5. Substitutions and partitions of the set of positive integers

There is a natural duality between the symbolic sequences, as they are studied throughout this book, and the *partitions of the set of positive integers* $\mathbb{N}^+$: given an infinite word (to the right) $\omega = \omega_1\omega_2\omega_3\ldots$ over the alphabet

$$\mathcal{A}_s := \{a_0, a_1, \ldots, a_s\},$$

where every symbol a_i eventually occurs in ω, it gives rise to a partition of $\mathbb{N}^+$ into $s+1$ parts:

$$\overset{\circ}{\bigcup_{0\le i\le s}} \chi(\omega; a_i) = \mathbb{N}^+, \tag{5.1}$$

which will be referred to as the *partition corresponding* to ω (throughout this chapter $\overset{\circ}{\bigcup}$ indicates a disjoint union); $\chi(\omega; a)$ is the characteristic set of ω with respect to a:

$$\chi(\omega; a) := \{n \in \mathbb{N}^+; \omega_n = a\}, a \in \mathcal{A}_s.$$

Conversely, any partition of $\mathbb{N}^+$ into sets $A_1, \ldots, A_s$ gives rise to a word defined by $\omega_n = i$ whenever $n \in A_i$.

We say that the partition (5.1) is *nonperiodic* (respectively *totally nonperiodic* if ω (respectively $\partial\chi(\omega; a_i)$ for all i) is not an ultimately periodic word (or sequence), and vice versa, where we mean by ∂C the sequence $(c_{n+1} - c_n)_{n\ge1}$ for a given sequence $C := (c_n)_{n\ge1}$.

Throughout this chapter, we also identify a set $\{s_n; n \in \mathbb{N}\} \subset \mathbb{N}^+$ such that $s_1 < s_2 < s_3 < \ldots$ with the sequence $(s_n)_{n\ge1}$.

In what follows, we shall give some classes of nonperiodic partitions of $\mathbb{N}^+$ (or some classes of nonperiodic infinite words), and some results and problems related to transcendence and complexity.

5.1 Beatty and Sturmian sequences, and associated partitions

5.1.1 Partitions associated with Beatty sequences

Let $[x](x \in \mathbb{R})$ be the largest integer not exceeding x, $\{x\}$ be the fractional part of $x \in \mathbb{R}$, i.e., $\{x\} := x - [x]$, and let $\mathbb{N}^+$ (respectively $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{R}^+$) denote the set of the positive integers (respectively the integers, the rational numbers, the real numbers, the positive numbers).

Let us recall that a *Beatty sequence* is a sequence of the form $([\alpha n+\beta])_{n\in\mathbb{N}}$. It is well-known that for positive numbers α and β, two Beatty sequences (or sets) $([\alpha n])_{n\geq 1}$ and $([\beta n])_{n\geq 1}$ make a partition of the set $\mathbb{N}^+$ into two parts if and only if α, β are irrationals numbers satisfying $1/\alpha + 1/\beta = 1$. This fact can be written in an equivalent form as follows. We have the following partition

$$\mathop{\overset{\circ}{\bigcup}}_{(\gamma_0,\gamma_1)\in A} \{[\gamma_0 n] + [\gamma_1 n]; n \in \mathbb{N}^+\} = \mathbb{N}^+, \tag{5.2}$$

$$\text{where } A := \{(1,\alpha), \alpha^{-1}(1,\alpha)\}, \text{with } \alpha > 0,$$

if and only if α is an irrational number.

Proposition 5.1.1 below is a generalization of (5.2), which gives a partition of $\mathbb{N}$ into $s+1$ parts by specific sums of Beatty sequences (see [342], Theorem 1).

We denote by $\alpha S, S+\alpha, S+T$, and ST the set $\{\alpha s; s \in S\}, \{s+\alpha; s \in S\}, \{s+t; s \in S, t \in T\}$, and $\{st; s \in S, t \in T\}$, respectively, for given sets $S, T \subset \mathbb{R}$, and for a given number $\alpha \in \mathbb{R}$.

Proposition 5.1.1. *Let s be a positive integer, and $\alpha_i > 0, \beta_i$ $(0 \leq i \leq s)$ be real numbers. Then the condition*

$$(\alpha_i^{-1}\mathbb{Z} - \alpha_i^{-1}\beta_i) \cap (\alpha_j^{-1}\mathbb{Z} - \alpha_j^{-1}\beta_j) \cap \mathbb{R}_+ = \emptyset \quad \textit{for all } i \neq j \tag{5.3}$$

is necessary and sufficient to have a partition

$$\mathop{\overset{\circ}{\bigcup}}_{(\underline{\gamma},\underline{\delta})\in L} \{ \sum_{(0\leq i\leq s)} [\gamma_j n + \delta_j]; n \in \mathbb{N}^+\} = \mathbb{N}^+ \tag{5.4}$$

where $(\underline{\gamma}, \underline{\delta}) = (\gamma_0, \gamma_1, \ldots, \gamma_s, \delta_0, \delta_1, \ldots, \delta_s)$, *and*

$$\begin{aligned} L := \{ (\underline{\alpha}, \underline{\beta});\ & \underline{\alpha} = \alpha_i^{-1}(\alpha_0, \alpha_1, \ldots, \alpha_s), \\ & \underline{\beta} = -\alpha_i^{-1}\{\beta_i\}(\alpha_0, \alpha_1, \ldots, \alpha_s) \\ & \qquad +(\{\beta_0\}, \{\beta_1\}, \ldots, \{\beta_s\}), 0 \leq i \leq s \}. \end{aligned}$$

We then define $\omega(L)$ as the word corresponding to this partition.

Setting $\alpha_0 = 1, \beta_i = 0$ for all i, we have

Corollary 5.1.1. *Let $s \in \mathbb{N}^+$, $\alpha_0 = 1$, $\alpha_i \in \mathbb{R}^+$ $(1 \le i \le s)$. Then the condition $\alpha_i \notin \mathbb{Q}$, and $\alpha_i/\alpha_j \notin \mathbb{Q}$ for all $1 \le i < j \le s$ implies*

$$\mathop{\overset{\circ}{\bigcup}}_{(\gamma_0,\dots,\gamma_s)\in C} \{ \sum_{0\le j\le s} [\gamma_j n]; n \in \mathbb{N}^+ \} = \mathbb{N}^+,$$

and vice versa, where $C := \{\alpha_i^{-1}(\alpha_0, \dots, \alpha_s), 0 \le i \le s\}$.

Remarks. If we take $s = 1$ in Corollary 5.1.1, we obtain (5.2). We remark that we can choose $\alpha_0 = 1$, $\beta_0 = 0$ in Proposition 5.1.1 without changing the form of the components of the partition. If $\alpha_0 = 1$, the partition (5.4) is nonperiodic if one of the α_i $(1 \le i \le s)$ is irrational, since the irrationality of f_{a_i}/f_{a_0} implies the nonperiodicity of $\omega(L)$, where f_{a_i} is the frequency, in the asymptotic sense (see Chap. 1), of a symbol a_i appearing in the word $\omega(L)$ corresponding to the partition (5.4).

Related to the condition (5.3), we can show the implications (5.8) $\Rightarrow$ (5.7) $\Rightarrow$ (5.6) $\Rightarrow$ (5.3) $\Rightarrow$ (5.5) in the case of $\alpha_0 = 1, \beta_0 = 0$, where (5.5)$-$(5.8) are the following conditions:

- $-\beta_i \notin \alpha_i\mathbb{N}^+ + \mathbb{Z}$ for all $1 \le i \le s$; (5.5)
- $\alpha_i\beta_j - \alpha_j\beta_i \notin \alpha_i\mathbb{Z} + \alpha_j\mathbb{Z}$ for all $1 \le i < j \le s$; (5.6)
- $1, \alpha_i, \beta_i$ are linearly independent over $\mathbb{Q}$ for each $1 \le i \le s$, and $(\alpha_i(\mathbb{Z} + \beta_i)\mathbb{Q}) \cap (\alpha_j(\mathbb{Z} + \beta_j)\mathbb{Q}) = \{0\}$ for all $1 \le i < j \le s$; (5.7)
- the $2s + 1$ numbers 1, and $\alpha_i, \alpha_i\beta_i (1 \le i \le s)$ are linearly independent over $\mathbb{Q}$. (5.8)

We remark that a result obtained by J. V. Uspensky [359] says the impossibility of having a partition into t parts by Beatty sequences for $t \ge 3$. Proposition 5.1.2 is a generalization of Proposition 5.1.1 (see [342], Theorem 3).

Proposition 5.1.2. *Let $f_i : \mathbb{R}_+ \cup \{0\} \to \mathbb{R}$ $(0 \le i \le s, 1 \le s \in \mathbb{N}^+)$ be a continuous, strictly monotone increasing function with $\lim_{x\to\infty} f_i(x) = \infty$ for all i. Then the condition*

$$f_i^{-1}(\mathbb{Z}) \cap f_j^{-1}(\mathbb{Z}) \cap \mathbb{R}_+ = \emptyset \ \text{ for all } i \ne j \tag{5.9}$$

is necessary and sufficient to have a partition

$$\mathop{\overset{\circ}{\bigcup}}_{0\le i\le s} \{ \sum_{(0\le j\le s)} ([f_j(f_j^{-1}(n + [f_i(0)]))] - [f_j(0)]); n \in \mathbb{N}^+ \} = \mathbb{N}^+. \tag{5.10}$$

The property $\lim_{x\to\infty} f_i(x) = \infty$ can be omitted from Proposition 5.1.2, at most, for s indices i. In that case, some of the components of the partition (5.10) turn out to be a finite set. We remark that in this sense, any partition of $\mathbb{N}^+$ into $s+1$ parts can be given by (5.10) under a suitable choice of the functions f_i (without loss of generality, we may assume that all the f_i are of $\mathcal{C}^\infty$ class with $f_0(x) = x$), that will be clear from the following argument.

Proof. The idea of the proof of the proposition is very simple. We denote by Π_i, $\Pi_i \subset \mathbb{R}^{n+1}$ (respectively $K \subset \mathbb{R}^{n+1}$) the set of hyperplanes (respectively the curve) defined by

$$\Pi_i := \{(x_0, \ldots, x_i, \ldots, x_s); x_j \in \mathbb{R}(j \neq i), x_i \in \mathbb{Z}\}(0 \leq i \leq s),$$

$$\Pi := \bigcup_{0 \leq i \leq s} \Pi_i \,;\, K := \{\underline{f}(x) = (f_0(x), \ldots, f_s(x)); x \in \mathbb{R}_+\}.$$

Secondly, we consider an infinite word $\omega = \omega(K) = \omega_1\omega_2\omega_3\ldots$ given by

$$\omega_n = a_i \text{ if } \underline{f}(x_n) \in \Pi_i \,,$$

where the sequence $(x_n)_{n\geq 1}$ is defined by

$$\{\underline{f}(x_n)\,;\, 0 < x_1 < x_2 < \cdots < x_n < \ldots\} := K \cap \Pi\,.$$

Note that the sequence $(x_n)_{n\geq 1}$ is well-defined, since the set $K \cap \Pi$ is a discrete one in $\mathbb{R}^{s+1}$ if the functions f_i are continuous, and strictly monotone increasing; and the word ω is well-defined by (5.9).

Under the assumption that the functions f_i are continuous, strictly monotone increasing, we can calculate the n-th term of the sequence (or the set) $\chi(\omega; a_i)$ by using the intermediate value theorem, and we can obtain Proposition 5.1.2. Taking K to be a half-line L, we get Proposition 5.1.1. For further details of the proof, see [342]. ∎

5.1.2 Billiards and Sturmian sequences

Proposition 5.1.1 has some connection with higher dimensional *billiards* (see also Chap. 6). Let $I^{s+1}(I := [0,1])$ be the unit cube of dimension $s+1$ with the faces

$$\{(x_1, \ldots, x_i, \ldots, x_t); x_j \in I \ (\forall j \neq i),\ x_i = 0, \text{ or } 1, \text{ for } 0 \leq i \leq s\}$$

labelled by a_i. Let a particle start at a point $\underline{\beta} \in [0,1)^{s+1}$ along a vector $\underline{\alpha} \in \mathbb{R}^{s+1}$ with the condition for α_i, β_j stated in Proposition 5.1.1, and be reflected at each face of I^{s+1} specularly. Then the word $\omega(L)$ where $L := \{\underline{\alpha}t + \underline{\beta}; t \in \mathbb{R}_+\}$ defined above coincides with a word obtained by writing down the label a_i of the faces which the particle hits in order of collision.

Let us recall that the complexity p_w of an infinite word w is the function $p_w : \mathbb{N} \to \mathbb{N}$ defined as the number of factors of length n of w (see Chap. 1). An infinite word w (or a sequence) is called *Sturmian* if there exists a positive integer s such that $p_w(n) = n + s$ for every n. It is known that if w is not an ultimately periodic word defined strictly over the alphabet $\mathcal{A}_s$, then $p_w(n) \geq n + s$, see [256, 146] and Proposition 1.2.1. It is a classical result that for $s = 1$, $\omega(L)$ is Sturmian over 2 letters provided that ω is not periodic. G. Rauzy has conjectured that $p_{\omega(L)}(n) = n^2 + n + 1$ for $s = 2$, when $\alpha_1, \alpha_2, \alpha_3$ are linearly independent over $\mathbb{Q}$, which was proved affirmatively in [40], see also [293, 294, 41]. An exact formula for $p_{\omega(L)}(n) = p_{\omega(L)}(n, s)$ as a function of n and s in the case where $\alpha_0, \dots, \alpha_s$ are linearly independent over $\mathbb{Q}$ was conjectured in [40]:

$$p_{\omega(L)}(n, s) = \sum_{0 \leq i \leq min\{n,s\}} n!s!/((n-i)!i!(s-1)!)$$

and proved affirmatively by Y. Baryshnikov [46]. Consequently, we have for every n, s:

$$p_{\omega(L)}(n, s) = p_{\omega(L)}(s, n) \text{ and } p_{\omega(L)}(n; 3) = n^3 + 2n + 1.$$

This was one of the conjectures, from which P. Arnoux and C. Mauduit derived the exact formula by adding some minor hypotheses. It is an interesting question (posed by C. Mauduit), to ask for a direct (or combinatorial) proof of the symmetry; it still remains mysterious why $p(n, s)$ is a symmetric function.

Quite recently, S. Ferenczi and C. Mauduit [146] have obtained the following remarkable result: the numbers having a Sturmian sequence consisting of any number of letters in the set $\{n \in \mathbb{N}^+ ; 0 \leq n \leq h\}$ as base $g(\geq h+1)$ expansion are transcendental (this was conjectured by C. Mauduit in 1989). They gave further results on transcendence of numbers having a sequence (or an infinite word) with low complexity as their expansion in base g.

5.2 Partitions given by substitutions

5.2.1 Existence theorem

By $\mathcal{A} = \mathcal{A}_s$ we mean the alphabet $\{a_0, a_1, \dots, a_s,\}(s \geq 1)$ as in Section 5.1.

We denote by $|W|$ the length of a finite word w, and by $|W|_a$ the number of occurrences of a symbol $a \in \mathcal{A}$ appearing in a word $W \in \mathcal{A}^*$. For a given sequence $C = (c_n)_{n \geq 1}$, $\int_i C$ indicates the sequence

$$\int_i C := (i + \sum_{1 \leq m \leq n-1} c_m)_{n \geq 1}.$$

Then we can show the following

Proposition 5.2.1. *Let σ be a substitution over the alphabet $\mathcal{A}$ defined by*

$$\sigma(a_j) := {a_0}^{k_{s-j}} a_{j+1} \ (1 \le j \le s-1), \ \sigma(a_s) := a_0,$$

where k_i $(1 \le i \le s)$ are integers satisfying $k_s \ge k_{s-1} \ge \cdots \ge k_0 = 1$. Let L_j be the set $\{|\sigma^j(a_0)|, |\sigma^j(a_0)| + |\sigma^j(a_1)|, \ldots, |\sigma^j(a_0)| + |\sigma^j(a_s)|\}(1 \le j \le s)$, and let $\tau_j : \mathcal{A}^* \to {L_j}^*$ be a monoid morphism defined by

$$\tau_j(a_i) := (|\sigma^j(a_0)|)^{k_{s-i}-1}(|\sigma^j(a_0)| + |\sigma^j(a_{i+1})|) \ (0 \le i \le s-1),$$

$$\tau_j(a_s) := |\sigma^j(a_0)|, 0 \le j \le s.$$

Then

$$\overset{\circ}{\bigcup_{0 \le j \le s}} \int_{|\sigma^j(a_0)|} \tau_j(\omega) = \mathbb{N}^+ \tag{5.11}$$

where ω is the fixed point of σ.

It is clear that (5.11) follows from $\chi(\omega; a_j) = \int_{|\sigma^j(a_0)|} \tau_j(\omega)$, which is Theorem 4 of [344].

Note that the partition (5.11) is a totally nonperiodic one for all $s \ge 1$, and all $k_i \in \mathbb{Z}$ satisfying $k_s \ge k_{s-1} \ge \cdots \ge k_0 = 1$; this follows from [344], Lemma 11:

$$\lim_{n \to \infty} |\sigma^n(a_0)|_{a_i} / |\sigma^n(a_0)| = \alpha^{s-i}/(\alpha^s + \alpha^{s-1} + \cdots + \alpha + 1), \tag{5.12}$$

where $\alpha > 1$ is an algebraic number with minimal polynomial $f(x) := x^{s+1} - \sum_{0 \le i \le s} k_i x^i$; the minimality follows from [344], Lemma 10.

We remark that in general, the partition (5.11) cannot be the partition of the form (5.4). For instance, suppose that (5.11) with $s = 2$, $k_1 = k_2 = 1$ coincides with (5.4) corresponding to some infinite word $\omega = \omega(L)$, where $L = \{t(1, \alpha_1, \alpha_2) + \beta; t \in \mathbb{R}_+\}$. Then (5.12) implies that $\alpha_i = \alpha^{-i}(i = 1, 2)$. The minimality of $f(x)$ implies that $1, \alpha_1, \alpha_2$ are linearly independent over $\mathbb{Q}$. Hence, $p_\omega(n) = n^2 + n + 1$, which contradicts that $p_\omega(n) = 2n + 1$ is the complexity of the fixed point of the substitution σ with $s = 2$, $k_1 = k_2 = 1$ (the fixed point is an *Arnoux-Rauzy sequence* , see [42], [146]).

On the other hand, in the case $s = 1$, the partition (5.11) turns out to be the partition (5.2); that will be seen by the following argument : Proposition 5.2.1 with $s = 1$ implies $f(x) = x^2 - kx - 1$ $(k := k_1)$, so that $\alpha = (k + (k^2 + 4))^{1/2}/2$. Setting $\chi(\omega; a_i) = \{{t_1}^{(i)} < {t_2}^{(i)} < \cdots < {t_n}^{(i)} < \ldots\}$, for $i = 0, 1$, we get by Proposition 5.2.1

$${t_n}^{(1)} = kn + {t_n}^{(0)}, \qquad \chi(\omega; a_0) \overset{\circ}{\bigcup} \chi(\omega; a_1) = \mathbb{N}^+. \tag{5.13}$$

Noting that the sets $\chi(\omega; a_i)$ are uniquely determined by (5.13), and

$$[\eta_1 n] = kn + [\eta_0 n], 1/\eta_0 + 1/\eta_1 = 1 \quad (\eta_0 := 1 + 1/\alpha, \quad \eta_1 := 1 + \alpha),$$

we obtain $\chi(\omega; a_i) = \{[\eta_i n];\ n \in \mathbb{N}^+\}$, for $i = 0, 1$.

5.2.2 Some transcendence results

Definition 5.2.1. *Let* $G = G_g := \{0, 1, \ldots, g-1\}$, *where* $2 \le g \in \mathbb{N}^+$. *Let* $\tau : \mathcal{A} \to G^\star$ *be a monoid morphism such that* $\tau(a) \neq \varepsilon$ *for all* $a \in \mathcal{A}$. *We denote by* $(0.\tau(w))_g$ *the number defined by* $\sum_{i\ge1} w_i/g^i$, *where* $\tau(w) = w_1w_2w_3\cdots \in G^\infty$, $w_i \in G$.
We say ω *is* transcendental *if* $(0.\tau(w))_g$ *is transcendental for an integer* g *and a morphism* τ.

The fixed point ω of σ is not only totally nonperiodic, but also transcendental:

Proposition 5.2.2. *([344], Theorem 3) Let* ω *be as in Proposition 5.2.1,* $g \ge 2$ *an integer,* τ *a monoid morphism such that* $\tau(a) \neq \varepsilon$ *for all* $a \in \mathcal{A}$, *and*

$$\mathit{rank}\ (|\tau(a_i)|_j)_{0\le i\le s,\, 0\le j\le g-1} > 1.$$

Then the number $(0.\tau(w))_g$ *is transcendental.*

The key for the proof of Proposition 5.2.2 is to show that the infinite word ω has a prefix which is a $(2+\epsilon)$-power of a nonempty word (see Proposition 5.3.6 below, and [344], Lemma 13); that can be connected with Roth's theorem. A stronger argument works in [146], where S. Ferenczi and C. Mauduit make use of a theorem of Ridout ([236], pp. 147-148) instead of Roth's theorem. We shall mention their results in the following section, see Proposition 5.3.6 below.

5.3 Similis partitions

5.3.1 A linguistic problem

Let D be a subset of $\mathbb{N}^+$, with $1 \in D$. In some cases, we can show that there exists a subset Γ such that

$$\overset{\circ}{\bigcup_{d\in D}} d\Gamma = \mathbb{N}^+ \qquad (D \neq \emptyset, \{1\}). \tag{5.14}$$

Such a partition will be referred to as a *similis partition* (of $\mathbb{N}^+$ with respect to D). Some results on similis partitions are to be found in [343]. Let us mention that a simple example of similis partitions comes from a linguistic phenomenon in Hungarian and Japanese language that are probably well-known to linguists, see [345, 355]: Numerals one, two, three, four, ... in Hungarian (respectively Japanese) are egy, kettŏ, három, négy, ... (hi, fu, mi, yo, ...). We thus can make the following diagram, where in each language, underlined consonants of two numerals in each row are common, or they have a resemblance (e.g., n and ny= palatalized n in the 3rd stage of the diagram); and simultaneously, in each row, the number corresponding to the right group is exactly twice the left:

	Γ		2Γ	
1):	1	egy (hi←fi←pi)	2	kettő (fu←pu)
2):	3	három (mi)	6	hat (mu)
3):	4	négy (yo)	8	nyolc (ya)
4):	5	öt (itsu←itu)	10	tíz (nto)

Here, among the numerals in Japanese language that are written in parentheses, for instance, itsu←itu indicates that the contemporary Japanese word itsu comes from the old Japanese word itu, that is a kind of palatalization. If we look at the numerals of older Japanese in parentheses, the consonants correspondence turns out to be an exact one. (Related to vowels, see, e.g., [168]; vowel harmony is also common in Hungarian and old Japanese.)

Considering what will happen, apart from numerals in natural language, when we formally extend the diagram downwards, we get a similis partition (5.14) with $D = \{1, 2\}$, which is uniquely determined. In fact, it is clear that $\gamma_1 := 1 \in \Gamma$, so that $2\gamma_1 \in 2\Gamma$, which gives the first stage 1) of the diagram. Now, consider the smallest positive integer γ_2 among the numbers that have not appeared in the stage 1). Then the minimality of γ_2 implies $\gamma_2 \in \Gamma$, otherwise $\gamma_2 \in 2\Gamma$, so that $\gamma_2 > \gamma_2/2 \in 2\Gamma$, i.e., the second stage is of the form $\gamma_2/2 \in \Gamma$, $\gamma_2 \in \Gamma$, which contradicts the minimality of γ_2. (Forget that $\gamma_2 \in \Gamma$ follows from $\gamma_2 = 3$ is odd; we shall see that $\gamma_3 (= 4)$ is even in the following argument.) Suppose that we have obtained a diagram with stages 1)-n). Consider the number γ_{n+1} defined to be the smallest positive integer that differs from all the numbers appearing in the stages 1)-n). Then $\gamma_{n+1} \in \Gamma$ follows from its minimality. We can continue the process, and we must have $\Gamma = \{\gamma_1, \gamma_2, \gamma_3, \dots\}$ as far as all the numbers $d\gamma_{n+1}$ $(d \in D)$ are different from the numbers $d\gamma_m$ $(d \in D, 1 \le m \le n)$. Hence, noting that the argument given above is valid for any nonempty, finite or infinite, subset D of $\mathbb{N}$, we obtain

Proposition 5.3.1. *If there exists a similis partition (5.14) for a given nonempty subset D of $\mathbb{N}^+$, then the partition is uniquely determined by the set D.*

5.3.2 The Hungarian-Japanese partition

On the other hand, it is clear that a similis partition (5.14) for $D = \{1, 2\}$ exists, since $\Gamma = \{2^{2j}m;\ j \ge 0, m \ge 1, m \text{ is odd }\}$ satisfies (5.14), that will be referred to as the *H.-J. (Hungarian-Japanese) partition.* The H.-J. partition can be easily generalized as

Proposition 5.3.2. *Let $D = D(k; q_1, \dots, q_k; e_1, \dots, e_k)$ be a set defined by*

$$D := \{\textstyle\prod_{1 \le i \le k} q_i^{j_i}\ ;\ 0 \le j_i \le e_i \quad (1 \le i \le k)\}, \tag{5.15}$$
$$k \ge 1, q_i \ge 2, e_i \ge 1\,(1 \le i \le k), G.C.D.(q_i, q_j) = 1 \textit{ for all } i \ne j.$$

Then

$$\Gamma := \{ \prod_{1 \le i \le k} q_i^{(e_i+1)j_i} m \,;\; j_i \ge 0, m \ge 1,\; m \not\equiv 0 \textit{ modulo } q_j^{e_j+1},\; 1 \le j \le k\}$$

satisfies (5.14).

We have conjectured that if a similis partition (5.14) is a partition of $\mathbb{N}^+$ into finite components, then there exist numbers k, and $q_1, \ldots, q_k, e_1, \ldots, e_k$ satisfying (5.15); that is probably still open. It is easily seen that there are no partitions (5.14) for some explicitly given D which are not of the form (5.15), see [343], Theorem 12. For example, if we take $D = \{1,2,3\}$, and trace the uniqueness proof of the uniqueness of (5.14) above, we see $2\gamma_4 = 12 = 3\gamma_2$, which contradicts the fact that (5.14) is a disjoint union. Proposition 5.3.2 can be extended to infinite partitions with respect to D given by (5.15) with $0 \le j_i$ for some indices i instead of $0 \le j_i \le e_i (1 \le i \le k)$:

$$\begin{aligned} D := \{ & \prod_{1 \le i \le k-h} q_i^{j_i} \cdot \prod_{k-h+1 \le i \le k} q_i^{j_i} \,; \\ & 0 \le j_i \le e_i \; (1 \le i \le k-h),\; 0 \le j_i \; (k-h+1 \le i \le k) \; \} \\ & k \ge 1, k \ge h \ge 1, \\ & q_i \ge 2, e_i \ge 1 \, (1 \le i \le k-h),\; G.C.D.(q_i, q_j) = 1 \text{ for all } i \ne j. \end{aligned}$$

For D given as above, we can show

$$\begin{aligned} \Gamma = \{ & \prod_{1 \le i \le k-h} q_i^{(e_i+1)j_i} m \,; \\ & j_i \ge 0,\; m \ge 1,\; m \not\equiv 0 \mod q_i^{e_i+1},\; 1 \le i \le k-h, \\ & G.C.D.(m, q_{k-h+1} \ldots q_k) = 1 \; \} \; (k \ge 2). \end{aligned}$$

If $k = h = 1$, then $\Gamma = \mathbb{N}^+ \backslash q_1 \mathbb{N}^+$, and the partition (5.14) is periodic (this case is not interesting). We remark that for some infinite partitions (5.14), D is not always of the form above. For instance, if we take $D = \{p_i^{j_i};\; j_i \ge 0, (0 \le i \le s)\}$ with prime numbers p_i $(p_0 > p_1 > \cdots > p_s,\; s \le 1)$, then

$$\Gamma = \{(p_0 \ldots p_s)^i m\},;\; i \ge 0,\; m \ge 1,\; G.C.D.(m, p_1 \ldots p_s) = 1 \; \}$$

satisfies (5.14). By the way, we remark that a sequence ω over $\mathcal{A}_s$ defined by

$$\begin{aligned} & \omega_n := a_i \text{ if } m_n \in \{p_i^j;\; j \ge 0\} \\ & (\{1 < m_1 < m_2 < \cdots < m_n < \ldots\} := D = \{p_i^{j_i};\; j_i \ge 0 \; (1 \le i \le s)\}) \end{aligned}$$

coincides with a word $\omega(L)$ defined by the billiards in I^{s+1} with $\underline{\alpha} = (\alpha_0 \ldots \alpha_s), \underline{\beta} = \underline{0}, \alpha_i = \log p_i / \log p_0$, see [294].

5.3.3 Some properties of the H.-J. partition

Return new to the first example of (5.14), the H.-J. partition. Let us show the following:

Proposition 5.3.3. *The word* $\omega = \omega_1\omega_2\omega_3\ldots(\omega_n \in S_1)$ *corresponding to the H.-J. partition is a totally nonperiodic word, which is a fixed point of a substitution.*

Proof. We mean by UV the set $\{uv; u \in U, v \in V\}$, by U^* the set $\{u_1\ldots u_n; u_i \in U(1 \le i \le n), n \ge 0\}$ for subsets U, V of a monoid, and by Γ (with $1 \in \Gamma$) the component of the H.-J. partition. Let E_g denote the base-g expansion of $\gamma \in \mathbb{N}$ $(E_g(0) := \varepsilon)$, and W^n $(W \in \mathcal{A}_g^*)$ be the word obtained by concatenating n copies of W.

We have $\gamma \in \Gamma$ if and only if $E_2(\gamma) = u0^{2n}$, where $n \ge 0$, and $u \in \{0,1\}^*$ is a word having 1 as prefix, and suffix. Hence $\gamma \in \Gamma$ if and only if $E_2(\gamma-1) = v1^{2n}$, where $v \in G_1^* = \{0,1\}^*$, and $n \ge 0$. Consequently the set $0^*E_2(\gamma-1);\ \gamma \in \Gamma\}$ coincides with the language accepted by the automaton M defined by

$$M := (\mathcal{A}_1, G_1, \delta, a_0, \{a_0\})$$

with the transition function δ

$$\delta(a_0, 0) := a_0,\ \delta(a_0, 1) := a_1,\ \delta(a_1, i) := a_0,\ (i = 0, 1).$$

For the definitions and notation related to automata, see [178]. Therefore, noting that $\omega = \delta(a_0, E_2(0))\ldots\delta(a_0, E_2(n-1))\ldots$, we see that ω is the fixed point of the substitution over $\mathcal{A}_1$ defined by

$$\sigma(a_0) := a_0a_1,\ \sigma(a_1) := a_0a_0. \tag{5.16}$$

Using this fact shown above, we can prove that ω is a totally nonperiodic word in the following manner. We remark that so far as similis partition are concerned, nonperiodicity implies total nonperiodicity. Hence it suffices to show the nonperiodicity of ω.

Suppose that ω is an ultimately periodic word, then $\theta := (0.\tau(\omega))_2 \in \mathbb{Q}$ $(\tau(a_i) := i)$. We put $a = a_0$, $b = a_1$, $\theta_n = (0.\tau(u_n)^*)_2$ $(u_n = \sigma^n(a))$, where u^* denotes the periodic word $uuu\ldots$ for a nonempty word u. We write $u\neg v$ if v is a prefix of u. The binary relation $\neg$ is transitive. In view of (5.16), we get $u_2 = abaa = u_1u_0^2$, so that $u_{n+2} = u_{n+1}u_n^2$ for all $n \ge 0$, $|u_n| = 2^n$, and $\omega\neg u_{n+1}\neg u_nu_{n-1}$. Hence, we obtain $|\theta - \theta_n| \le 2^{-3.2^{n-1}}$. For any $n \ge 1$, we can put $\theta_n = E_2^{-1}(u)/(2^!2^n - 1)$ with certain $u \in 1G_1^*$. Let θ_n equal P_n/Q_n, $G.C.D.(P_n, Q_n) = 1$. Then $|\theta_n - P_n/Q_n| \le Q_n^{-3/2}$, which together with $\theta \in \mathbb{Q}$ implies that $\{P_n/Q_n; n \ge 0\}$ is a finite set. Therefore $\theta_i = \theta_{i+j}$ for some $i \ge 0$ and $j \ge 1$, so that $u_{i+j} = u_i^{2^j}$. Since $u_{j+1} = u_{i+j-1}u_{i+j-1}^2$, we get $u_{i+j-1} = u_i^{2^{j-1}}$, and inductively, $u_{i+1} = u_i^2$. By $u_{i+1} = u_i u_{i-1}^2$, we get $u_i = u_{i-1}^2$. Repeating the argument, we obtain $u_1 = u_0^2$ which contradicts $u_1 = ab \neq aa = u_0^2$. ∎

Remark. By direct computation, we see that $\partial\Gamma = 2112221121121122\ldots$ for the H.-J. partition. We can show that the sequence (or word) $\partial\Gamma$ is the fixed point of a substitution over $\{1,2\}$ in the following manner. Let ω be the fixed point of the substitution σ (5.16). Noting that bb does not occur in ω, we can factorize ω into two words $A := ab$ and $B := a$, and we get a new word $\tilde{\omega}$ over $\{A, B\}$:

$$\begin{array}{rcccccccccccccc} \omega = & ab & a & a & ab & ab & ab & a & a & ab & a & a & ab & \ldots, \\ \tilde{\omega} = & A & B & B & A & A & A & B & B & A & B & B & A & \ldots, \end{array}$$

and noting that ω is the fixed point of σ, we see that $\tilde{\omega}$ is the fixed point of a substitution τ over $\{A, B\}$:

$$\tau : \quad \begin{array}{ll} A = ab & \longrightarrow \sigma(ab) = abaa = ABB \\ B = a & \longrightarrow \sigma(a) = ab = A. \end{array}$$

Since $\partial\Gamma = \zeta(\tilde{\omega})$ with $\zeta(A) = 2,\ \zeta(B) = 1$, $\partial\Gamma$ is the fixed point of the substitution $2 \to 211,\ 1 \to 2$.

5.3.4 General case

We can generalize all the statements given above for the H.-J. partition to those for the partition with $D = \{q^i;\ 0 \le i \le e\}$ as in Proposition 5.3.4-5.3.5, by considering the automaton

$$M_{e,q} := (\mathcal{A}_e, G_q, \delta, a_0, \{a_0\})$$

with a transition function $\delta{=}\delta_{e,q}$ defined by

$$\begin{array}{l} \delta(a_i, j) := a_0,\ \ \delta(a_i, q-1) := a_{i+1},\ \ (0 \le i \le e-1,\ \ 0 \le j \le q-2), \\ \delta(a_e, j) := a_0,\ \ (0 \le j \le q-1) \end{array}$$

Proposition 5.3.4. *Let ω be the word corresponding to a similis partition (5.14) with respect to $D = \{q^i;\ 0 \le i \le e\}$ $(e \ge 1,\ q \ge 2)$. Then ω is a totally nonperiodic word over $\mathcal{A}_e$, which is the fixed point of a substitution over $\mathcal{A}_e$ defined by*

$$\sigma(a_i) := a_0^{q-1} a_{i+1}\ \ (0 \le i \le e-1),\ \ \sigma(a_e) := a_0^q. \tag{5.17}$$

Proposition 5.3.5. *Let (5.14) be a similis partition with respect to D as in Proposition 5.3.4. Let $\tau : \mathcal{A}_e^* \to \mathcal{A}_e^*, \kappa : \mathcal{A}_e^* \to \{1,2\}^*$ be the morphisms defined for $q = 2$ by;*

$$\begin{aligned}
&\tau(a_i) := a_0 a_{i+1}, \ \ (0 \le i \le e-2),\\
&\tau(a_{e-1}) := a_0 a_e^2,\\
&\tau(a_e) := a_0,\\
&\kappa(a_i) = 2 \ \ (0 \le i \le e-1),\\
&\kappa(a_e) = 1,
\end{aligned}$$

and for $q > 2$:

$$\begin{aligned}
&\tau(a_0) := a_0^{q-2} a_1,\\
&\tau(a_i) := a_0^{q-2} a_1 a_0^{q-2} a_{i+1}, \ \ (0 \le i \le e-1),\\
&\tau(a_e) := a_0^{q-2} a_1 a_0^{q},\\
&\kappa(a_i) = 1 \ \ (0 \le i \le e-1),\\
&\kappa(a_e) = 2.
\end{aligned}$$

Then $\partial\Gamma$ is a nonperiodic word over $\{1,2\}$, which is given by

$$\partial\Gamma = \kappa(\omega'),$$

where ω' is the fixed point of τ.

Remark. Let ω be as in Proposition 5.3.4. Then, in view of the locally catenative formula $\sigma^{n+e}(a_0) = (\sigma^{n+e-1}(a_0))^{q-1} \dots (\sigma^{n+1}(a_0))^{q-1}(\sigma^n(a_0))^q$, we can easily find that the frequency of a_i appearing in ω is rational for all i. This fact together with the nonperiodicity of ω implies that a similis partition with respect to D given by (5.15) can neither be a partition (5.4) nor a partition (5.11).

Proof. Let us consider a similis partition (5.14) with respect to $D = \{1,2,3,6\}$. Then $\Gamma = \{2^{2i}3^{2j}k;\ i \ge 0,\ j \ge 0,\ k \ge 1,\ G.C.D.(2.3,k) = 1\}$ which equals $\{2^{2i}k;\ i \ge 0,\ (2,k) = 1\} \cap \{3^{2i}k;\ i \ge 0,\ (3,k) = 1\}$. It is clear that

$$\Gamma = \chi(\omega(1,2);a_0) \cap \chi(\omega(1,3);a_0),$$

where $\omega(e,q)$ is the fixed point of a substitution over $\mathcal{A}_e$ defined by (5.17). In general, we can show, considering the languages accepted by the automata $M_{e_i,q_i,j_i} := (S_{e_i}, G_{q_i}, \delta_{e_i,q_i}, a_0, \{a_{j_i}\})$ $(0 \le j_i \le e_i \ \ (1 \le i \le k))$, that

$$q_1^{j_1} \dots q_k^{j_k}\Gamma = \bigcap_{1 \le i \le k} \chi(\omega(e_i,q_i);a_{j_i}) \ \ (0 \le j_i \le e_i \ \ (1 \le i \le k)) \tag{5.18}$$

holds for any finite similis partition with respect to D given by (5.15). We denote by Ω the word (strictly over $\prod_{1 \le i \le k}(e_i+1)$ letters) corresponding to a partition given by Proposition 5.3.2. Then, it follows from (5.18) that Ω is an interpretation of $\omega(e_i,q_i)$ (i.e., whenever the i-th symbol counted from the beginning differs from the j-th symbol in ω, then the same happens in Ω, see [306]). Hence, Ω is not an ultimately periodic word by Proposition 5.3.4. Therefore, any similis partition given according to Proposition 5.3.2 is totally nonperiodic. ■

Related to the transcendence of the word $\partial\Gamma$ and the word corresponding to a finite similis partition, the following result obtained by S. Ferenczi and C. Mauduit is useful (let us recall that a substitution σ over the alphabet $\mathcal{A}$ is called *primitive* if $(|\sigma^n(a)|_b)_{a,b\in\mathcal{A}}$ is a positive matrix for some n):

Proposition 5.3.6. *([146], Proposition 5) If the expansion of θ in some base k is not ultimately periodic fixed point of a primitive substitution, and does contain at least one word of the form $v^{2+\varepsilon}$ (that is, vvv' for a nonempty word v and a prefix v' of v with $|v'| \geq \varepsilon|v|$ $(\varepsilon > 0)$), then θ is transcendental.*

If we apply Proposition 5.3.6 to the word ω in Proposition 5.3.4, and note that the transcendence of ω implies the transcendence of Ω, we obtain Proposition 5.3.7 (respectively Proposition 5.3.8) by Proposition 5.3.4 (respectively Proposition 5.3.5) as follows:

Proposition 5.3.7. *Let (5.14) be a finite similis partition with respect to D given by (5.15) with the word Ω corresponding to (5.14). Then Ω is transcendental.*

Proposition 5.3.8. *Let Γ be as in Proposition 8. Then $\partial\Gamma$ is transcendental.*

5.4 Log-fixed points and Kolakoski words

5.4.1 Definition

A word (or a sequence) ξ' over $\{1, 2\}$ is referred to as a *Kolakoski word* if the word defined by its run-lengths is equal to ξ itself:

$$\begin{aligned} \xi' = {} & 22\ 11\ 2\ 1\ 22\ 1\ 22\ 11\ 2\ 11\ 22\ 1\ 2\ 11\ 2\ 1\ 22\ \ldots \\ & 2\ \ 2\ \ 1\,1\,2\ \ 1\,2\ \ 2\ 1\ 2\ \ 2\ \ 1\,1\,2\ \ 1\,1\,2 \cdots = \xi', \end{aligned}$$

where we mean by a run a maximal subword consisting of identical letters, see [208], and [116] with the references therein. The word $\xi := 1\xi'$ is the only other word having this property. It can be easily seen that ξ is not an ultimately periodic word, see [358]. Related to the complexity $p(n) := p_\xi(n)$, F. M. Dekking has shown that

$$\forall n, p(n) \leq n^{7.2};$$

he has furthermore conjectured (see [114, 116]) that

$$p(n) \asymp n^{log 2/log(3/2)}.$$

Let $\mathcal{A}$ be an alphabet with Card $\mathcal{A} \geq 2$. We denote by $\mathcal{A}^\times$ the set

$$\mathcal{A}^\times := (\mathcal{A}^* \cup \mathcal{A}^\infty) \backslash (\bigcup_{a \in \mathcal{A}} \mathcal{A}^*\{a^*\}).$$

i.e., $\mathcal{A}^\times$ is the set of all finite or infinite words that are different from all the words of the form ua^* ($u \in \mathcal{A}^*, a \in \mathcal{A}$). We shall write $\mathcal{A}^*w$ instead of $\mathcal{A}^*\{w\}$. For any word $\omega = \omega_1\omega_2 \ldots \omega_n \cdots \in \mathcal{A}^\times$, we can define two words log ω, and base ω by

$$\begin{aligned} &\log \omega := e_1 e_2 e_3 \ldots, \qquad \text{base } \omega := b_1 b_2 b_3 \ldots, \\ &\text{if } \ \omega = b_1^{e_1} b_2^{e_2} b_3^{e_3} \ldots \ (e_i \geq 1, \ b_i \in \mathcal{A}, \ b_i \neq b_{i+1}, \text{ for all } i \geq 1). \end{aligned}$$

In what follows, we take $\mathcal{A} \subset \mathbb{N}^+$.

Definition 5.4.1. *A word $\omega \in \mathcal{A}^\times$ satisfying $\omega = log\ \omega$ will be referred to as a* log-fixed point. *The* Kolakoski word *ξ is defined to be a log-fixed point with base $\xi = (12)^*$.*

If Card $\mathcal{A} = 2$, then Card $\{\omega \in \mathcal{A}^\infty ; \omega = \log \omega\} = 2$; if Card $\mathcal{A} \geq 3$, then the set $\{\omega \in \mathcal{A}^\infty ; \omega = \log \omega\}$ has continuum cardinality, since so does the set $\{\text{base } \omega\}$. It can be easily seen by a similar manner to that given by [358] that all the log-fixed points are not ultimately periodic.

Consider for instance a log-fixed point ω with base $\omega = (26)^*$, and factorize it into the words of length 2:

$$\begin{array}{lllllllllll} \omega & = 22 & 66 & 22 & 22 & 22 & 66 & 66 & 66 & 22 & 66 \ldots \\ & = \mathrm{A} & \mathrm{B} & \mathrm{A} & \mathrm{A} & \mathrm{A} & \mathrm{B} & \mathrm{B} & \mathrm{B} & \mathrm{A} & \mathrm{B} \ldots \\ & = W_1 & W_2 & W_3 & W_4 & W_5 & W_6 & W_7 & W_8 & W_9 & W_{10} \ldots . \end{array}$$

Then it is clear that W_i is $A := 22$ or $B := 66$ (since the length of the period of base ω is 2, which divides $2, 6 \in \mathcal{A}$), and ω as a word over $\{A, B\}$ is invariant under the morphism

$$\begin{aligned} A = 22 &\longrightarrow 2266 = AB \\ B = 66 &\longrightarrow 222222666666 = AAABBB. \end{aligned}$$

Note that such an argument does not work at all for the Kolakoski words, but it can be applied to some general cases:

Proposition 5.4.1. *Let $s \geq 1$ be an integer, $\mathcal{A} = \{a_0, \ldots, a_s\} \subset \mathbb{N}^+$ such that s divides a_i for all $1 \leq i \leq s$. Let σ be a substitution over $\{A_0, \ldots, A_s\}$ defined by $\sigma(A_i) = A_0^{a_i/s} \ldots A_s^{a_i/s}$ $(0 \leq i \leq s)$, and let Ω be its fixed point. Then the log-fixed point ω with base $\omega = (a_0 \ldots a_s)^*$ can be given by $\omega = \tau(\Omega)$, where τ is the morphism defined by $\tau(A_i) = a_i^s$ $(0 \leq i \leq s)$.*

5.4.2 Generalized substitutions

Return now to the Kolakoski word ξ'. Consider what could be the substitution σ in Proposition 5.1.1 for the infinite word ξ' in a formal sense. It would become a "substitution" defined by

$$\sigma(A_0) = A_0 A_1,\ \sigma(A_1) = A_0^{1/2} A_1^{1/2},$$

where we mean by $A^{1/2}$ a half of a symbol. Define $(W_1 W_2 \dots W_n)^{1/2}$ (each W_i is a symbol, or a half-symbol) to be a "word" $W_1 \dots W_{[n/2]}$ (respectively $W_1 \dots W_{[n/2]} W_{[n/2]+1}^{1/2}$, which is possibly a "word" containing a fourth of a symbol) for even n (respectively odd n), and define $\sigma(W^{1/2})$ to be a "word" $\sigma(W)^{1/2}$. Consider an infinite word $\Omega = \lim \sigma^n(A_0)$, then

$$A_0 \xrightarrow{\sigma} A_0 A_1 \xrightarrow{\sigma} A_0 A_1 A_0^{1/2} A_1^{1/2} \xrightarrow{\sigma} A_0 A_1 A_0^{1/2} A_1^{1/2} A_0 A_0^{1/2} \xrightarrow{\sigma} \dots$$

$$\longrightarrow \Omega = A_0 A_1 A_0^{1/2} A_1^{1/2} A_0 A_0^{1/2} A_0 A_1 A_0^{1/2} A_0 A_1 A_0^{1/2} A_1^{1/2} A_0 \dots .$$

We can define the sequence Ω to be the fixed point of a substitution over an alphabet $\{a,\ b,\ c,\ d\}$ in the usual sense, where we identify $a = A_0,\ b = A_1,\ c = A_0^{1/2},\ d = A_1^{1/2}$. The following question is thus natural: can we find any relation between ξ' and Ω ? (Probably, not!; then we must find a better treatment for half-symbols.) It will be remarkable that the word ξ is a fixed point of of the map

$$\begin{aligned} \Psi : (\mathbb{N}^+)^\infty \cup ((\mathbb{N}^+)^* \backslash \mathbb{N}^* 1) &\to \{1,\ 2,\ 3\}^\infty, \\ \Psi(\omega) &:= B_3(B_{2\to 3}(\varphi(c(\omega))) + 1/2), \end{aligned}$$

where $c,\ \varphi,\ B_{2\to 3},\ B_3$ are maps defined as follows:

1) $c : (\mathbb{N}^+)^\infty \cup (\mathbb{N}^* \backslash (\mathbb{N}^+)^* 1) \to I = [0,1],\ c(a_1 a_2 a_3 \dots) := [0; a_1, a_2, a_3, \dots]$ for $a_1 a_2 a_3 \dots \in (\mathbb{N}^+)^\infty \cup ((\mathbb{N}^+)^* \backslash (\mathbb{N}^+)^* 1)$, where the right-hand side denotes the usual continued fraction expansion;
2) $\varphi : I \to I$ is the so called *question-mark-function* introduced by Minkowski determined by the following conditions:
   (i) the map φ is continuous with $\varphi(0) = 0,\ \varphi(1) = 1$,
   (ii) $\varphi((p+p')/(q+q')) = (\varphi(p/q) + \varphi(p'/q'))/2$ for all $p, q, p', q' \in \mathbb{N}^+ \cup \{0\}$ such that $p/q,\ p'/q' \in I,\ p'q - pq' = \pm 1$;
3) $B_{2\to 3} : I \to [0, 1/2], B_{2\to 3}((0.b_1 b_2 b_3 \dots)_2) := (0.b_1 b_2 b_3 \dots)_3$ for $b_1 b_2 b_3 \dots \in \{0,1\}^\infty \backslash \{0,1\}^* 0^*$ $(B_{2\to 3}(0) := 0)$;
4) $B_3 : I \to \{0,1,2\}^\infty,\ B_3(x) = c_1 c_2 c_3 \dots$ for $x = (0.c_1 c_2 c_3 \dots)_3$ with $c_1 c_2 c_3 \dots \in \{0,1,2\}^\infty \backslash \{0,1,2\}^* 0^*$ $(B_3(0) := 0^*)$.

We can see that ξ is uniquely determined by $\Psi(\xi) = \xi$ by the fact

$$\varphi([0; a_1, a_2, a_3, \dots]) = (0.\ 0^{a_1}\ 1^{a_2}\ 0^{a_3}\ 1^{a_4}\ 0^{a_5}\ \dots)_2, \tag{5.19}$$

see [270].

The problem of the existence of frequencies of words is known as Keane's problem: the question is whether the frequency of 1 in ξ exists, and whether it equals $1/2$, [200]. This is still open. If the frequency does not exist, or if it equals $1/2$ (it probably does!), then it is easy to see that the words ω corresponding to the partition (5.4), (5.12), or (5.15) cannot be the word ξ.

5.4.3 Some connections with continued fraction expansions

Instead of (5.19), we may ask about the existence of a number $x \in I$ satisfying

$$x = [0; a_1, a_2, a_3, \dots] = (0.a_1 a_2 a_3)_g, \tag{5.20}$$

where

$$a_n \in \mathbb{Z},\ 1 \le a_n \le g-1\ (i \ge 1). \tag{5.21}$$

Such a number x exists for a square number $g = h^2$, for $2 \le h \in \mathbb{Z}$, since $[0; h] = 1/h = h/g = (0.h)_g$; this is not interesting. Now, we ask for an irrational number $x \in I$ satisfying (5.20) with (5.21). If we take $g = 10$, then by simple computations, we can show that such a number does not exist. If we consider (5.20) with

$$a_1 a_2 a_3 \cdots \in (\{0, \dots, g-1\}^\infty \backslash \{0, \dots, g-1\}^* 0^* \tag{5.22}$$

instead of (5.21), then it seems very likely that a number $x \in I$ satisfying (5.20) exists; a computation says that

$$[0; 3, 3, 5, 8, 3, 4, 7, \dots] = (0.3358347\dots)_{11},$$

where we mean, for example,

$$[0; 3, 3, 5, \dots, 1, 1, 9, 10, \underbrace{0, \dots, 0}_{\text{odd number of 0s}}, 2, 9, \dots] = [0; 3, 3, 5, \dots, 1, 1, 9, 10+2, 9, \dots],$$

$$[0; 3, 3, 5, \dots, 1, 1, 9, 10, \underbrace{0, \dots, 0}_{\text{even number of 0s}}, 2, 9, \dots] = [0; 3, 3, 5, \dots, 1, 1, 9, 10, 2, 9, \dots].$$

The difficulty in showing the existence of a number x satisfying (5.20) for $g = 11$ comes from the possibility of a long run of 0's. Probably, the length of a run of 0's which begins by the n-th symbol counted from the begining is bounded by a function of n taking sufficiently small values; and probably, such an irrational number x, satisfying (5.20) with (5.22) exists for infinitely many g. It is clear that if an irrational number x, satisfying (5.20) with (5.21) exists, then x is an irrational number being different from all the quadratic irrationals. Note that a periodic, or a non-periodic infinite continued fraction with (5.22) can be a rational number, for instance $[0;3,1,0,3,0,0,0,5,0,0,0,0,0,7,\ldots] = [0;3,\infty] = 1/3$, $[0;3,1,0,7,0,7,0,7,\ldots] = 1/3$.

5.5 Problems

Let us end this chapter by evoking some open problems.

1. We denote by $\psi_i(z)$ the analytic function on the unit disc defined by

$$\psi_i(z) = \psi_i(z;\omega) := \sum_{n\in\chi(\omega;a_i)} z^n \quad (0 \le i \le s)$$

for $\omega \in \mathcal{A}^\infty$, where $\mathcal{A} = \{a_0,\ldots,a_s\}$, and we take ω to be the word $\omega(L)$ defined by the billiard as in Section 5.1 with $\{t\underline{\alpha} + \underline{\beta}; t \in \mathbb{R}_+\}$. Then

$$\psi_i(z) = \sum_{1\le n<\infty} z^{\sum_{0\le j\le s}[\alpha_i^{-1}\alpha_j n - \alpha_i^{-1}\{\beta_i\}\alpha_j + \{\beta_j\}]}$$

follows from Proposition 5.1.1. We suppose that $\alpha_0,\ldots,\alpha_s$ are linearly independent over $\mathbb{Q}$.

Question. Is the number $\sum_{0\le i\le s} c_i \cdot \psi_i(g^{-1})$ $(c_i, g \in \mathbb{Z},\ g \ge 2)$ always transcendental except for the case where $c_i = c$ for all i?

It follows from Proposition 2 in [146], that for $s = 1$, $\sum_{0\le i\le s} c_i \cdot \psi_i(g^{-1})$ $(c_0 \ne c_1)$ is transcendental since ω is Sturmian for $s = 1$, as we have mentioned in Section 5.1. For a proof of the transcendence of $\omega = \omega(L)$, it suffices to show the transcendence of $\psi_i(g^{-1};s)$ for some i. Taking $s = 2$, $\alpha_0 = 1$, $\beta_0 = 0$, $0 < \beta_i < 1$, we have

$$\psi_0(z) = \sum_{1\le n<\infty} z^{[(\alpha_1+1)n+\beta_1]+[\alpha_2 n+\beta_2]}. \tag{5.23}$$

Question. Can we show the transcendence of the value $\psi_0(g^{-1})$? (Probably, yes; (5.23) is a simple expression similar to that in the case $s = 1$.) For problems related to linear independence and transcendence for $\psi_i(z)$, see [342], (i)-(v), pp.213.

2. It is difficult to show that there is no number x satisfying (5.20) with (5.22) for $g = 10$. The difficulty comes from that, for example,

$$[0;2,0,2,1,0,0,9,0,8,\ldots] = 0.202100908\ldots \text{ (in base 10)}$$

may be a solution for (5.20).
We may ask about the existence of a number x satisfying (5.20) for irrational g, for example,

$$[0; 3, 2, 4, 6, 9, 8, 2, \ldots] = 0.3246982\ldots \text{ (in base } \beta = ((1+5^{1/2})/2)^5)$$

may be such a number.
It is easy to show that there exists a number $\beta = \beta(\omega)$ satisfying

$$[0; a_1, a_2, a_3, \ldots] = 0.a_1a_2a_3\ldots \text{ (in base } \beta) \tag{5.24}$$

for any given $\omega = a_1a_2a_3 \cdots \in \{0, 1, \ldots, h\}^\infty \backslash \{0, 1, \ldots, h\}^* 0^*$. For instance, for the Kolakoski sequence ξ,

$$[0; 1, 2, 2, 1, 1, 2, 1, 2, 2, \ldots] = (0.122112122\ldots)_\beta, \ \beta = 2.837559\ldots;$$

for the fixed point of a substitution $1 \to 10, \ 0 \to 1$,

$$[0; 1, 0, 1, 1, 0, 1, 0, 1, 1, \ldots] = (0.101101011\ldots)_\beta, \ \beta = 2.729451\ldots.$$

Question. Can we show the transcendence of such a number $\beta(\omega)$ for a nonperiodic fixed point ω of a substitution?

Let us mention two conjectures.

Conjecture 5.5.1. For any integer $g = h^2+h+i (i = 0, 1, \ h \geq 3)$, there exists an irrational number satisfying (5.20) with (5.22); such an irrational number is always transcendental.

Conjecture 5.5.2. The number $\beta(\omega)$ defined by (5.24) is transcendental for any nonperiodic word ω.

3. Let ω be the word corresponding to the partition (5.10), i.e., $\omega = \omega(K)$ for a curve $K = \{f_0(x), \ldots, f_s(x); \ x \in \mathbb{R}_+\}$ for f_i as in Proposition 5.1.2. Suppose that $f_i(x) \in \mathbb{Q}(x)$ for all i.

   Question. Can we show that $p_{\omega(K)}(n)$ is bounded by a polynomial in n (respectively s) for fixed s (respectively n) (see [342], (vi,vii), p.214)? x

Acknowledgements. The author would like to thank Valérie Berthé and Sébastien Ferenczi for their careful reading of the original manuscript, and for their modifications to make it suitable as a chapter of this book.