\alpha \in \mathcal{A}$ , from which no path leads to  $\mathcal{A}_1$ . Then there is a shortest  $\alpha \in \mathcal{A}$ , from which no path leads to shorter elements of  $\mathcal{A}$ . Let the length of  $\alpha$  be  $s \leq p$ , assuming  $\alpha \neq T_p$  (the case  $\alpha = T_p$  is not hard to investigate, it will be considered later). This  $\alpha$  can be written in the form  $\alpha = t_1 t_2 \dots t_{s-1} j$  with  $j < t_s$ , since  $\alpha \in \mathcal{A}_s$ .

**Proposition 6.** *For this  $\alpha$  there exists an index  $u$ , for which  $t_{u+1} \dots t_{s-1}j = \overline{T}_{s-u}$ .*

PROOF. First we investigate, for which sequences  $\xi \in K$  holds  $\alpha\xi \in K$ , too. Let  $\xi = k_1 \dots k_r$ . According to conditions in (8),

$$(12) \quad t_{u+1} \dots t_{s-1}jk_1 \dots k_v \leq T_p \quad \text{and} \quad \overline{t_{u+1} \dots t_{s-1}j} \overline{k_1 \dots k_v} \leq T_p$$

must be satisfied for every  $0 \leq u < s$  and  $v = p + u - s$ . If for all index  $u$

$$t_{u+1} \dots t_{s-1}j < T_{s-u} \quad \text{and} \quad \overline{t_{u+1} \dots t_{s-1}j} < T_{s-u},$$

then conditions in (12) clearly hold. This means, that  $\alpha\xi$  is an allowed sequence part,  $K_\alpha = \alpha K$ , i.e. we would be able to go from  $\alpha$  to an arbitrary  $\xi \in K$ . However, from  $\xi \in \mathcal{A}_1$  edges lead to every word, so this is a contradiction, since the graph would be strongly connected.

So now we must have an index  $u$ ,  $0 \leq u < s$ , for which either

$$t_{u+1} \dots t_{s-1}j = T_{s-u} \quad \text{or} \quad t_{u+1} \dots t_{s-1}j = \overline{T}_{s-u}.$$

In the first case

$$T_{s-u} = t_{u+1} \dots t_{s-1}j < t_{u+1} \dots t_{s-1}t_s \leq T_{s-u},$$

since  $j < t_s$ , and  $\sigma^j(T) \leq T$ . Thus, this case is not possible.  $\square$

*Investigation in case  $t_{u+1} \dots t_{s-1}j = \overline{T}_{s-u}$ .* Now

$$\alpha = t_1 t_2 \dots t_u \underbrace{t_{u+1} \dots t_{s-1}j}_{\overline{T}_{s-u}}, \quad \text{hence with } v = s - u,$$

$$K_\alpha = T_u K_{\overline{T}_v} = T_u \overline{K_{T_v}}.$$

Since  $K = \bigcup_{\alpha \in \mathcal{A}} K_\alpha$ , set  $K_{T_v}$  can be written in the form  $\sum K_\beta$  where the words  $\beta$  are in  $\mathcal{A}$ :  $T_v i \in \mathcal{A}_{v+1}$ ,  $T_v t_{v+1} i \in \mathcal{A}_{v+2}, \dots$ , and so eventually

$$K_{T_v} = \sum_{\beta \in \bigcup_{\gamma=v+1}^{p+1} \mathcal{A}_\gamma} K_\beta.$$

At this time edges lead from  $\alpha$  with label  $t_1 t_2 \dots t_u$  to all of the elements in  $\bigcup_{\gamma=v+1}^{p+1} \overline{\mathcal{A}}_\gamma$ . However, by reason of our assumption from  $\alpha$  we can

not reach a word shorter than  $\alpha$ , thus  $\overline{\mathcal{A}}_{v+1} = \overline{\mathcal{A}}_{v+2} = \dots = \overline{\mathcal{A}}_{s-1} = \emptyset$ , and edges lead from  $\alpha$  to the elements of the sets  $\overline{\mathcal{A}}_s, \overline{\mathcal{A}}_{s+1}, \dots, \overline{\mathcal{A}}_{p+1}$ . From this follows  $\mathcal{A}_{v+1} = \mathcal{A}_{v+2} = \dots = \mathcal{A}_{s-1} = \emptyset$ , too. We can observe that if  $\beta$  is another element of  $\mathcal{A}_s$ ,  $\beta \neq \alpha$ , then there is an edge from  $\alpha$  to  $\overline{\beta}$ , so  $\beta$  has the same property as  $\alpha$ , no path leads from  $\beta$  to shorter elements.

Let  $h^* < s$  be the greatest index, for which  $\mathcal{A}_{h^*} \neq \emptyset$ . Obviously  $h^* < v + 1$ . Using, that the sets  $\mathcal{A}_i$  are empty for  $i = v + 1, \dots, s - 1$  we have  $K_{T_v} = K_{T_{s-1}}$ , i.e.  $K_\alpha = T_u \overline{K_{T_v}} = T_u \overline{K_{T_{s-1}}}$ , and  $K_{T_{h^*}} = K_{T_{s-1}}$ . Here  $\alpha \in K_{T_{s-1}}$ , so this set is not empty.

We can now divide  $\mathcal{G}(\mathcal{A})$  into two parts. Let  $\mathcal{G}_1$  be the strongly connected graph part. All of the words with length  $\leq s - 1$  belong here. In the second part  $\mathcal{G}_2$  are all of the words, from which it is not possible to go to words in  $\mathcal{G}_1$ .  $\mathcal{G}_2$  contains only words longer than  $s - 1$ .

Our goal in the following is to specify and count the words in  $\mathcal{G}_2$ . To achieve this, let now  $\gamma$  be an arbitrary  $m$ -length word with  $m \geq s$ , from which we are not able to go to any word shorter than  $s$ .

**Proposition 7.** *In case  $m = p$ ,  $\gamma = T_p$  we have  $t_{u+1} \dots t_p = \overline{T_{p-u}}$  with  $p - u \geq h^*$ .*

PROOF. Let us make  $p$ -long examinations for  $\gamma\xi$ ,  $\xi \in K$ , as in Proposition 6:

$$t_1 t_2 \dots \underbrace{t_{u+1} \dots t_p k_1 \dots k_u}_{\text{...}} k_{u+1} \dots$$

Similarly as in Proposition 6 we find an index  $u$ , for which

- a)  $t_{u+1} \dots t_p = T_{p-u}$  or
- b)  $t_{u+1} \dots t_p = \overline{T_{p-u}}$ .

In case a) we would have  $T_p = T_u T_{p-u}$ , however this was excluded by Proposition 2. Thus, only case b) is possible.

Let us assume, that  $p - u < h^*$ . Let  $\tilde{\gamma}$  be an arbitrary  $h^*$ -long word from  $\mathcal{A}$ . Then  $K_{\tilde{\gamma}} \subseteq K_{T_{p-u}}$ . However, from  $T_p$  edges lead to the words of  $\overline{K_{T_{p-u}}}$ , since  $T_p = T_u \overline{T_{p-u}}$ , so eventually we would be able to go from  $T_p$  to the word  $\tilde{\gamma}$ . Its length  $h^*$  is smaller than  $s$ , but this is not possible, since  $m = p \geq s$ .

### 5. The dimension of the graph

In case  $p \geq m \geq s$ ,  $\gamma = T_{m-1}l$  with  $\gamma \in \mathcal{A}_m$  we clearly have

$$(13) \quad \gamma = T_u \overline{T_v} \text{ with } v \geq h^*,$$

according to Proposition 6. Thus, together with the result of Proposition 7, (13) holds for every word  $\gamma$  in  $\mathcal{G}_2$ .

So  $K_\gamma = T_u \overline{K_{T_v}} = T_u \overline{K_{T_{s-1}}}$ , and using the notation  $\gamma = T_{u_\gamma} \overline{T_{v_\gamma}}$  we get

$$K_{T_{s-1}} = \sum_{\lambda(\gamma) \geq s} K_\gamma = \sum_{\lambda(\gamma) \geq s} T_{u_\gamma} \overline{K_{T_{v_\gamma}}} = \sum_{\lambda(\gamma) \geq s} T_{u_\gamma} \overline{K_{T_{s-1}}},$$

where  $\lambda(\gamma)$  denotes the length of the word  $\gamma$ .

Let  $\mathcal{M}_{s-1} = \langle K_{T_{s-1}}, \Theta \rangle$ . Using our former results we get

$$\begin{aligned} \langle K_{T_{s-1}}, \Theta \rangle &= \sum_{\lambda(\gamma) \geq s} \langle T_{u_\gamma} \overline{K_{T_{v_\gamma}}}, \Theta \rangle \\ &= \sum_{\lambda(\gamma) \geq s} \{t_1 \Theta + \cdots + t_{u_\gamma} \Theta^{u_\gamma} + \Theta^{u_\gamma} \langle \overline{K_{T_{s-1}}}, \Theta \rangle\}. \end{aligned}$$

We know, that  $\langle \overline{C}, \Theta \rangle = L - \langle C, \Theta \rangle$ , since  $\sum \overline{c_i} \Theta^i = \sum (k - c_i) \Theta^i = L - \sum c_i \Theta^i$ . Applying this  $\Theta^{u_\gamma} \langle \overline{K_{T_{s-1}}}, \Theta \rangle = \Theta^{u_\gamma} L - \Theta^{u_\gamma} \langle K_{T_{s-1}}, \Theta \rangle$ , and eventually

$$\mathcal{M}_{s-1} = \sum_{\lambda(\gamma) \geq s} \{t_1 \Theta + \cdots + t_{u_\gamma} \Theta^{u_\gamma} + \Theta^{u_\gamma} L - \Theta^{u_\gamma} \mathcal{M}_{s-1}\}.$$

Thus  $\mathcal{M}_{s-1}$  is a self-similar set, we produce it with the following mapping:  $\varphi_\gamma(x) = t_1 \Theta + \cdots + t_{u_\gamma} \Theta^{u_\gamma} + \Theta^{u_\gamma} L - \Theta^{u_\gamma} x$ , thus  $\mathcal{M}_{s-1} = \sum_{\lambda(\gamma) \geq s} \varphi_\gamma(\mathcal{M}_{s-1})$ . Let us denote the self-similarity dimension of  $\mathcal{M}_{s-1}$  by  $\eta$ . Then

$$1 = \sum_{\lambda(\gamma) \geq s} \Theta^{u_\gamma \eta}.$$

To compute this dimension we have to specify the number of the appropriate  $\gamma$ -s. Let  $C(n) = \#\{\gamma \mid \lambda(\gamma) \geq s, u_\gamma = n\}$ , and so  $1 = \sum C(n) \Theta^{n\eta}$ . Let us introduce a function  $g_1$  in the following manner:

$$g_1(y) = \sum C(n) y^n - 1.$$

Thus,  $g_1(0) = -1$ . Using the fact, that the linear combination of coefficients  $C(i)$  with numbers  $\Theta^{n\eta}$  (they are all less than 1) produces 1,  $g_1(1) > 0$  holds. So between 0 and 1 there exists a root  $y_1$ , for which  $g_1(y_1) = 0$ . The self-similarity dimension searched for is the number  $\eta_1$  for which  $\Theta^{\eta_1} = y_1$ .

*Upper estimation of  $C(n)$ .* Since  $u_\gamma = n$ , thus

$$(14) \quad \gamma = T_n \overline{T_{v_\gamma}} = T_{n+v_\gamma-1} \overline{t_{v_\gamma}},$$

where in the last equality we used, that  $\gamma$  is a word in  $\mathcal{A}$ , and so it can be written in the form  $T_{\lambda(\gamma)-1}l$ .

**Proposition 8.** *For  $v_\gamma$  we have  $h^* \leq v_\gamma \leq n$ .*

PROOF. a) Analysing (14) we conclude, that from  $\gamma$  edges lead to the elements of  $\overline{K_{T_{v_\gamma}}}$ , but to words shorter than  $h^*$  surely do not. Thus  $h^* \leq v_\gamma$ .

b) Let us assume, that  $v_\gamma > n$ . From (14) we get  $T_{n+v_\gamma-1} = T_n \overline{T_{v_\gamma-1}}$ , so in the case  $v_\gamma - 1 \geq n$  it would derive  $T_p = T_n \overline{T_n} V$ , where  $V$  is an arbitrary sequence part. However, this case was excluded by Proposition 3.

Since  $v_\gamma$  is a word length, it can have only  $n - h^* + 1$  different values. Thus  $C(n) \leq n - h^* + 1$ , since the prefixes of the words  $\gamma$  were fixed (with  $T_{u_\gamma} = T_n$ ). So the number of appropriate  $\gamma$ -s is specified with the number of possible  $v_\gamma$ -s. Let us introduce

$$g_2(y) = \sum_{n \geq h^*} (n - h^* + 1) y^n - 1.$$

For this function  $g_2(0) = -1$  and  $g_2(1) > 0$ , since its coefficients are not smaller, than those of  $g_1$ . By the same reason for the root of this function  $y_2 \leq y_1$  holds. Let now  $\Theta^{\eta_2} = y_2$ . Taking into account, that the numbers  $\Theta$ ,  $y_1$  and  $y_2$  are less than 1 and  $\log y_2 \leq \log y_1$ , we get  $\eta_2 \geq \eta_1$ . We can

write

$$\begin{aligned}
 g_2(y) &= y^{h^*}(1 + y + y^2 + \dots) + \sum_{n \geq h^*} (n - h^*)y^n - 1 \\
 &= \frac{y^{h^*}}{1 - y} + y^{h^*+1}(1 + 2y + 3y^2 + \dots) - 1 \\
 &= \frac{y^{h^*}}{1 - y} + y^{h^*+1}(1 + y + y^2 + \dots)' - 1 \\
 &= \frac{y^{h^*}}{1 - y} + \frac{y^{h^*+1}}{(1 - y)^2} - 1 = \frac{y^{h^*} - y^2 + 2y - 1}{(1 - y)^2}.
 \end{aligned}$$

Let  $\kappa$  be the numerator. For  $\kappa(y)$  we have  $\kappa(0) = -1$  and  $\kappa(1) = 1$ , so this function has a root between 0 and 1.

Let us recall the definitions of  $\mathcal{G}_1$  and  $\mathcal{G}_2$  in Section 4, where part  $\mathcal{G}_2$  contains the  $\gamma$  words. Using, that all of the elements of  $\mathcal{A}_1$  are in  $\mathcal{G}_1$ , let  $P$  be the set of those sequences, which contain only digits  $1, 2, \dots, k-1$ . Then

$$P = \sum_{i=1}^{k-1} iP, \quad \text{i.e.} \quad \langle P, \Theta \rangle = \sum_{i=1}^{k-1} i\Theta + \Theta \langle P, \Theta \rangle,$$

so set  $\langle P, \Theta \rangle$  can be constructed from itself with similarities  $h_i(x) = i\Theta + \Theta x$ . Thus, its self similarity dimension  $\xi$  can be computed with  $1 = (k-1)\Theta^\xi$ , from which  $\Theta^\xi = \frac{1}{k-1}$ , i.e.

$$(15) \quad \xi = \frac{\log \frac{1}{k-1}}{\log \Theta} = \frac{\log(k-1)}{\log \beta}.$$

Now,  $\dim \mathcal{G}_1 \geq \xi$  and  $\dim \mathcal{G}_2 \leq \eta_1$ . Thus, if  $\xi > \eta_2 \geq \eta_1$  i.e. for the roots  $\frac{1}{k-1} < y_2 \leq y_1$  holds, then  $\dim \mathcal{G}_1 = \dim \mathcal{G} (= \dim_H \langle K, \Theta \rangle)$ , where  $\dim_H$  denotes the Hausdorff dimension). This is guaranteed now by condition  $\kappa(\frac{1}{k-1}) < 0$ . Here we used a theorem of R. D. MAULDIN and S. C. WILLIAMS, which in our case guarantees that the dimension of a graph containing strongly connected components is the maximum of the dimensions of the components ([7]).

We have the following cases:

1.  $k \geq 4$ ,  $h^* \geq 1$ . Then  $\kappa(\frac{1}{k-1}) = (\frac{1}{k-1})^{h^*} - (\frac{1}{k-1})^2 + \frac{2}{k-1} - 1 < 0$ .

2.  $k = 3$ ,  $h^* \geq 3$ . Still true, that  $\kappa(\frac{1}{2}) = (\frac{1}{2})^{h^*} - \frac{1}{4} + 1 - 1 < 0$ .
3.  $k = 3$ ,  $h^* = 2$ . Then  $\kappa(\frac{1}{2}) = (\frac{1}{2})^2 - \frac{1}{4} + 1 - 1 = 0$ , so  $\xi = \eta_2 \geq \eta_1$ , i.e. for the roots  $\frac{1}{k-1} = y_2 \leq y_1$ . This means, that  $\dim \mathcal{G}_1 \geq \dim \mathcal{G}_2$ .
4.  $k = 2$ . Unfortunately, with this method we can not reach any result.

Considering now case  $\alpha = T_p$  we can argue as follows. For  $k \geq 2$  the structure of  $\mathcal{G}_1$  is more complicated, than that of  $\mathcal{G}_2$ , which contain only words  $T_p$  and  $\overline{T}_p$ . Namely,  $\dim \mathcal{G}_2 = 0$ , since  $K_{T_p}$  is a countable set. So  $\dim \mathcal{G}_1 = \dim \mathcal{G}$ .  $\square$

Thus, in Sections 4 and 5 we have proved eventually

**Theorem 4** (Main result for the dimension of the univoque set). *Let us consider all of the (number) systems in finite subcase of the big case. Let us construct a graph  $\mathcal{G}$ , which describe the structure of the univoque set, according to results in Section 3. Let  $\xi$  be specified by equation (15).*

1. *If the graph is strongly connected, then  $\dim_H \langle K, \Theta \rangle = \dim \mathcal{G} \geq \xi$ .*
2. *If the graph is not strongly connected, let  $h^*$  be the greatest index, for which  $\mathcal{A}_{h^*} \neq \emptyset$ . If  $k \geq 4$  or  $k = 3$  and  $h^* \geq 2$ , then we have  $\dim_H \langle K, \Theta \rangle = \dim \mathcal{G} = \dim \mathcal{G}_1$  with  $\dim \mathcal{G}_1 \geq \xi$ , where  $\mathcal{G}_1$  is the strongly connected part containing the elements of  $\mathcal{A}_1$ .*

## 6. Reduced representation

**Proposition 9.** *Let  $\alpha$  be an arbitrary word from  $\mathcal{A}$ . To specify the dimension of  $\mathcal{G}(\mathcal{A})$ , it is enough to investigate a simplified graph, which contains only either of  $\alpha$  or  $\overline{\alpha}$ , with keeping up the relations between the nodes.*

PROOF. As in (11), let  $\beta_1, \dots, \beta_n$  be the endpoints of the edges coming from  $\alpha$ , and  $\delta_1, \dots, \delta_n$  be the corresponding labels. Let  $\delta_s = u_1^{(s)} \dots u_{r_s}^{(s)}$ . Then  $K_\alpha = \sum u_1^{(s)} \dots u_{r_s}^{(s)} K_\beta$ . If  $\beta_s \in \mathcal{A}_j$  then for  $H_\alpha = \langle K_\alpha, \Theta \rangle$  formula (11) holds. If  $\beta_s \in \overline{\mathcal{A}}_j$  then

$$(16) \quad H_\alpha = \sum \left( u_1^{(s)} \Theta + \dots + u_{r_s}^{(s)} \Theta^{r_s} + \Theta^{r_s} (L - H_{\overline{\beta_s}}) \right)$$

holds, using that  $H_{\overline{\alpha}} = L - H_\alpha$ . Since the dimension of the sets  $\mathcal{B}$  and  $L - \mathcal{B}$  are the same, it is enough to consider from words  $\beta_s$  only those, which are either in  $\mathcal{A}_j$ , or in  $\overline{\mathcal{A}}_j$ .

Thus, because of the symmetry of the univoque set, we are able to simplify  $\mathcal{G}(\mathcal{A})$  as follows.

a) Let  $\mathcal{B}$  be the set  $\{1, \dots, [\frac{k}{2}]\} \cup_{i=2}^{p+1} \mathcal{A}_i$ . It is clear that  $\mathcal{B} \cup \overline{\mathcal{B}} = \mathcal{A}$ , and that  $\mathcal{B} \cap \overline{\mathcal{B}}$  is empty if  $k$  is odd, and contains only the element  $[\frac{k}{2}]$  if  $k$  is even.

b) We construct the directed multigraph  $\mathcal{G}(\mathcal{B})$  from  $\mathcal{G}(\mathcal{A})$ . Assume that  $\alpha, \beta \in \mathcal{B}$  and in  $\mathcal{G}(\mathcal{A})$  there is an edge from  $\alpha$  to  $\beta$  labelled with  $\delta$ . Then we keep this edge and label in  $\mathcal{G}(\mathcal{B})$ . Assume that  $\alpha \in \mathcal{B}, \beta \notin \mathcal{B}$  and in  $\mathcal{G}(\mathcal{A})$  there is an edge from  $\alpha$  to  $\beta$  labelled with  $\delta$ . Then  $\overline{\beta} \in \mathcal{G}(\mathcal{B})$ . In  $\mathcal{G}(\mathcal{B})$  we lead an edge from  $\alpha$  to  $\overline{\beta}$  and label it with  $(\delta, *)$ .

We apply this rule for all  $\alpha \in \mathcal{B}$ . □

In the following we call the original graph as *normal representation*, and the simplified one as *reduced representation*. Our former results clearly hold in the reduced representation, too.

## 7. Examples

*Case*  $1 = 3\Theta + \Theta^2 + \Theta^3$ . In this system  $\beta \approx 3.3830$ ,  $T = 310310\dots$ , i.e.  $T$  is periodic with  $p = 3$ , and  $K = V_3$ . Thus, sets  $\mathcal{A}_i$  are the following:  $\mathcal{A}_1 = \{1, 2\}$ ,  $\mathcal{A}_2 = \{30\}$ ,  $\mathcal{A}_3 = \{\emptyset\}$ ,  $\mathcal{A}_4 = \{310\}$ , and  $\overline{\mathcal{A}}_2 = \{03\}$ ,  $\overline{\mathcal{A}}_4 = \{023\}$ .

We can symbolize the structure of the univoque set in normal and reduced representation with graphs shown in Figure 1 respectively.

*Figure 1: The graphs representing the structure of set  $K$*

In both of the representations these graphs contain two strongly connected parts. In reduced representation we leave the nodes beginning with 2 and 0. The remaining ones are 1, 30 and 310.

We specify the Hausdorff dimension using reduced representation (in normal representation we can work in the same manner). For the computation of the Hausdorff dimension the method presented in [3] will be applied.

For the first graph part

$$q_1^s = 2\lambda \cdot q_1^s, \text{ from which immediately } \lambda = \frac{1}{2}.$$

For the second graph part

$$q_{30}^s = \lambda \cdot q_{310}^s + \lambda \cdot q_{30}^s, \quad q_{310}^s = \lambda^2 \cdot q_{30}^s + \lambda^2 \cdot q_{310}^s$$

after substitution

$$q_{310}^s = \lambda^2 \frac{\lambda \cdot q_{310}^s}{1 - \lambda} + \lambda^2 \cdot q_{310}^s, \quad 0 = \lambda^2 + \lambda - 1,$$

the solution of which is

$$\lambda = \frac{\sqrt{5} - 1}{2}.$$

Thus, for the dimensions of the graph parts we have

$$0.3948 \approx s_1 = \frac{\log(\frac{\sqrt{5}+1}{2})}{\log \beta} < \frac{\log 2}{\log \beta} = s_2 \approx 0.5687,$$

i.e. the dimension of the whole graph is  $s_2$ .

*Remark.* In [5] we have investigated this system using a full representation, with all possible condition 3-s. With this method the graph was very complicated, and the computation of the Hausdorff dimension was lengthy.

*Cases with  $T = 42110001\eta$ , where  $\eta$  is an arbitrary allowable sequence part.* Choosing  $p = 8$  and  $u = 4$  we have  $t_{u+1} \dots t_p = 0001 < 0233 = \bar{T}_{p-u}$ . Thus,  $K = U_8$ , according to Theorem 1. For  $T_8 = 42110000$  relation  $K = U_8$  still holds, so with reduction we get  $T_4 = 4211$  with  $K = U_4$ .

With  $T_4 = 4210$  we have finally  $K = V_4$ . Thus, sets  $\mathcal{A}_i$  are the following:

$\mathcal{A}_1 = \{1, 2, 3\}$ ,  $\mathcal{A}_2 = \{40, 41\}$ ,  $\mathcal{A}_3 = \{420\}$ ,  $\mathcal{A}_4 = \{\emptyset\}$ ,  $\mathcal{A}_5 = \{4210\}$ ,  
and  $\overline{\mathcal{A}}_2 = \{04, 03\}$ ,  $\overline{\mathcal{A}}_3 = \{024\}$ ,  $\overline{\mathcal{A}}_5 = \{0234\}$ .

We can symbolize the structure of the univoque set in normal and reduced representation with graphs shown in Figure 2 respectively.

*Figure 2: The graph representing the structure of set  $K$*

In both of the representations the graphs are strongly connected. In normal representation we marked some nodes with  $*$ , from these nodes edges lead to all of the others. In reduced representation we leave the nodes beginning with 3 and 0. The remaining ones are 1, 2, 41, 40, 420 and 4210. From nodes denoted with  $**$ , double edges lead to all of the other nodes. The labels of these edges are  $i$  and  $(i, *)$ , respectively. However, since  $2 = \overline{2}$ , there is only one edge from 1 to 2 and from 2 to 2. As above, we specify the Hausdorff dimension using reduced representation

$$q_1^s = 2\lambda \cdot q_1^s + \lambda \cdot q_2^s + 2\lambda \cdot q_{41}^s + 2\lambda \cdot q_{40}^s + 2\lambda \cdot q_{420}^s + 2\lambda \cdot q_{4210}^s = q_2^s$$

$$q_{41}^s = \lambda \cdot q_1^s$$

$$q_{40}^s = \lambda \cdot q_{40}^s + \lambda \cdot q_{420}^s + \lambda \cdot q_{4210}^s + \lambda \cdot q_{41}^s$$

$$q_{420}^s = \lambda^2 \cdot q_{40}^s + \lambda^2 \cdot q_{420}^s + \lambda^2 \cdot q_{4210}^s + \lambda^2 \cdot q_{41}^s = \lambda \cdot q_{40}^s$$

$$q_{4210}^s = \lambda^3 \cdot q_{40}^s + \lambda^3 \cdot q_{420}^s + \lambda^3 \cdot q_{4210}^s + \lambda^3 \cdot q_{41}^s = \lambda^2 \cdot q_{40}^s.$$

Substituting  $q_{420}^s$  and  $q_{4210}^s$ , we get an equation system for  $\lambda$ ,  $q_1^s$ ,  $q_{41}^s$  and  $q_{40}^s$ . This system yields  $0 = 3\lambda^4 + 2\lambda^3 - 4\lambda + 1$ , which real solution in the interval  $[0, 1/2]$  is  $\lambda \approx 0,2626$ .

$$\text{Thus, the dimension of the univoque set is } s = \frac{\log \frac{1}{\lambda}}{\log \beta} \approx \frac{\log 3,8077}{\log \beta}.$$

## References

- [1] Z. DARÓCZY and I. KÁTAI, Univoque sequences, *Publ. Math. Debrecen* **42** (1993), 397–407.
- [2] Z. DARÓCZY and I. KÁTAI, On the structure of univoque numbers, *Publ. Math. Debrecen* **46** (1995), 385–408.
- [4] G. A. EDGAR, Measure, Topology and Fractal Geometry, *Springer Verlag, New York*, 1994.
- [4] GÁBOR KALLÓS, The Structure of the Univoque Set in the Small Case, *Publ. Math. Debrecen* **54** (1999), 153–164.
- [5] GÁBOR KALLÓS, Hausdorff Dimension of Univoque Sets, *Acta Cybernetica* **14** (1999), 303–314.
- [6] I. KÁTAI and G. KALLÓS, On the Set for which 1 is Univoque, *Publ. Math. Debrecen* (to appear).
- [7] R. D. MAULDIN and S. C. WILLIAMS, Hausdorff Dimension in Graph Directed Constructions, *Trans. Amer. Math. Soc.* **309** (1988), 811–819.
- [8] W. PARRY, On the  $\beta$  Expansions of Real Numbers, *Acta Math. Hung.* **11** (1960), 401–406.

GÁBOR KALLÓS  
DEPARTMENT OF COMPUTER SCIENCE  
SZÉCHENYI ISTVÁN COLLEGE  
H-9026 GYŐR, HÉDERVÁRI ÚT 3.  
HUNGARY

E-mail: kallos@rs1.szif.hu

(Received April 4, 2001; revised April 4, 2001)