Trends in Mathematics
Information Center for Mathematical Sciences
Volume 7, Number 2, December, 2004, Pages 39–47

# STURMIAN WORDS AND $\beta$-SHIFTS: DYNAMICS AND TRANSCENDENCE

DONG PYO CHI AND DOYONG KWON

ABSTRACT. For a given Sturmian word $s$, we consider the minimal $\beta$-shift containing the shift space generated by $s$. In this note we characterize such $\beta$, and investigate its dynamical, topological and number theoretical properties.

## 1. INTRODUCTION.

Sturmian words are infinite words over binary alphabet $A$ whose factors of length $n$ are exactly $n+1$. It is known that Sturmian words are aperiodic infinite words with minimal complexity [8]. These critical words admit some equivalent definitions in different manners. They can be coded from the irrational billiards on a unit square or equivalently from the irrational rotation on $\mathbb{R}/\mathbb{Z}$ under a certain partition [13]. And they also carry the balanced properties which will be defined later [13].

Let $\beta > 1$ be a real number. We consider $\beta$-transformation $T_\beta$ on $[0,1]$ defined by $T_\beta : x \mapsto \beta x \mod 1$. Then the $\beta$-expansion of $x \in [0,1]$, denoted by $d_\beta(x)$, is a sequence of integers determined by the following rule:

$$d_\beta(x) = (x_i)_{i\geq 1} \text{ if and only if } x_i = \lfloor \beta T_\beta^{i-1}(x) \rfloor,$$

where $\lfloor t \rfloor$ is the largest integer not greater than $t$. If $\beta$ is not an integer, $d_\beta(x)$ is a sequence over the alphabet $A = \{0, 1, \cdots, \lfloor \beta \rfloor\}$. When $\beta$ is an integer, the digits $x_i$ belong to $A = \{0, 1, \cdots, \beta - 1\}$. The $\beta$-shift $S_\beta$ is the closure of $\{d_\beta(x) | x \in [0,1)\}$ with respect to the topology of $A^{\mathbb{N}}$. In [14], Parry completely characterized $S_\beta$ in terms of $d_\beta(1)$ and the lexicographic order given in $A^{\mathbb{N}}$. From Parry's result we note that the collection of $\beta$-shifts is totally ordered. The main concern of this article is about the minimal $\beta$-shift containing the shift space generated by a Sturmian word.

For $1 < \beta < 2$, we call $\beta$ a self-Sturmian number if the set $\{T_\beta^n 1\}_{n\geq 0}$ is infinite and

$$1 - \frac{1}{\beta} \leq T_\beta^n 1 \leq 1 \quad \text{for all } n \geq 0.$$

We show that if $S_\beta$ minimally contains the shift space generated by some Sturmian word then $\beta$ is a self-Sturmian number. And conversely we also show that $d_\beta(1)$ is a Sturmian word for every self-Sturmian number $\beta$. The definition of self-Sturmian number is naturally generalized for $\beta \geq 2$. Then the alphabet involved is extended

2000 *Mathematics Subject Classification:* 68R15, 11J91, 37B10.

from $\{0,1\}$ to $\{a,b\}$ with $0 \le a < b = \lfloor\beta\rfloor$. This gives a large class of specified $\beta$-transformations $T_\beta$ and moreover for such $\beta$, the diameters of closure of $\{T_\beta^n 1\}_{n\ge 0}$ are minimal in a certain sense. We also prove the transcendence of such $\beta$. That is a partial answer to the question posed by Blanchard [5].

## 2. Preliminaries.

In this section we briefly review the terminology on words.

An *alphabet* is a finite set $A$ which may have an order or not. We consider the free monoid $A^*$ generated by $A$. We mean by a *word*, any element in $A^*$, whereas a *(right) infinite word* is one in $A^{\mathbb{N}}$. $A^*$ has a natural binary operation called the *concatenation* and the *empty word* $\varepsilon$ serves as its identity element. For a word $w$, we denote by $|w|$ the *length* of $w$, i.e., if $w = a_1a_2\cdots a_n$ and all $a_i \in A$, then $|w| = n$. Sometimes a word of length 1 is called a *letter*.

A (finite) word $w$ is said to be a *factor* (a *prefix*, a *suffix*, respectively) of a finite or infinite word $u$ provided $u$ is expressed as $u = xwy$ ($u = wy$, $u = xw$, respectively) for some words $x$ and $y$ with $x$ finite. The set of factors of a word $x$ is denoted by $F(x)$. Among factors of $x$, we mean by $F_n(x)$, the set of factors with length $n$. For $X \subset A^*$ or $X \subset A^{\mathbb{N}}$, we define

$$F(X) := \bigcup_{x\in X} F(x).$$

$F_n(X)$ is defined in a similar way. We say a subset $X$ of $A^*$ is *factorial* if all factors of its elements are again in $X$, that is, $x \in X$ implies $F(x) \subset X$.

The *reversal* of a word $w = a_1a_2\cdots a_n$, where $a_1, a_2, \cdots, a_n$ are letters, is the word $\tilde{w} = a_na_{n-1}\cdots a_1$. And a *palindrome word* is a word $w$ such that $w = \tilde{w}$. In particular the empty word and all letters are palindrome words.

The *complexity function* of an infinite word $x$ is, for each integer $n \ge 0$, the cardinality of the set of factors of length $n$ in $x$. In formula, one can write

$$P(x,n) := \mathrm{Card}(F_n(x)).$$

For a subset $X$ of $A^{\mathbb{N}}$, $P(X,n)$ can be defined similarly. The *frequency* of a word is a partial function that measures the relative occurrence of $w$ in an infinite word $x$. Suppose $\mu_x^N(w)$ is the number of occurrences of $w$ in the prefix of length $N+|w|-1$ of $x$. In other words, if $w = a_1a_2\cdots a_m$ and $x = x_1x_2\cdots$, then we have

$$\mu_x^N(w) = \mathrm{Card}\left(\{j | x_{1+j} = a_1, \cdots, x_{m+j} = a_m, 0 \le j \le N-1\}\right).$$

Then the frequency of $w$ in $x$ is

$$\mu_x(w) = \lim_{N\to\infty} \frac{1}{N}\mu_x^N(w), \quad \text{if the limit exists.}$$

If an alphabet has an order, it can be extended to $A^*$ and $A^{\mathbb{N}}$ lexicographically which is called the *lexicographic order*. More precisely, suppose $A$ is an ordered alphabet. Given $x, y \in A^{\mathbb{N}}$, we denote $x < y$ if there exist a nonnegative integer $m$ such that $x_1 = y_1, \cdots, x_{m-1} = y_{m-1}$ and $x_m < y_m$. Here $x_i, y_j$'s are all letters. For finite words $x = x_1\cdots x_m$ and $y = y_1\cdots y_n$, we write $x < y$ if $x_1\cdots x_m 00\cdots <$

$y_1 \cdots y_n 00 \cdots$.

The set $A^{\mathbb{N}}$ is well endowed with a metric in a sense that the metric generates the usual product topology of $A^{\mathbb{N}}$. For any $x, y \in A^{\mathbb{N}}$, we define the distance between $x$ and $y$ by $d(x, y) = 2^{-n}$, where $n = \min\{k \geq 0 | x_k \neq y_k\}$.

## 3. Sturmian words and lexicographic order.

A *Sturmian word* is an infinite word $s$ with its complexity function $P(s, n) = n+1$ for any nonnegative integer $n$. Since $P(s, 1) = 2$, Sturmian words are forced to be infinite words over the alphabet $\{0, 1\}$ by renaming if necessary. Thus we assume $A = \{0, 1\}$ unless stated explicitly.

*Example 1.* The Fibonacci word is an infinite word defined by

$$f_0 = 0,\ f_1 = 01,\ f_{n+2} = f_{n+1} f_n,\ n \geq 0.$$

We note the sequence of $|f_n|$ is the famous integer sequence of Fibonacci numbers. By recursive arguments, we get

$$f = \lim_{n\to\infty} f_n = 0100101001001010010100100101001001 \cdots .$$

One can also find that $f$ is Sturmian. See Chapter 2 of Lothaire's book [12].

In the present section, we give two alternative characterizations of Sturmian words. The first one is the 'balanced property.'

The *height* $h(x)$ of a word $x$ is the number of the occurrences of 1 in $x$. We say a subset $X$ of $A^*$ is *balanced* if for any $x, y \in X$, $|h(x) - h(y)| \leq 1$ whenever $|x| = |y|$. An infinite word $s$ is also called *balanced* if $F(s)$ is balanced. Morse and Hedlund [13] showed the following.

**Theorem 3.1.** *Suppose $s$ is an infinite word. Then $s$ is Sturmian if and only if $s$ is aperiodic and balanced.*

In [8], Coven and Hedlund described the balanced property in more detail.

**Theorem 3.2.** *Let $X$ be a factorial subset of $A^*$. Then $X$ is unbalanced if and only if there exists a palindrome word $w$ such that both $0w0$ and $1w1$ lie in $X$.*

The *slope* of a nonempty word $x$ is the real number $\pi(x) = h(x)/|x|$. Let $x$ be an infinite balanced word, and $x_n$ be the prefix of $x$ with length $n \geq 1$. Then we see that the sequence $(\pi(x_n))_{n\geq 1}$ converges to some value $\pi(x)$(see [12]), which we will call the *slope* of $x$.

*Example 2.* The slope of Fibonacci word can be easily computed. Let $F_n = |f_n|$. Then $h(f_n) = F_{n-2}$ and

$$\pi(f) = \lim_{n\to\infty} \frac{F_{n-2}}{F_n} = \frac{1}{\tau^2},$$

where $\tau = (1 + \sqrt{5})/2$.

Next we survey Sturmian words in a different point of view.

For a real number $t$, $\lceil t \rceil$ the smallest integer not less than $t$, and $\{t\}$ is the fraction part of $t$, i.e., $t = \lfloor t \rfloor + \{t\}$. Let $\alpha$, $\rho$ be two real numbers with $0 \le \alpha \le 1$. We now define two infinite words over $\{0, 1\}$. Consider, for nonnegative integer $n$,

$$s_{\alpha,\rho}(n) = \lfloor \alpha(n+1) + \rho \rfloor - \lfloor \alpha n + \rho \rfloor,$$

$$s'_{\alpha,\rho}(n) = \lceil \alpha(n+1) + \rho \rceil - \lceil \alpha n + \rho \rceil.$$

The infinite words $s_{\alpha,\rho}$, $s'_{\alpha,\rho}$ are termed a *lower mechanical word* and an *upper mechanical word* respectively with *slope* $\alpha$ and *intercept* $\rho$. And *mechanical words* refer to either lower or upper mechanical words. Until now we have introduced the terminology 'slope' twice for infinite words. But two definitions of slope coincide for infinite balanced words. One notes that if $\rho$ and $\rho'$ differ by an integer, then $s_{\alpha,\rho} = s_{\alpha,\rho'}$ and $s'_{\alpha,\rho} = s'_{\alpha,\rho'}$ hold. Hence with no loss of generality, we assume $0 \le \rho < 1$ through the article unless stated explicitly. If $\alpha n + \rho$ is not an integer for any $n \ge 0$, we have $s_{\alpha,\rho} = s'_{\alpha,\rho}$. Otherwise in the case that $\alpha n + \rho$ is an integer for some $n > 0$, we get

$$s_{\alpha,\rho}(n-1) = 1, \quad s_{\alpha,\rho}(n) = 0,$$

$$s'_{\alpha,\rho}(n-1) = 0, \quad s'_{\alpha,\rho}(n) = 1.$$

Thus if $\alpha$ is irrational, then $s_{\alpha,\rho}$, $s'_{\alpha,\rho}$ are the same possibly except only one factor of length at most 2. Worthy of comment is the special case where $\rho$ has the value 0. If $\alpha$ is irrational, we see

$$s_{\alpha,0} = 0c_\alpha, \quad s'_{\alpha,0} = 1c_\alpha$$

for some infinite word $c_\alpha$. Here the word $c_\alpha$ is called the *characteristic word* of $\alpha$. Morse and Hedlund [13] also characterized Sturmian words in terms of mechanical words.

**Theorem 3.3.** *Suppose $s$ is an infinite word. Then $s$ is Sturmian if and only if $s$ is irrationally mechanical.*

If $s$ is some Sturmian word, we denote by $\overline{\mathcal{O}}(s)$ the shift space generated by $s$, i.e., the orbit closure of $s$. For the shift map we write $\sigma$. It is well known that irrational rotations on a circle are ergodic and their orbits are all dense.

**Proposition 3.1** ([7])**.** *Let $s$ be a Sturmian word with slope $\alpha$. Then $\overline{\mathcal{O}}(s)$ is the set of all mechanical words of slope $\alpha$.*

In the consecutive section, what we need critically is the lexicographic orders between Sturmian words.

**Theorem 3.4** ([7])**.** *Let $\alpha$ be an irrational number in $(0, 1)$. Then*

$$0c_\alpha < s_{\alpha,\rho} < 1c_\alpha \text{ for any } 0 < \rho < 1.$$

The next corollary indeed was known to Borel and Laubie [6].

**Corollary 3.4.1.** *Let $s$ be a Sturmian word with a slope $\alpha$. Then $1c_\alpha$ is the maximal element and $0c_\alpha$ the minimal one in $\overline{\mathcal{O}}(s)$. In particular we have*

$$1c_\alpha > \sigma^n(1c_\alpha) \quad \textit{and} \quad 0c_\alpha < \sigma^n(0c_\alpha)$$

*for all positive integers $n$.*

Given a Sturmian word $s$, the theorem above presents an explicit algorithm to find the maximal and minimal points of $\overline{\mathcal{O}}(s)$.

*Example 3.* Let $f$ be the Fibonacci word. For any $x \in \overline{\mathcal{O}}(f)$, we know

$$0c_{\tau^{-2}} \le x \le 1c_{\tau^{-2}} \quad \text{or}$$

$$0010010100100101001010 0100 \cdots \le x \le 1010010100100101001010 0100 \cdots .$$

And the inequality is best possible.

If two Sturmian words have different slopes, the lexicographic order is given by the following theorem.

**Theorem 3.5** ([7])**.** *For two irrational $\alpha$, $\beta$ in $(0,1)$, we have*

$$c_\alpha < c_\beta \quad \textit{if and only if} \quad \alpha < \beta.$$

## 4. $\beta$-shifts and self-Sturmian numbers.

Recall a $\beta$-shift $S_\beta$ is the closure of all $\beta$-expansions of real numbers in $[0,1)$. Just as the number 1 dominates any number in $[0,1)$, so does $d_\beta(1)$ in $S_\beta$ with respect to the lexicographic order. Parry [14] showed:

**Theorem 4.1.** *Given $\beta > 1$, let $s$ be an element of $\{0,1,\cdots,\lfloor\beta\rfloor\}^{\mathbb{N}}$.*
*If $d_\beta(1)$ is not finite, then $s$ belongs to $S_\beta$ if and only if*

$$\sigma^n(s) \le d_\beta(1) \quad \textit{for all } n \ge 0.$$

*If $d_\beta(1) = d_1 \cdots d_m 00\cdots$, then $s$ belongs to $S_\beta$ if and only if*

$$\sigma^n(s) \le d_1 \cdots d_{m-1}(d_m - 1)d_1 \cdots d_{m-1}(d_m-1)d_1 \cdots \quad \textit{for all } n \ge 0.$$

Moreover Parry also characterized sequences that can be $\beta$-expansions of 1 for some $\beta > 1$. Such sequences obey the next rule.

**Theorem 4.2.** *A sequence $s \in \{0,1,\cdots,\lfloor\beta\rfloor\}^{\mathbb{N}}$ is a $\beta$-expansion of 1 for some $\beta$ if and only if $\sigma^n(s) < s$ for $n \ge 1$. In the case, such $\beta$ is unique.*

The following proposition that is also due to Parry furnishes a total order to the collection of all $\beta$-shifts.

**Proposition 4.1.** *Suppose $\beta$, $\gamma$ $> 1$. Then*

$$\beta < \gamma \quad \textit{if and only if} \quad d_\beta(1) < d_\gamma(1).$$

**Corollary 4.1.1.** *If $\beta < \gamma$, then $S_\beta \subset S_\gamma$.*

With the above results in mind, we consider the minimal $\beta$-shift containing the orbit closure of a Sturmian word.

Let $s$ be a Sturmian word of slope $\alpha$. Then any $t \in \overline{\mathcal{O}}(s)$ lies between $0c_\alpha$ and $1c_\alpha$. For this we use the notation as $0c_\alpha \leq \overline{\mathcal{O}}(s) \leq 1c_\alpha$. This notation represents $\beta$-shift as $0^\infty \leq S_\beta \leq d_\beta(1)$. In both cases, two inequalities are best possible. By Theorem 4.2 and Corollary 3.4.1, there exists a unique $\beta \in (1,2)$ such that $d_\beta(1) = 1c_\alpha$. Thus we deduce that such $\beta$-shift is the minimal one we are searching for. Moreover the closure of $\{\sigma^n(d_\beta(1))\}_{n\geq 0}$ is equal to $\overline{\mathcal{O}}(s)$ and $0c_\alpha$, $1c_\alpha$ are accumulation points by Proposition 3.1. Here the minimal point informs us that $d_\beta(1-1/\beta) = 0c_\alpha$. We state these facts as a theorem.

**Theorem 4.3.** *Let $s$ be a Sturmian word of slope $\alpha$. If $S_\beta$ is the smallest $\beta$-shift containing $\overline{\mathcal{O}}(s)$, then $\overline{\{\sigma^n(d_\beta(1))\}_{n\geq 0}} = \overline{\mathcal{O}}(s)$ and $\beta$ is the unique solution of*

$$1 = \sum_{n=0}^{\infty} \frac{s'_{\alpha,0}(n)}{x^{n+1}}.$$

In terms of $\beta$-transformation $T_\beta$, the theorem means

**Corollary 4.3.1.** *For $\beta$ appearing in the theorem, the closure of $\{T_\beta^n 1\}_{n\geq 0}$ is contained in $[1-1/\beta, 1]$. Moreover $1-1/\beta$ and $1$ are accumulation points.*

Because of Corollary 4.3.1, we have a good reason to name such numbers after Sturm.

*Definition.* $\beta \in (1,2)$ is a *self-Sturmian number* if $d_\beta(1)$ is aperiodic and $1-1/\beta \leq T_\beta^n 1 \leq 1$ for any $n \geq 0$.

Corollary 4.3.1 reads as follows: If $S_\beta$ minimally contains all Sturmian words of the same slope, then $\beta$ is a self-Sturmian number.

**Theorem 4.4** ([7]). *If $\beta$ is a self-Sturmian number, then $d_\beta(1)$ is the Sturmian word $1c_\alpha$ for some $\alpha$.*

Now there is something to declare about pure real dynamics.

**Corollary 4.4.1.** *For any $\beta \in [1,2]$, either $\overline{\{T_\beta^n 1\}_{n\geq 0}}$ is finite or has a diameter not less than $1/\beta$.*

## 5. Dynamics of Sturmian $\beta$-transformations.

¿From now on we extend the alphabet $A = \{0,1\}$ to $\{a,b\}$ with $0 \leq a < b$.

In [5], Blanchard suggested the study of real numbers according to the ergodic properties of their $\beta$-shifts and classified $\beta$-shifts into five categories. The terminology on language theory used in the next definition is referred to [5] or the bibliography therein.

*Definition.* The set of real numbers greater than 1 is categorized into five classes according to their $\beta$-shifts:

- $\beta \in \mathcal{C}_1$ if and only if $S_\beta$ is a shift of finite type.
- $\beta \in \mathcal{C}_2$ if and only if $S_\beta$ is sofic.

- $\beta \in \mathcal{C}_3$ if and only if $S_\beta$ is specified.
- $\beta \in \mathcal{C}_4$ if and only if $S_\beta$ is synchronizing.
- $\beta \in \mathcal{C}_5$ if and only if $S_\beta$ has none of the above properties.

¿From the definition we see the following inclusions.

$$\emptyset \neq \mathcal{C}_1 \subset \mathcal{C}_2 \subset \mathcal{C}_3 \subset \mathcal{C}_4 \subset (1, \infty), \ \mathcal{C}_5 = (1, \infty) \setminus \mathcal{C}_4.$$

For any $\beta$ contained in some classes, the morphology of its $\beta$-expansion $d_\beta(1)$ is totally understood by Parry and Bertrand-Mathis.

**Proposition 5.1** ([14, 4])**.** *The following equivalences hold.*

- *$\beta \in \mathcal{C}_1$ if and only if $d_\beta(1)$ is finite.*
- *$\beta \in \mathcal{C}_2$ if and only if $d_\beta(1)$ is ultimately periodic.*
- *$\beta \in \mathcal{C}_3$ if and only if there exists $n \in \mathbb{N}$ such that the number of consecutive $0$'s in $d_\beta(1)$ is less than $n$, or equivalently the origin is not an accumulation point of $\{T_\beta^n 1\}_{n \geq 0}$.*
- *$\beta \in \mathcal{C}_4$ if and only if some word of $F(S_\beta)$ does not appear in $d_\beta(1)$, or equivalently $\{T_\beta^n 1\}_{n \geq 0}$ is not dense in $[0, 1]$.*
- *$\beta \in \mathcal{C}_5$ if and only if all words of $F(S_\beta)$ appear at least once in $d_\beta(1)$, or equivalently $\{T_\beta^n 1\}_{n \geq 0}$ is dense in $[0, 1]$.*

On the other hand Schmeling [16] determined each size of the classes, which is one of the questions asked by Blanchard [5].

**Proposition 5.2.** *$\mathcal{C}_3$ has Hausdorff dimension $1$ and $\mathcal{C}_5$ has full Lebesgue measure.*

Now we concentrate on a special class of real numbers that is contained in $\mathcal{C}_3$. For an infinite word $x$, $\mathrm{alph}(x)$ is the set of letters involved in $x$.

*Definition.* Let $\beta > 1$. We call $\beta$ a *self-Sturmian number* if $d_\beta(1)$ is a Sturmian word over a binary alphabet $A = \{a, b\}$, $0 \leq a < b = \lfloor \beta \rfloor$. In particular, $\beta$ is *maximally self-Sturmian* if it is self-Sturmian and $\mathrm{alph}\,(d_\beta(1)) = \{\lfloor \beta \rfloor - 1, \lfloor \beta \rfloor\}$.

*Remark.* Not every Sturmian word, of course can be $d_\beta(1)$. In such cases, $d_\beta(1) = D(1c_\alpha)$ for some irrational $\alpha \in (0, 1)$, where $D$ is a morphism defined by $D(0) = a$ ($a = \lfloor \beta \rfloor - 1$ if $\beta$ is maximally self-Sturmian) and $D(1) = b = \lfloor \beta \rfloor$. If $\beta \in (1, 2)$, then $\beta$ is self-Sturmian in the sense of the previous section, too.

To each $\beta > 1$, we associate a real number that is the diameter of $T_\beta$-orbit of 1.

*Definition.* $\mathrm{diam} : (1, \infty) \to [0, 1]$ is a function defined by the diameter of $T_\beta$-orbit of 1, that is,

$$\mathrm{diam}(\beta) := \mathrm{diam}\{T_\beta^n 1\}_{n \geq 0} = \sup\{|x - y| : x, y \in \{T_\beta^n 1\}_{n \geq 0}\}.$$

One can note that if $\beta \in \mathcal{C}_1$ or $\beta \in (1, \infty) \setminus \mathcal{C}_3$, then $\mathrm{diam}(\beta) = 1$ since both 0 and 1 lie in the closure of $\{T_\beta^n 1\}_{n \geq 0}$. Thus noteworthy is only the case $\beta \in \mathcal{C}_3 \setminus \mathcal{C}_1$. We get immediately

**Proposition 5.3.** *Suppose $\beta$ is self-Sturmian and* $\mathrm{alph}(d_\beta(1)) = \{a, b\}$ *with $0 \leq a < b = \lfloor \beta \rfloor$. Then*

- $\beta \in \mathcal{C}_3 \setminus \mathcal{C}_2$.
- $\mathrm{diam}(\beta) = \frac{b-a}{\beta}$.

*Moreover $\beta$ is maximally self-Sturmian if and only if $\beta \notin \mathcal{C}_2$ and $1-1/\beta \le T_\beta^n 1 \le 1$ for any $n \ge 0$.*

The above proposition implies that for a maximal self-Sturmian number $\beta$, its diameter is minimal in the following sense.

**Corollary 5.3.1.** *For any $\beta > 1$, either $\beta \in \mathcal{C}_2$ or $\mathrm{diam}(\beta) \ge 1/\beta$.*

Proposition 5.2 shows the set of self-Sturmian numbers is of Lebesgue measure zero. Then what about the size of $\overline{\{T_\beta^n 1\}}_{n\ge 0}$ for a fixed self-Sturmian number $\beta$? The next theorem answers this question.

**Theorem 5.1** ([7])**.** *If $\beta$ is self-Sturmian and $\mathrm{alph}(d_\beta(1)) = \{a, b\}$ with $0 \le a < b = \lfloor \beta \rfloor$, then $\overline{\{T_\beta^n 1\}}_{n\ge 0}$ is of Lebesgue measure zero.*

## 6. Transcendence of self-Sturmian numbers.

We know that $\beta$ is an algebraic integer for all $\beta \in \mathcal{C}_2$. Then are there transcendental numbers in $\mathcal{C}_3$, $\mathcal{C}_4$, and $\mathcal{C}_5$? This was questioned by Blanchard in his paper [5]. From Schmeling's results on the sizes of the classes, $\mathcal{C}_5$ is abundant in transcendental numbers. But a transcendental number reported in $\mathcal{C}_3$ is, to the knowledge of authors, only the Komornik-Loreti constant $\delta = 1.787231650\cdots$. See [11, 1, 2].

This section tells us that all self-Sturmian numbers are transcendental. That enriches $\mathcal{C}_3$ with transcendental numbers of continuum cardinality. In fact Sturmian words hitherto have given birth to transcendental numbers in other manners. Ferenczi and Mauduit showed real numbers whose expansions in some integer base are Sturmian are transcendental [9]. Moreover, they also generalized the transcendence results for Arnoux-Rauzy sequences (on 3 letters), and later it was extended to Arnoux-Rauzy sequences on any $k$ letters by Risley and Zamboni [15]. Recently it is known that if the sequence of partial quotients of the continued fraction expansion of a positive real number is Sturmian, then the number is transcendental [3]. We are now in a position to state the main result of this section.

**Theorem 6.1** ([7])**.** *Every self-Sturmian number is transcendental, that is, if $d_\beta(1)$ is Sturmian, then $\beta$ is transcendental.*

*Example 4.* We assume $d_\beta(1) = 31311313113113131131313113113\cdots$, where the sequence is obtained by substituting 3 for 1 and 1 for 0 in the sequence $1c_{\tau^{-2}}$ of Example 3. Then such $\beta \in (3, 4)$ exists and is transcendental. Furthermore, $\mathrm{diam}\{T_\beta^n 1\}_{n\ge 0} = 2/\beta$ and $\overline{\{T_\beta^n 1\}}_{n\ge 0}$ is of Lebesgue measure zero.

## References

[1] J.-P. Allouche and M. Cosnard. The Komornik-Loreti constant is transcendental. *Amer. Math. Monthly* **107** (2000) 448-449.

[2] J.-P. Allouche and M. Cosnard. Non-integer bases, iteration of continuous real maps, and an arithmetic self-similar set. *Acta Math. Hungar.* **91** (2001) 325-332.

[3] J.-P. Allouche, J.L. Davison, M. Queffélec and L.Q. Zamboni. Transcendence of Sturmian or morphic continued fractions. *J. Number Theory* **91** (2001) 39-66.

[4] A. Bertrand-Mathis. Développement en base $\theta$ et répartition modulo 1 de la suite $(x\theta^n)$. *Bull. Soc. Math. France* **114** (1986) 271-324.

[5] F. Blanchard. $\beta$-expansions and symbolic dynamics. *Theoret. Comput. Sci.* **65** (1989) 131-141.

[6] J.-P. Borel and F. Laubie. Quelques mots sur la droite projective rèelle. *J. Théor. Nombres Bordeaux* **5** (1993) 23-51.

[7] D.P. Chi and D.Y. Kwon. Sturmian words, $\beta$-shifts, and transcendence. *Theoret. Comput. Sci.* **321** (2004) 395-404.

[8] E.M. Coven and G.A. Hedlund. Sequences with minimal block growth. *Math. Systems Theory* **7** (1973) 138-153.

[9] S. Ferenczi and C. Mauduit. Transcendence of numbers with a low complexity expansion. *J. Number Theory* **67** (1997) 146-161.

[10] F. Hofbauer. $\beta$-shifts have unique maximal measure. *Monatsh. Math.* **85** (1978) 189-198.

[11] V. Komornik and P. Loreti. Unique developments in non-integer bases. *Amer. Math. Monthly* **105** (1998) 636-639.

[12] M. Lothaire. *Algebraic combinatorics on words.* Cambridge University Press, 2002.

[13] M. Morse and G.A. Hedlund. Symbolic dynamics II: Sturmian sequences. *Amer. J. Math.* **62** (1940) 1-42.

[14] W. Parry. On the $\beta$-expansion of real numbers. *Acta Math. Acad. Sci. Hungar.* **11** (1960) 401-416.

[15] R.N. Risley and L.Q. Zamboni. A generalization of Sturmian sequences: combinatorial structure and transcendence. *Acta Arith.* **95** (2000) 167-184.

[16] J. Schmeling. Symbolic dynamics for $\beta$-shift and self-normal numbers. *Ergodic Theory Dynam. Systems* **17** (1997) 675-694.

School of Mathematical Sciences, Seoul National University, Seoul 151-747, Korea.
*E-mail address*: dpchi@math.snu.ac.kr

School of Computational Sciences, Korea Institute for Advanced Study, Seoul 130-722, Korea.
*E-mail address*: doyong@kias.re.kr