Direct definition of a ternary infinite square-free sequence

Tetsuo Kurosaki

Department of Physics, Graduate School of Science, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Bank of Japan, Hongoku-cho, Nihonbashi, Chuo-ku, Tokyo 103-8660, Japan

Abstract

We propose a new ternary infinite (even full-infinite) square-free sequence. The sequence is defined both by an iterative method and by a direct definition. Both definitions are analogous to those of the Thue-Morse sequence. The direct definition is given by a deterministic finite automaton with output. In short, the sequence is automatic.

Key words:
Thue-Morse sequence, Square-free sequence, Automatic sequence, Combinatorial problems, Symbolic dynamics

1 Introduction

First of all, we introduce several notations to be used in this paper. Let Σ be a non-empty and finite set called *alphabet* and Σ^* be the free monoid generated by Σ. A sequence consists of elements of the free monoid, i-th element of which is usually assigned a symbol with a subscript i. Specifically, a sequence $\mathbf{a}$ over Σ, or $\mathbf{a} \in \Sigma^*$, with the length $n + 1$ is expressed as $\mathbf{a} = \{a_i\}_{i=0}^n \equiv a_0 a_1 \cdots a_n$ where $\forall a_i \in \Sigma$. We usually make the numbering of elements of a sequence start with 0, whereas this numbering will be altered later. A sequence can be either empty or infinite. Concatenation of two finite sequences is denoted by juxtaposition. For example, $\{a_i\}_{i=0}^n \{a_{n+i+1}\}_{i=0}^m = \{a_i\}_{i=0}^{n+m+1}$. We say that a sequence $\mathbf{a}$ contains another sequence $\mathbf{u}$, called a factor of $\mathbf{a}$, if there are $\mathbf{r}$ and $\mathbf{r}^*$ such that $\mathbf{a} = \mathbf{r u r}^*$. The concept of the factor is defined to an infinite sequence as well.

Thue introduced a simple and mysterious binary infinite sequence $\mathbf{t} = \{t_i\}_{i=0}^{\infty} \in \{0,1\}^*$ and discussed the combinatorial properties of it [10, 11]. It starts such that

$$\mathbf{t} = 0110100110010110\cdots. \tag{1}$$

Morse rediscovered the sequence $\mathbf{t}$ in the context of differential geometry [8, 9]. After the two pioneers, the sequence $\mathbf{t}$ is now called the *Thue-Morse sequence*, which has a long history of research [1, 2, 5, 7].

There exist several ways of the mathematical definition of the Thue-Morse sequence $\mathbf{t}$. A popular one is based on a recursive construction. Define an operator $\psi : \{0,1\}^* \to \{0,1\}^*$ which doubles the length of the sequence such that $\psi(\mathbf{a}) = \mathbf{a}\bar{\mathbf{a}}$, where $\bar{\mathbf{a}}$ is the sequence obtained from $\mathbf{a}$ by exchanging 0 and 1. We find that $\psi(0) = 01$, $\psi^2(0) = 0110$ and $\psi^{\infty}(0) = \lim_{n \to \infty} \psi^n(0)$ converges toward $\mathbf{t}$ for the ordinary topology, which is a fixed point of ψ. Another definition is more direct. Let $s_2(i)$ be the sum of digits in the binary representation of an integer i and then $s_2(i)$ corresponds to t_i under modulo 2. From the latter definition, we can construct a deterministic finite automaton with output (DFAO) which takes as input the binary representation of i and yields t_i. This means that the Thue-Morse sequence is 2-automatic [1, 4].

The Thue-Morse sequence is known to appear in various distinct fields such as combinatorics, number theory, group theory, real analysis, information science, solid state physics and so on (see, for instance, a review [2]). In terms of combinatorial aspects, a remarkable fact is that a *square-free* sequence can be derived by utilizing the Thue-Morse sequence as the auxiliary sequence [10, 11]. A sequence $\mathbf{u}$ over Σ is said to be square-free if $\mathbf{u}$ contains no square, i.e., no two consecutive repeating blocks $\mathbf{ww}$ with $\mathbf{w}$ a non-empty and finite sequence of Σ^*. For $i \geq 0$, let v_i be the number of 1's between i-th and $(i+1)$-st appearance of 0 in the sequence $\mathbf{t}$. Then, $\mathbf{v} = \{v_i\}_{i=0}^{\infty} = 21020121012\cdots$ is a ternary infinite square-free sequence over $\Sigma = \{0, 1, 2\}$. The concept of square-free is significant in that it leads to a branch of combinatorics, called *combinatorics on words* [7]. Many sequences other than $\mathbf{v}$ are known as ternary infinite square-free sequences. In particular, the sequences of Arshon [3], Leech [6], Zech [12] and so on are not only square-free but also automatic, due to the fact that they can be defined by a uniform tag sequence [4].

In this paper, we present a new method to define a ternary infinite square-free sequence. First, the sequence is defined by an iterative method and then redefined by a direct definition. Since the direct definition is brought by an automaton operation, we find that the sequence is automatic. Our definitions are quite straightforward in contrast with those of the Thue-Morse sequence.

2 Construction of a ternary infinite square-free sequence

Ternary sequences considered in this section are those over $\{1, 2, 3\}$.

Let σ and ρ be morphisms or permutations on $\{1, 2, 3\}^*$, namely, for $\mathbf{a} \in \{1, 2, 3\}^*$. They are defined as follows, respectively:

- $\sigma(\mathbf{a})$ is the sequence obtained from $\mathbf{a}$ by exchanging 1 and 2.
- $\rho(\mathbf{a})$ is the sequence obtained from $\mathbf{a}$ by exchanging 2 and 3.

In addition, define an operator $\varphi : \{1, 2, 3\}^* \rightarrow \{1, 2, 3\}^*$ which triples the length of the sequence such that $\varphi(\mathbf{a}) = \sigma(\mathbf{a}) \, \mathbf{a} \, \rho(\mathbf{a})$.

Theorem 1 *The sequence $\varphi^n(2)$ is square-free for arbitrary integer n.*

It is easy to write explicitly $\varphi^n(2)$ for $n = 1, 2, 3$. They are given by

$$\varphi^1(2) = 123,$$
$$\varphi^2(2) = 213123132,$$
$$\varphi^3(2) = 123213231213123132312132123,$$

and we can see that they are square-free.

The length of $\varphi^n(2)$ is 3^n. We shall prove two lemmas about $\varphi^n(2) \equiv \{a_i^n\}_{i=0}^{3^n-1}$. These lemmas are very simple, but crucial to prove Theorem 1. The main idea is to decompose $\varphi^n(2)$ into 3^{n-1} consecutive ternarys $a_{3k}^n a_{3k+1}^n a_{3k+2}^n$.

Lemma 2 *Each ternary $a_{3k}^n a_{3k+1}^n a_{3k+2}^n$ $(k = 0, 1, \cdots, 3^{n-1} - 1)$ is a certain permutation of 123.*

PROOF. Obviously, the iterative operations for the sequence 2 by σ and ρ, in short, φ, preserve the property stated by Lemma 2. $\square$

Remark that Lemma 2 guarantees that, if two elements of the ternary are provided, one can guess the remaining one element of it.

Define an operator f which extracts a_{3k+1}^n from each ternary $a_{3k}^n a_{3k+1}^n a_{3k+2}^n$ for $k = 0, 1, \cdots, 3^{n-1} - 1$, i.e., $f(\varphi^n(2)) = \{a_{3i+1}^n\}_{i=0}^{3^{n-1}-1}$. The idea of extracting the central element of the ternary comes from the analogy of the Cantor set, well known as the fractal set.

3

Lemma 3 *The sequence* $f(\varphi^n(2)) = \{a_{3i+1}^n\}_{i=0}^{3^{n-1}-1}$ *is equivalent to* $\varphi^{n-1}(2) = \{a_i^{n-1}\}_{i=0}^{3^{n-1}-1}$.

PROOF. Obviously, $f(\varphi(2)) = f(123) = 2 = \varphi^0(2)$. Assume that, for $n \geq 1$, $f(\varphi^n(2)) = \varphi^{n-1}(2)$ is also true. Note that the operators φ and f are commutable, and then $f(\varphi^{n+1}(2)) = \varphi(f(\varphi^n(2))) = \varphi^n(2)$. Therefore, $f(\varphi^n(2)) = \varphi^{n-1}(2)$ is proved for arbitrary n, by a mathematical induction. $\square$

By using Lemmas 2 and 3, we can prove Theorem 1.

PROOF of Theorem 1. It is sufficient to prove that $\varphi^n(2) = \{a_i^n\}_{i=0}^{3^n-1}$ does not contain a square with any length 2ℓ, i.e., $\{a_i^n\}_{i=p}^{p+2\ell-1}$ such that $a_k^n = a_{k+\ell}^n$ $(k = p, p+1, \cdots, p+\ell-1)$, for any p and ℓ. To complete the proof, we divide the problem into several cases, according to the value of ℓ, whereas we employ a mathematical induction and a reduction to absurdity. It is clear that $\varphi^1(2) = 123$ is square-free. Hereafter, we consider $\varphi^{n+1}(2)$ for individual cases divided as to ℓ, provided that $\varphi^n(2)$ is square-free for $n \geq 1$.

Case 1: $\ell = 1, 2$

It is obvious that both $\sigma(\varphi^n(2))$ and $\rho(\varphi^n(2))$ are also square-free. Therefore, if $\varphi^{n+1}(2)$ contains a square, it should straddle over the joint between $\sigma(\varphi^n(2))$ and $\varphi^n(2)$, or between $\varphi^n(2)$ and $\rho(\varphi^n(2))$. However, noting that $\varphi^n(2)$ starts and ends such that $123 \cdots 123$ for odd n, and $213 \cdots 132$ for even n, $\varphi^{n+1}(2)$ is in the form of $213 \cdots 213123 \cdots 123132 \cdots 132$ for odd n, and $123 \cdots 231213 \cdots 132312 \cdots 123$ for even n, respectively. From the above observations, it is shown that $\varphi^{n+1}(2)$ does not contain any square with the length either 2 or 4.

Case 2: $\ell \equiv 0 \pmod 3$

In this case, if $\varphi^{n+1}(2)$ contains a square with the length 2ℓ, $\varphi^n(2)$ contains a square with the length $2\ell/3$, according to Lemma 3. This contradicts the assumption that $\varphi^n(2)$ is square-free.

Case 3: $\ell \geq 4$ and $\ell \not\equiv 0 \pmod 3$

In this case, the proof is rather complicated than those in the other cases. It is necessary to further divide the cases, according to the values of ℓ and p. Here, we omit a superscript $n+1$ from a_i^{n+1} for simplicity: $\varphi^{n+1}(2) = \{a_i\}_{i=0}^{3^{n+1}-1}$.

4

For example, we consider the case that $\ell = 3m + 1$ and $p = 3q$ where m and q are certain integers within an appropriate range. Assume that $\varphi^{n+1}(2)$ contains a square such that $a_k = a_{k+3m+1}$ for $k = 3q, 3q+1, \cdots, 3(q+m)$. Note a couple of ternarys $a_{3q}a_{3q+1}a_{3q+2}$ and $a_{3(q+m)}a_{3(q+m)+1}a_{3(q+m)+2}$. One can deduce that $a_{3q+2} = a_{3(q+m)}$ because $a_{3q} = a_{3(q+m)+1}$ and $a_{3q+1} = a_{3(q+m)+2}$ from the assumption, and Lemma 2 guarantees that a_{3q+2} $(a_{3(q+m)})$ is different from a_{3q} and a_{3q+1} $(a_{3(q+m)+1}$ and $a_{3(q+m)+2})$. The similar deductions for each couple of ternarys lead to

$$a_{3q+2} = a_{3q+5} = \cdots = a_{3(q+m)-1} = a_{3(q+m)} = a_{3(q+m+1)} = \cdots = a_{3(q+2m)},$$

and hence $\varphi^{n+1}(2)$ contains a square with the length 2: $a_{3(q+m)-1} = a_{3(q+m)}$.

For another example, we consider the case that $\ell = 3m + 2$ and $p = 3q + 1$, and assume that $\varphi^{n+1}(2)$ contains a square such that $a_k = a_{k+3m+2}$ for $k = 3q+1, 3q+2, \cdots, 3(q+m)+2$. $a_{3q} = a_{3(q+m+1)+2}$ is deduced noting a couple of ternarys $a_{3q}a_{3q+1}a_{3q+2}$ and $a_{3(q+m+1)}a_{3(q+m+1)+1}a_{3(q+m+1)+2}$. It follows that

$$a_{3q} = a_{3q+3} = \cdots = a_{3(q+m)} = a_{3(q+m+1)+2} = a_{3(q+m+2)+2} = \cdots = a_{3(q+2m)+2}.$$

Note a couple of consecutive ternarys $a_{3(q+m)}a_{3(q+m)+1}a_{3(q+m)+2}$ and $a_{3(q+m+1)}a_{3(q+m+1)+1}a_{3(q+m+1)+2}$ in turn. Since $\varphi^n(2)$ is square-free, Lemma 3 allows one to deduce $a_{3(q+m)+1} \neq a_{3(q+m+1)+1}$. Combine this fact and $a_{3(q+m)} = a_{3(q+m+1)+2}$, and then $a_{3(q+m)+1} = a_{3(q+m+1)}, a_{3(q+m)+2} = a_{3(q+m+1)+1}$ from Lemma 2. Therefore, $\varphi^{n+1}(2)$ contains a square with the length 4.

Under modulo 3, ℓ and p are divided into two and three cases, respectively. There exist six cases in total. We can prove that, if $\varphi^{n+1}(2)$ contains a square, it contains another square with the length 2 or 4 at the joint of the square, for other four cases as well as for the above two cases. To summarize, Case 3 can be reduced to Case 1, by utilizing Lemmas 2 and 3.

From the above considerations for three cases, we can conclude that $\varphi^{n+1}(2)$ does not contain a square with any length 2ℓ, and thus Theorem 1 has been established by a mathematical induction. □

Note that $\varphi^\infty(2)$ does not yet converge toward a definite sequence. We should alter the numbering of elements of $\varphi^n(2)$, in order to derive an infinite sequence. Taking the symmetry of the ternary into account, it is required to employ a full-infinite sequence, instead of a semi-infinite sequence. When we set the numbering of elements of $\varphi^n(2)$ in such a way that $\varphi^n(2) \equiv \{a_i\}_{i=-(3^n-1)/2}^{(3^n-1)/2}$, the sequence $\varphi^n(2)$ is included in the middle of $\varphi^{n+1}(2)$, and then the (full-) infinite square-free sequence $\varphi^\infty(2) \equiv \{a_i\}_{i=-\infty}^{\infty}$ can be defined. $\varphi^\infty(2)$ is

quite similar to the Arshon sequence [3], but essentially different, because the Arshon sequence is semi-infinite.

3   Direct definition of a ternary infinite square-free sequence

In this section, we redefine the square-free sequence derived in section 2, by a direct definition. The key to a direct definition is to represent an integer i not by an ordinary ternary representation but by a *balanced* ternary representation. The latter uses the digits $-1, 0, +1$ whereas the former uses the digits $0, 1, 2$.

Let the balanced ternary representation of an integer i $(-\infty < i < \infty)$ be

$$i = \sum_{n=0}^{\nu} u_n 3^n \equiv [u_\nu \cdots u_1 u_0]_3 \tag{2}$$

where $u_n \in \Sigma = \{-1, 0, +1\}$ and $u_\nu \neq 0$. Note that this representation of i is uniquely determined.

Here, permutations $\pi_{-1}, \pi_0, \pi_{+1}$ over $\{-1, 0, +1\}$ are defined in the following manner, respectively:

- π_{-1} exchanges -1 and 0.
- π_0 is the identical permutation.
- π_{+1} exchanges 0 and $+1$.

In terms of $\pi_{-1}, \pi_0, \pi_{+1}$, we define the infinite sequence $\{b_i\}_{i=-\infty}^{\infty} \in \{-1, 0, +1\}^*$ such that

$$b_i = \pi_{u_\nu} \circ \cdots \circ \pi_{u_1} \circ \pi_{u_0}(0) \tag{3}$$

with $i = [u_\nu \cdots u_1 u_0]_3$. For example, $b_8 = \pi_{+1}(\pi_0(\pi_{-1}(0))) = -1$ and $b_{-17} = \pi_{-1}(\pi_{+1}(\pi_0(\pi_{+1}(0)))) = -1$ because $8 = 3^2 - 1$ and $-17 = -3^3 + 3^2 + 1$, respectively.

Finally, we redefine σ and ρ on $\{-1, 0, +1\}^*$. It follows that $\varphi^n(0)$ denotes the sequence obtained from $\varphi^n(2) = \{a_i\}_{i=-(3^n-1)/2}^{(3^n-1)/2}$ by replacing 1,2 and 3 with -1, 0 and $+1$, respectively.

Based on the above notations, a direct definition of a ternary infinite square-free sequence is derived:

6

Theorem 4 $\{b_i\}_{i=-(3^n-1)/2}^{(3^n-1)/2}$ is equivalent to $\varphi^n(0)$ for arbitrary integer n. Therefore, the ternary infinite sequence $\{b_i\}_{i=-\infty}^{\infty}$ is square-free according to Theorem 1.

PROOF. The proof proceeds via a mathematical induction. For $n = 1$, it is obvious that $b_{-1}b_0b_{+1} = \varphi(0)$. Assume that $\{b_i\}_{i=-(3^n-1)/2}^{(3^n-1)/2} = \varphi^n(0)$. For $(3^n - 1)/2 < i \le (3^{n+1}-1)/2$, $i = 3^n + i'$ where $|i'| \le (3^n-1)/2 = 3^{n-1}+3^{n-2}+\cdots+1$, namely, the highest digit of i in the balanced ternary representation is 1. Then, we obtain $\{b_i\}_{i=(3^n-1)/2+1}^{(3^{n+1}-1)/2} = \{\pi_{+1}(b_{i'})\}_{i'=-(3^n-1)/2}^{(3^n-1)/2} = \rho(\varphi^n(0))$. Similarly, we obtain $\{b_i\}_{i=-(3^{n+1}-1)/2}^{-(3^n-1)/2-1} = \sigma(\varphi^n(0))$. It follows that

$$\{b_i\}_{i=-(3^{n+1}-1)/2}^{(3^{n+1}-1)/2} = \sigma(\varphi^n(0))\ \varphi^n(0)\ \rho(\varphi^n(0)) = \varphi^{n+1}(0). \tag{4}$$

Therefore, the proof of Theorem 4 has been completed. □

The formula (3) can be regarded as the definition of a DFAO which takes as input the balanced ternary representation of an integer i, read in order from low to high digits. Indeed, the three function π_a denote the state transition function among the three state q_a/a ($a = -1, 0, +1$), where, by convention, the state labeled q_a/a indicates that the output associated with the state q_a is a and the initial state is $q_0/0$. The state transition diagram for this DFAO is drawn as a directed graph in Figure 1. Therefore, the sequence $\{b_i\}_{i=-\infty}^{\infty}$ is automatic as well as the Arshon sequence.
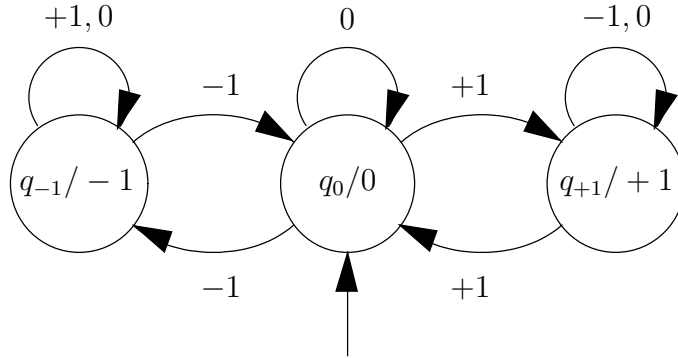

Fig. 1. DFAO generating the sequence $\{b_i\}_{i=-\infty}^{\infty}$

The two equivalent ways of the definition of the ternary infinite square-free sequence, $\varphi^{\infty}(0)$ and $\{b_i\}_{i=-\infty}^{\infty}$, are almost parallel with those of the Thue-Morse sequence. φ is to $\varphi^{\infty}(0)$ what ψ is to the Thue-Morse sequence $\psi^{\infty}(0)$. Furthermore, the balanced ternary representation is to the function b_i (3) what the binary representation is to the function $s_2(i)$. In this sense, our sequence is said to be the counterpart of the Thue-Morse sequence, i.e., a natural extension from binary to ternary.

4 Conclusion and discussion

We have proposed a novel construction of the ternary infinite square-free sequence, generated by the automaton operation with the balanced ternary representation of an integer. The sequence presented in this paper may be favorable for various applications, because of its easiness to generate. The nature of *square-free* is intriguing and useful itself. We hope that the properties of our sequence are further clarified and utilized in various fields such as information science, random system on physics and quasiperiodic tiling.

Acknowledgement

References

[1] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*, Cambridge University Press, Cambridge (2003).

[2] J.-P. Allouche and J. Shallit, The ubiquitous Prouhet-Thue-Morse sequence, in Sequences and Their applications, Proc. SETA'98 (Ed. C. Ding, T. Helleseth and H. Niederreiter), Springer-Verlag, New York (1999) pp. 1-16. (also available at http://www.cs.uwaterloo.ca/~shallit/Papers/ubiq.ps)

[3] S. Arshon, Démonstration de l'existence des suites asymétriques infinies, Mat. Sb. 44 (1937) 769-777.

[4] A. Cobham, Uniform tag sequences, Math. Systems Theory 6 (1972) 164-192.

[5] G. A. Hedlund, Remarks on the work of Axel Thue on sequences, Nordisk Mat. Tidskr. 15 (1967) 147-150.

[6] J. Leech, A problem on strings of beads, Math. Gazette 41 (1957) 277-278.

[7] M. Lothaire, *Combinatorics on Words*, Encyclopedia of Mathematics 17, Addison-Wesley, Reading, MA (1983).

[8] M. Morse, Recurrent geodesics on a surface of negative curvature, Trans. Amer. Math. Soc. 22 (1921) 84-100.

[9] M. Morse and G. A. Hedlund, Symbolic dynamics, Amer. J. Math. 60 (1938) 815-866.

[10] A. Thue, Über unendliche Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat.

Nat. Kl. 7 (1906) 1-22. (reprinted in: Selected Mathematical Papers of Axel Thue (Ed. T. Nagell), Universitetsforlaget, Oslo (1977) pp. 139-158)

[11] A. Thue, Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen, Norske Vid. Selsk. Skr. I. Mat. Nat. Kl. 1 (1912) 1-67. (reprinted in: Selected Mathematical Papers of Axel Thue (Ed. T. Nagell), Universitetsforlaget, Oslo (1977) pp. 413-478)

[12] T. Zech, Wiederholungsfreie Folgen, Z. Angew. Math. Mech. 38 (1958) 206-209.