Bulletin
de la SOCIÉTÉ MATHÉMATIQUE DE FRANCE

THUE–MORSE–STURMIAN WORDS AND CRITICAL BASES FOR TERNARY ALPHABETS

Wolfgang Steiner

Tome 148
Fascicule 4

2020

SOCIÉTÉ MATHÉMATIQUE DE FRANCE

pages 597-611

Le *Bulletin de la Société Mathématique de France* est un périodique trimestriel de la Société Mathématique de France.

Fascicule 4, tome 148, décembre 2020

Diffusion

Maison de la SMF
Case 916 - Luminy
13288 Marseille Cedex 9
France
commandes@smf.emath.fr

AMS
P.O. Box 6248
Providence RI 02940
USA
www.ams.org

Tarifs

Vente au numéro : 43 € ($ 64)
Abonnement électronique : 135 € ($ 202),
avec supplément papier : Europe 179 €, hors Europe 197 € ($ 296)
Des conditions spéciales sont accordées aux membres de la SMF.

Secrétariat : Bulletin de la SMF

Bulletin de la Société Mathématique de France
Société Mathématique de France
Institut Henri Poincaré, 11, rue Pierre et Marie Curie
75231 Paris Cedex 05, France
Tél : (33) 1 44 27 67 99 • Fax : (33) 1 40 46 90 96
bulletin@smf.emath.fr • smf.emath.fr

ISSN 0037-9484 (print) 2102-622X (electronic)

Directeur de la publication : Fabien DURAND

Bull. Soc. Math. France
148 (4), 2020, p. 597-611

THUE–MORSE–STURMIAN WORDS AND CRITICAL BASES FOR TERNARY ALPHABETS

by Wolfgang Steiner

Abstract. — The set of unique β-expansions over the alphabet $\{0,1\}$ is trivial for β below the golden ratio and uncountable above the Komornik–Loreti constant. Generalisations of these thresholds for three-letter alphabets were studied by Komornik, Lai and Pedicini (2011, 2017). We use a class of S-adic words, including the Thue–Morse sequence (which defines the Komornik–Loreti constant) and Sturmian words (which characterise generalised golden ratios) to determine the value of a certain generalisation of the Komornik–Loreti constant to three-letter alphabets.

Résumé (*Mots de Thue–Morse–Sturm et bases critiques pour les alphabets ternaires*). — L'ensemble des β-développements uniques avec l'alphabet $\{0,1\}$ est trivial pour β au-dessous du nombre d'or et non dénombrable au-dessus de la constante de Komornik–Loreti. Des généralisations de ces seuils pour les alphabets de trois lettres furent étudiées par Komornik, Lai et Pedicini (2011, 2017). Nous utilisons une classe de mots S-adiques comprenant la suite de Thue–Morse (qui définit la constante de Komornik–Loreti) et les mots sturmiens (qui caractérisent les nombres d'or généralisés) pour déterminer la valeur d'une certaine généralisation de la constante de Komornik–Loreti aux alphabets de trois lettres.

Texte reçu le 18 février 2019, modifié le 25 février 2020, accepté le 30 avril 2020.

Wolfgang Steiner, Université de Paris, IRIF, CNRS, F-75006 Paris, France • *E-mail :* steiner@irif.fr

Mathematical subject classification (2010). — 11A63, 68R15.

Key words and phrases. — β-expansion, Unique expansion, Golden ratio, Komornik–Loreti constant, Thue–Morse sequence, Sturmian words, S-adic words.

This work was supported by the Agence Nationale de la Recherche through the project *CODYS* (ANR-18-CE40-0007).

BULLETIN DE LA SOCIÉTÉ MATHÉMATIQUE DE FRANCE 0037-9484/2020/597/$ 5.00

doi:10.24033/bsmf.2817

1. Introduction and main results

For a base $\beta > 1$ and a sequence of digits $u_1 u_2 \cdots \in A^\infty$, with $A \subset \mathbb{R}$, let

$$\pi_\beta(u_1 u_2 \cdots) = \sum_{k=1}^{\infty} \frac{u_k}{\beta^k};$$

we say that $u_1 u_2 \cdots$ is a β-expansion of this number. This paper deals with *unique β-expansions over A*, that is, with

$$U_\beta(A) = \{\mathbf{u} \in A^\infty \colon \pi_\beta(\mathbf{u}) \neq \pi_\beta(\mathbf{v}) \text{ for all } \mathbf{v} \in A^\infty \setminus \{\mathbf{u}\}\}.$$

We know from [5] that $U_\beta(\{0,1\})$ is trivial if and only if $\beta \le \frac{1+\sqrt{5}}{2}$, where trivial means that $U_\beta(\{0,1\}) = \{\overline{0}, \overline{1}\}$, $\overline{a}$ being the infinite repetition of a. Therefore,

$$\mathcal{G}(A) = \inf\{\beta > 1 \colon |U_\beta(A)| > 2\}$$

is called the *generalised golden ratio* of A. By [6] the set $U_\beta(\{0,1\})$ is uncountable if and only if β is larger than or equal to the Komornik–Loreti constant $\beta_{\mathrm{KL}} \approx 1.787$; we call

$$\mathcal{K}(A) = \inf\{\beta > 1 \colon U_\beta(A) \text{ is uncountable}\}$$

the *generalised Komornik–Loreti constant* of A. (We can replace *uncountable* throughout the paper by *has the cardinality of the continuum.*) The precise structure of $U_\beta(\{0,1\})$ was described in [8]. For integers $M \ge 2$, $\mathcal{G}(\{0,1,\dots,M\})$ was determined by [2], and $U_\beta(\{0,1,\dots,M\})$ was described in [12, 1].

For $x, y \in \mathbb{R}$, $x \neq 0$, we have $(xu_1 + y_1)(xu_2 + y_2) \cdots \in U_\beta(xA + y)$ if and only if $u_1 u_2 \cdots \in U_\beta(A)$, thus $\mathcal{G}(xA + y) = \mathcal{G}(A)$ and $\mathcal{K}(xA + y) = \mathcal{K}(A)$. Hence, the only two-letter alphabet to consider is $\{0,1\}$. A three-letter alphabet $\{a_1, a_2, a_3\}$ with $a_1 < a_2 < a_3$ can be replaced by $\{0, 1, \frac{a_3 - a_1}{a_2 - a_1}\}$ or $\{0, 1, \frac{a_3 - a_1}{a_3 - a_2}\}$. Since $\frac{a_3 - a_1}{a_2 - a_1}$ and $\frac{a_3 - a_1}{a_3 - a_2}$ are on opposite sides of 2 (or both equal to 2), we can restrict ourselves to alphabets $\{0, 1, m\}$, $m \in (1, 2]$. Of course, it is also possible to restrict ourselves to $m \ge 2$, as in [9] (note that the alphabet $\{0, 1, m\}$ can be replaced by $\{0, 1, \frac{m}{m-1}\}$), but we find it easier to work with $m \le 2$. We write

$$U_\beta(m) = U_\beta(\{0,1,m\}), \quad \mathcal{G}(m) = \mathcal{G}(\{0,1,m\}), \quad \mathcal{K}(m) = \mathcal{K}(\{0,1,m\}).$$

It was established in [9, 14, 3] that the generalised golden ratio $\mathcal{G}(m)$ is given by mechanical words, i.e., Sturmian words and their periodic counterparts; in particular, we can restrict ourselves to sequences $\mathbf{u} \in \{0,1\}^\infty$. Calculating $\mathcal{K}(m)$ seems to be much harder, since this restriction is not possible. Therefore, we study

$$\mathcal{L}(m) = \inf\{\beta > 1 \colon U_\beta(m) \cap \{0,1\}^\infty \text{ is uncountable}\},$$

following [11], where this quantity was determined for certain intervals. We give a complete characterisation in Theorem 1.1 below.

To this end, we use the substitutions (or morphisms)

$$L: 0 \mapsto 0, \quad M: 0 \mapsto 01, \quad R: 0 \mapsto 01,$$
$$1 \mapsto 01, \quad 1 \mapsto 10, \quad 1 \mapsto 1,$$

which act on finite and infinite words by $\sigma(u_1u_2\cdots) = \sigma(u_1)\sigma(u_2)\cdots$. The monoid generated by a set of substitutions S (with the usual product of substitutions) is denoted by S^*. An infinite word $\mathbf{u}$ is a *limit word* of a sequence of substitutions $(\sigma_n)_{n\ge1}$ (or an S*-adic word* if $\sigma_n \in S$ for all $n \ge 1$), if there is a sequence of words $(\mathbf{u}^{(n)})_{n\ge1}$ with $\mathbf{u}^{(1)} = \mathbf{u}$, $\mathbf{u}^{(n)} = \sigma_n(\mathbf{u}^{(n+1)})$, for all $n \ge 1$. The sequence $(\sigma_n)_{n\ge1}$ is called *primitive*, if for each $k \ge 1$, there is an $n \ge k$, such that both words $\sigma_k\sigma_{k+1}\cdots\sigma_n(0)$ and $\sigma_k\sigma_{k+1}\cdots\sigma_n(1)$ contain both letters 0 and 1. For $S = \{L, M, R\}$, this means that there is no $k \ge 1$, such that $\sigma_n = L$ for all $n \ge k$ or $\sigma_n = R$, for all $n \ge k$. Let $\mathcal{S}_S$ be the set of limit words of primitive sequences of substitutions in S^∞. Then $\mathcal{S}_{\{L,R\}}$ consists of *Sturmian words*, and $\mathcal{S}_{\{M\}}$ consists of the *Thue–Morse word* $0\mathbf{u} = 0110100110010110\cdots$, which defines the Komornik–Loreti constant by $\pi_{\beta_{\mathrm{KL}}}(\mathbf{u}) = 1$, and its reflection by $0 \leftrightarrow 1$. We call the elements of $\mathcal{S}_{\{L,M,R\}}$, which to our knowledge have not been studied yet, *Thue–Morse–Sturmian words*. For details on S-adic and other words, we refer to [15, 4].

For $\mathbf{u} \in \{0,1\}^\infty$ and $m \in (1, 2]$, define $f_\mathbf{u}(m)$ (if $\mathbf{u}$ contains at least two 1s) and $g_\mathbf{u}(m)$ as the unique positive solutions of

$$f_\mathbf{u}(m)\,\pi_{f_\mathbf{u}(m)}(\sup O(\mathbf{u})) = m \quad \text{and} \quad (g_\mathbf{u}(m) - 1)(1 + \pi_{g_\mathbf{u}(m)}(\inf O(\mathbf{u}))) = m,$$

respectively, where $O(u_1u_2\cdots) = \{u_ku_{k+1}\cdots : k \ge 1\}$ denotes the shift orbit, and infinite words are ordered by the *lexicographic order*. For the existence and monotonicity properties of $f_\mathbf{u}(m)$ and $g_\mathbf{u}(m)$, see [3, Lemmas 3.11 and 3.12] and Lemma 2.1 below. We define $\mu_\mathbf{u}$ by

$$f_\mathbf{u}(\mu_\mathbf{u}) = g_\mathbf{u}(\mu_\mathbf{u}),$$

i.e. $f_\mathbf{u}(\mu_\mathbf{u}) = g_\mathbf{u}(\mu_\mathbf{u}) = \beta$ with $\beta\,\pi_\beta(\sup O(\mathbf{u})) = (\beta - 1)(1 + \pi_\beta(\inf O(\mathbf{u})))$.

The main result of [9] on generalised golden ratios of three-letter alphabets can be written as

$$\mathcal{G}(m) = \begin{cases} f_{\sigma(\overline{0})}(m) & \text{if } m \in [\mu_{\sigma(1\overline{0})}, \mu_{\sigma(\overline{0})}],\ \sigma \in \{L,R\}^*M, \\ g_{\sigma(\overline{0})}(m) & \text{if } m \in [\mu_{\sigma(\overline{0})}, \mu_{\sigma(0\overline{1})}],\ \sigma \in \{L,R\}^*M, \\ f_{\overline{1}}(m) & \text{if } m \in [\mu_{0\overline{1}}, 2], \\ 1 + \sqrt{m} & \text{if } m = \mu_\mathbf{u},\ \mathbf{u} \in \mathcal{S}_{\{L,R\}}; \end{cases}$$

cf. [3, Proposition 3.18], where substitutions $\tau_h = L^hR$ are used, and $f, g, \mu, \mathcal{S}$ are defined slightly differently. Our main theorem looks similar, but we need $\{L, M, R\}$ instead of $\{L, R\}$, and the roles of f and g are exchanged.