Applications of standard Sturmian words to elementary number theory

Tom C. Brown

Citation data: T.C. Brown, *Applications of standard Sturmian words to elementary number theory.*, WORDS (Rouen, 1999). Theoret. Comput. Sci. **273** (2002), no. 1–2, 5–9.

Abstract

This note is a short survey, by no means intended to be complete, of some of the descriptions which have been given of standard Sturmian words, and of some of the applications of these descriptions to elementary number theory.. ('Elementary number theory' is interpreted fairly broadly.) The descriptions and applications below have appeared before, expect for Fact 4, Application 1, and the proof of Application 2.

Keywords: Sturmian words; Continued fractions; Zeckendorf representations

1 Introduction

In 1876, Smith [28] proved the following remarkable fact.

Let α be an irrational real number with $0 < \alpha < 1$. Define the infinite binary word f_α by $f_\alpha = f_\alpha(1)f_\alpha(2)\cdots f_\alpha(n)\cdots$ where $f_\alpha(n) = [(n+1)\alpha] - [n\alpha]$, $n \geq 1$, $[\cdot]$ denoting the greatest integer function. Assume that α has the simple continued fraction expansion $\alpha = [0, a_1, a_2, \ldots, a_n, \ldots]$, and inductively define finite words $g_0 = 0$, $g_1 = 0^{a_1-1}1$, $g_n = g_{n-1}^{a_n} g_{n-2}$, $n \geq 2$, and then set $g_\alpha = \lim_{n\to\infty} g_n$. Smith proved that $f_\alpha = g_\alpha$.

The definition of g_α just given is equivalent to the usual modern definition *standard Sturmian word with slope* α. See, for example, [10–15], or the chapter on Sturmian words in [4].

2 Descriptions of standard Sturmian words

The following notation is fixed throughout the statements of the four facts below. Let α be an irrational real number with $0 < \alpha < 1$, and with simple continued fraction expansion $\alpha = [0, a_1, a_2, \ldots, a_n, \ldots]$. Let f_α be the infinite binary word $f_\alpha = f_\alpha(1)f_\alpha(2)\cdots f_\alpha(n)\cdots$, where $f_\alpha(n) = [(n+1)\alpha] - [n\alpha]$, $n \geq 1$. Furthermore, for $n \geq 1$ let $p_n/q_n = [0, a_1, a_2, \ldots, a_n]$, the nth *convergent* of α, where p_n and q_n are relatively prime positive integers, and when $n \geq 1$ let X_n denote the initial segment (prefix) of f_α of length q_n, that is, $X_n = f_\alpha(1)f_\alpha(2)\cdots f_\alpha(q_n)$, $n \geq 1$. (Although for convenience we shall often define $X_0 = 0$, it is not necessarily true that X_0 is a prefix of f_α.)

We can now restate Smith's result in the following way, as Fact 1.

Fact 1. *(Fraenkel et al. [18], Rosenblatt [24], Shallit [27], Smith [28], Stolarsky [30]): For each $n \geq 2$, $X_n = X_{n-1}^{a_n} X_{n-2}$, where $X_0 = 0$ and $X_1 = 0^{a_1-1}1$. (Here $X_{n-1}^{a_n}$ denotes $X_{n-1}X_{n-1}\cdots X_{n-1}$, with a_n repetitions. If $a_1 = 1$, then $X_1 = 1$.) We emphasize that X_n is a prefix of f_α when $n \geq 1$, but $X_0 = 0$ need not be a prefix of f_α.*

As an illustration of this, if we take $\alpha = [0,2,1,1,\ldots]$, then $f_\alpha = 0100101001001\cdots$, the famous *Fibonacci word*. Here we have $f_\alpha = \lim_{n\to\infty} X_n$, where $X_0 = 0$, $X_1 = 01$, $X_n = X_{n-1}X_{n-2}$, and X_n has *length* $|X_n| = q_n$, where $q_0 = 1$, $q_1 = 2$, $q_n = q_{n-1} + q_{n-2}$, $n \geq 2$.

Fact 2. *(Bernoulli [3], Cristoffel [12], Markoff [22], Venkov [31]): For each $t \geq 1$, define the morphism k_t by $k_t(0) = 0^{t-1}1$, $k_t(1) = 0^t1$. For each $m \geq 1$, define $c_m = k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_m}(0)$, where $\circ$ denotes composition. Then $f_\alpha = c_1c_2c_3\cdots c_m\cdots$. In fact, for each $n \geq 1$, $f_\alpha = (c_1c_2c_3\cdots c_n)(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_n}(f_{\alpha_n})$, where α_n is defined by $\alpha = [0,a_1,a_2,\ldots,a_n+\alpha_n]$.*

Fact 3. *(Brown [7]): For each $t \geq 1$, define the morphism h_t by $h_t(0) = 0^{t-1}1$, $h_t(1) = 0^{t-1}10$. Then for each $n \geq 1$, $f_\alpha = h_{a_1} \circ h_{a_2} \circ \cdots \circ h_{a_n}(f_{\alpha_n})$, where α_n is define by $\alpha_n = [0,a_1,a_2,\ldots,a_n+\alpha_n]$.*

The following fact is apparently new, although it follows easily from Fact 2.

Fact 4. *Define $Z_0 = 0$, $Z_1 = 0^{a_1-1}1$, $Z_n = Z_{n-1}^{a_n-1}Z_{n-2}Z_{n-1}$, $n \geq 2$. Then $f_\alpha = Z_1Z_2Z_3\cdots Z_n\cdots$.*

To prove Fact 4, define c_m as in Fact 2. Then by induction on m, $c_m = Z_m$, $m \geq 1$. The induction step is:

$$\begin{aligned} c_m &= k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-1}}(k_{a_m}(0)) = k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-1}}(0^{a_m-1}1) \\ &= c_{m-1}^{a_m-1}(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-1}})(1) = c_{m-1}^{a_m-1}(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-2}})(0^{a_{m-1}}1) \\ &= c_{m-1}^{a_m-1}(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-2}})(0)(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-2}})(0^{a_{m-1}-1}1) \\ &= c_{m-1}^{a_m-1}(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-2}})(0)(k_{a_1} \circ k_{a_2} \circ \cdots \circ k_{a_{m-1}})(0) \\ &= c_{m-1}^{a_m-1}c_{m-2}c_{m-1} = Z_{m-1}^{a_m-1}Z_{m-2}Z_{m-1} = Z_m \end{aligned}$$

Hence, $f_\alpha = c_1c_2c_3\cdots c_m\cdots = Z_1Z_2Z_3\cdots Z_n\cdots$.

3 Applications

Application 1. *Let F_n be the nth Fibonacci number, defined as usual by $F_0 = 1$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$, $n \geq 2$. We show that the Fibonacci word $f = 0100101001001\cdots$ satisfies $f = \tilde{X}_1\tilde{X}_2\tilde{X}_3\cdots\tilde{X}_n\cdots$, where X_n is the initial segment of f of length F_n, and $\tilde{X}_n$ is the* reversal *of X_n. (Note that here the length of X_n is different from the length of X_n in the illustration following Fact 1.)*

To see this, take $\beta = [0,1,1,1,\cdots]$. Then, according to Fact 1 applied to f_β with $X_0 = 0$, $X_1 = 1$, $X_n = X_{n-1}X_{n-2}, n \geq 2$, we know that for each $n \geq 1$, X_n is an initial segment of f_β. The length of X_n is $|X_n| = F_n$, where $F_0 = 1$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$, $n \geq 2$. On the other hand, according to Fact 4, with $Z_0 = 0$, $Z_1 = 1$, $Z_n = Z_{n-2}Z_{n-1}$, $n \geq 2$, we know that $f_\beta = Z_1Z_2Z_3\cdots Z_n\cdots$. By induction, it is easy to see that $Z_n = \tilde{X}_n$, $n \geq 0$, where $\tilde{X}_n$ is the *reversal* of X_n. Thus, $f_\beta = \tilde{X}_1\tilde{X}_2\tilde{X}_3\cdots\tilde{X}_n\cdots$, where X_n is the initial segment of f_β of length F_n. Since (as is easily seen) f_β is the *complement* of the Fibonacci word (0's and 1's interchanged), it follows that the Fibonacci word $f = 0100101001001\cdots$ has the same property: $f = \tilde{X}_1\tilde{X}_2\tilde{X}_3\cdots\tilde{X}_n\cdots$, where X_n is the initial segment of f of length F_n.

This can be compared with [14], where it is shown that $f = 0100101001001\cdots = \tilde{X}_3\tilde{X}_5\tilde{X}_7\cdots\tilde{X}_{2n+1}\cdots$. (However, note that in [14] it is shown in addition that the division $f = 0100101001001\cdots = \tilde{X}_3\tilde{X}_5\tilde{X}_7\cdots\tilde{X}_{2n+1}\cdots$ has a certain *minimal* property with respect to the lexicographical order.)

Application 2. *(Carstens et al. [11]). Let α be an irrational real number with $\alpha > 0$, and with $1/\alpha = [a_0, a_1, a_2, \cdots]$. (Here we do not assume $\alpha < 1$.) Define two infinite words $C_\alpha = C_\alpha(1)C_\alpha(2)\cdots C_\alpha(n)\cdots$ and $B_\alpha = b_1b_2\cdots b_n\cdots$ as follows. For each $n \geq 1$, $C_\alpha(n) = |\alpha\mathbb{N} \cap (n, n+1)|$; that is, C_α is the number of positive integer multiples of α which lie between n and $n+1$. Each $b_k, k \geq 0$, is a word on the two symbols a_0 and a_0+1, defined inductively by $b_0 = a_0$, $b_1 = a_0^{a_1-1}(a_0+1)$, $b_k = b_{k-1}^{a_k-1}b_{k-2}b_{k-1}$, $k \geq 2$. Then $C_\alpha = B_\alpha$.*

The proof in [11] is rather long and complicated. We now give a short proof based on Fact 4.

Proof. Assume first that $\alpha > 1$, so that $a_0 = 0$, and both C_α and B_α are binary words on $0, 1$. Now C_α is the characteristic function of the set $\{[k\alpha] : k \geq 1\}$, since $C_\alpha(n) = 1 \Leftrightarrow |\alpha\mathbb{N} \cap (n, n+1)| = 1 \Leftrightarrow (\exists k, n < k\alpha < n+1) \Leftrightarrow (\exists k, n = [k\alpha]) \Leftrightarrow n \in \{[k\alpha] : k \geq 1\}$. On the other hand, it is easy to show (using $\alpha > 1$) that the characteristic function of $\{[k\alpha] : k \geq 1\}$ is in fact $f_{1/\alpha}$, where $f_{1/\alpha}(n) = [(n+1)1/\alpha] - [n1/\alpha]$, $n \geq 1$. Since $1/\alpha = [0, a_1, a_2, \ldots]$, and the b_k's are defined exactly as the Z_k's are defined in Fact 4 (applied to $1/\alpha$), Fact 4 now tells us that $C_\alpha = f_{1/\alpha} = Z_1Z_2\cdots = b_1b_2\cdots = B_\alpha$.

Next, assume that $\alpha < 1$. Let $\frac{1}{\alpha} = [a_0, a_1, a_2, \ldots]$, so that $\alpha = 1/a_0 + \alpha_1$, where $\alpha_1 = [0, a_1, a_2, \ldots] < 1$. It is easy to see that $C_\alpha(n) = a_0 + f_{\alpha_1}(n)$, $n \geq 1$ (where as usual $f_{\alpha_1}(n) = [(n+1)\alpha_1] - [n\alpha_1]$). From the definition of the word B_α, we see (using Fact 4 applied to α_1) that if the symbol a_0 is replaced by 0, and a_0+1 is replaced by 1, the word B_α is transformed into the word f_{α_1}. Since $C_\alpha(n) = a_0 + f_{\alpha_1}(n)$, $n \geq 1$, it follows that $C_\alpha = B_\alpha$. ☐

Application 3. *In [1, 5, 13] are three independent proofs of (a generalization of) the following result. A very simple proof using Fact 1 is in [2]. (See also [6].) Let $\tau = (-1+\sqrt{5})/2$, let $F_0 = F_1 = 1$, $F_n = F_{n-1} + F_{n-2}$, $n \geq 2$. Then*

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{f_\tau(k)}{2^k} &= \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^{[k/\tau]} = \sum_{k=1}^{\infty} \frac{[k\tau]}{2^k} \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{(2^{F_k}-1)(2^{F_{k-1}}-1)} \\ &= [0, 2^0, 2^1, 2^1, 2^2, 2^3, \ldots, 2^{F_n}, \ldots].\end{aligned}$$

Application 4. *(Brown [8]). Assume $0 < \alpha < 1$, with $\alpha = [0, a_1, a_2, \ldots, a_n, \ldots]$, and for $n \geq 1$, $p_n/q_n = [0, a_1, a_2, \ldots, a_n]$, where p_n and q_n are relatively prime positive integers. For $m \geq 2$, we find the* Zeckendorf representation of *$m-1$ (see, for example, [16]) by subtracting the largest possible q_i from $m-1$, and repeating, until finally we have $m-1 = \sum_{j=1}^{t} z_j q_{j-1}$. Then $[m\alpha] = \sum_{j=1}^{t} z_j p_{j-1}$.*

(The result in Application 4, in a somewhat more complex form, appears in [17]. The proof given in [8] uses in a simple way a generalization of Fact 1.)

Application 5. *(Brown and Shiue [9]). By induction on n, using Fact 1, one can show that for $n \geq 1$, $\sum_{k=1}^{q_n}[k\alpha] = \frac{1}{2}(p_n q_n - q_n + p_n + (-1)^n)$. (The proof given in [9], however, does not use Fact 1.) More generally, for any $m \geq 1$, if $m = \sum_{j=1}^{t} z_j q_{j-1}$ is the* Zeckendorf representation *of m (see Application 4) then*

$$\sum_{k=1}^{m}[k\alpha] = \frac{1}{2}\sum_{j=1}^{t} z_j(z_j p_{j-1} q_{j-1} - q_{j-1} + p_{j-1} + (-1)^{j-1}) + \sum_{1\leq i<j\leq t} z_i z_j p_{j-1} q_{j-1}.$$

This leads to a formula for the function $C_\alpha(m) = \sum_{k=1}^{m}(\{k\alpha\} - \frac{1}{2})$ ($\{\cdot\}$ denotes the fractional part), which was used in [9] to give some improvements on results of Hardy and Littlewood [19, 20], and Ostrowski [23], and a simplified proof of a theorem of Sós [29]. (See also [25, 26].) One of the main results of [19, 20, 23] is that if $\alpha = [0, a_1, a_2, \ldots]$ and $a_i \leq A$ for all i, then, for some positive constant c_A, $C_\alpha(m) > c_A \log m$ holds for infinitely many m, and $C_\alpha(m) < -c_A \log m$ holds for infinitely many m. We show in [9] (among other things) that the same conclusions hold if the a_i are bounded infinitely often *on the average*, that is, if $1/t\sum_{j=1}^{t} a_j \leq A$ for infinitely many t.

One can also use Application 5 to prove an old formula of Lerch [21]: $\sum_{k=1}^{m}[k\alpha] + \sum_{k=1}^{[m\alpha]}[k1/\alpha] = m[m\alpha]$. This proof seems a little complicated, and it was asked at the conference whether there is not a more direct proof. Within a day or so, Jano Manuch showed the author a very short completely elementary proof.

References

[1] W.W. Adams and J.L. Davison, *A remarkable class of continued fractions*, Proc. Amer. Math. Soc. **65** (1977), 194–198.

[2] P.G. Anderson, T.C. Brown, and P.J.-S. Shiue, *A simple proof of a remarkable continued fraction identity*, Proc. Amer. Math. Soc. **123** (1995), 2005–2009.

[3] J. Bernoulli III, *Sur une nouvelle espece de calcul*, Recueil pour les Astronomes **1** (Berlin, 1772), 255–284.

[4] J. Berstel and P. Séébold, *Sturmian words*, Algebraic Combinatorics on Words (M. Lothaire, ed.), Encyclopedia of Mathematics and its Applications, vol. 90, Cambridge University Press, Cambridge, 2002, pp. 45–110.

[5] P.E. Böhmer, *Über die transzendenz gewisser dyadischer brüche*, Math. Ann. **96** (1926), 367–377, erratum 96 (1926) 735.

[6] J.M. Borwein and P.B. Borwein, *On the generating function of the integer part:* $[n\alpha + \gamma]$, J. Number Theory **43** (1993), 293–318.

[7] T.C. Brown, *A characterization of the quadratic irrationals*, Canad. Math. Bull. **34** (1991), 36–41.

[8] ———, *Descriptions of the characteristic sequence of an irrational*, Canad. Math. Bull. **36** (1993), 15–21.

[9] T.C. Brown and P.J.-S. Shiue, *Sums of fractional parts of integer multiples of an irrational*, J. Number Theory **50** (1995), 181–192.

[10] A. Carpi and A. de Luca, *Special factors, periodicity, and an application to Sturmian words*, Acta Inform. **36** (2000), 983–1006, no. 12.

[11] H. Carstens, W. Deuber, W. Thumser, and E. Koppenrade, *Geometrical bijections in discrete lattices*, Combin. Probab. Comput. **8** (1999), 109–129.

[12] E.B. Christoffel, *Observatio arithmetica*, Ann. Mat. **6** (1875), 145–152.

[13] L.V. Danilov, *Some classes of transcendental numbers*, Math. Notes Acad. Sci. USSR **12** (1972), 524–527.

[14] A. de Luca, *A division property of the fibonacci word*, Inform. Process. Lett. **54** (1995), 307–312.

[15] ———, *Sturmian words: structure, combinatorics, and their arithmetics*, Theoret. Comput. Sci. **183** (1997), 45–82.

[16] A.S. Fraenkel, *Systems of numeration*, Amer. Math. Monthly **92** (1985), 105–114.

[17] A.S. Fraenkel, J. Levitt, and M. Shimshoni, *Characterization of the set of values of $f(n)=[n\alpha]$, $n=1,2,\cdots$*, Discrete Math. **2** (1972), 335–345.

[18] A.S. Fraenkel, A. Mushkin, and U. Tassa, *Determination of $[n\theta]$ by its sequence of differences*, Canad. Math. Bull. **21** (1978), 441–446.

[19] G.H. Hardy and J.E. Littlewood, *Some problems of diophantine approximation: the lattice-points of a right-angled triangle*, Proc. London Math. Soc. **20** (1922), 15–36.

[20] ———, *Some problems of diophantine approximation: the lattice-points of a right-angled triangle (second memoir)*, Abh. Math. Sem. Univ. Hamburg **1** (1922), 212–249.

[21] M. Lerch, *Question 1547*, l'Intermédiaire des Mathématiciens **11** (1904), 145–146.

[22] A.A. Markoff, *Sur une question de Jean Bernoulli*, Math. Ann. **19** (1882), 27–36.

[23] A. Ostrowski, *Bemerkungen zur theorie der Diophantischen approximationen*, Abh. Math. Sem. Univ. Hamburg **1** (1922), 77–98.

[24] J. Rosenblatt, *The sequence of greatest integers of an arithmetic progression*, J. London Math. Soc. **17** (1978), 213–218.

[25] J. Schoissengeier, *On the discrepency of $(n\alpha)$*, Acta Arithmetica **64** (1984), 241–279.

[26] ———, *Abschätzungen für $\sum_{n\leq N} b_1(n\alpha)$*, Monatshefte Math. **102** (1986), 59–77.

[27] Jeffrey Shallit, *Characteristic words as fixed points of homomorphisms*, Tech. Report CS-91-72, Univ. of Waterloo, Dep. of Computer Science, 1991.

[28] H.J.S. Smith, *Note on continued fractions*, Messenger Math. (6) (1876), 1–14.

[29] Vera T. Sós, *On the theory of diophantine approximations i (on a problem of a. ostrowski)*, Acta Math. **8** (1957), 461–472.

[30] K.B. Stolarsky, *Beatty sequences, continued fractions, and certain shift operators*, Canad. Math. Bull. **19** (1976), 473–482.

[31] B.A. Venkov, *Elementary number theory*, pp. 65–68, Wolters-Noordhoff, Groningen, 1970, Translated and edited by H. Alderson.