Discrete Mathematics 339 (2016) 1459–1470

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

# On formal inverse of the Prouhet–Thue–Morse sequence

Maciej Gawron *, Maciej Ulas

*Jagiellonian University, Faculty of Mathematics and Computer Science, Institute of Mathematics, Łojasiewicza 6, 30 - 348 Kraków, Poland*

ARTICLE INFO

*Article history:*
Received 18 June 2015
Received in revised form 9 December 2015
Accepted 14 December 2015
Available online 11 January 2016

*Keywords:*
Prouhet–Thue–Morse sequence
Automatic sequences
Regular sequences

ABSTRACT

Let $p$ be a prime number and consider a $p$-automatic sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ and its generating function $U(X) = \sum_{n=0}^{\infty} u_nX^n \in \mathbb{F}_p[[X]]$. Moreover, let us suppose that $u_0 = 0$ and $u_1 \neq 0$ and consider the formal power series $V \in \mathbb{F}_p[[X]]$ which is a compositional inverse of $U(X)$, i.e., $U(V(X)) = V(U(X)) = X$. In this note we initiate the study of arithmetic properties of the sequence of coefficients of the power series $V(X)$. We are mainly interested in the case when $u_n = t_n$, where $t_n = s_2(n) \pmod 2$ and $\mathbf{t} = (t_n)_{n\in\mathbb{N}}$ is the Prouhet–Thue–Morse sequence defined on the two letter alphabet $\{0, 1\}$. More precisely, we study the sequence $\mathbf{c} = (c_n)_{n\in\mathbb{N}}$ which is the sequence of coefficients of the compositional inverse of the generating function of the sequence $\mathbf{t}$. This sequence is clearly 2-automatic. We describe the sequence $\mathbf{a}$ characterizing solutions of the equation $c_n = 1$. In particular, we prove that the sequence $\mathbf{a}$ is 2-regular. We also prove that an increasing sequence characterizing solutions of the equation $c_n = 0$ is not $k$-regular for any $k$. Moreover, we present a result concerning some density properties of a sequence related to $\mathbf{a}$.


## 1. Introduction

Let $k \in \mathbb{N}_{\geq 2}$ and consider a $k$-automatic sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$. Let us recall that the sequence $\mathbf{u}$ is $k$-automatic if its $n$th term is generated by a finite automaton with $n$ in base $k$ as the input. One can prove that this property is equivalent to the fact that the family of sequences (called $k$-kernel of $\mathbf{u}$)

$$\mathcal{K}(\mathbf{u}) := \{(u_{k^a n+b})_{n\in\mathbb{N}} : a \in \mathbb{N}, 0 \leq b \leq k^a - 1\}$$

is finite. The simplest $k$-automatic sequences are periodic sequences.

A famous 2-automatic sequence which is not periodic is the Prouhet–Thue–Morse sequence $\mathbf{t} = (t_n)_{n\in\mathbb{N}}$ (PTM sequence for short). In order to define the sequence $\mathbf{t}$ (on alphabet $\{0, 1\}$) let $n \in \mathbb{N}$ be written in base 2, i.e., $n = \sum_{i=0}^{k} \epsilon_i 2^i$, where $\epsilon_i \in \{0, 1\}$ for $i = 0, 1, \ldots, k$. Then we define the sum of digits function $s_2 : \mathbb{N} \to \mathbb{N}$ as $s_2(n) = \sum_{i=0}^{k} \epsilon_i$. This function satisfies the obvious recurrence relations:

$$s_2(0) = 0, \qquad s_2(2n) = s_2(n), \qquad s_2(2n+1) = s_2(n) + 1$$

for $n \geq 0$. The sum of digits function allows us to define the PTM sequence $\mathbf{t} = (t_n)_{n\in\mathbb{N}}$, where $t_n = s_2(n) \pmod 2$. We thus have

$$t_0 = 0, \qquad t_{2n} = t_n, \qquad t_{2n+1} = 1 - t_n$$

* Corresponding author.
*E-mail addresses:* maciej.gawron@uj.edu.pl (M. Gawron), maciej.ulas@uj.edu.pl (M. Ulas).

http://dx.doi.org/10.1016/j.disc.2015.12.016
0012-365X/

for $n \geq 0$. In particular, from the above relations we immediately deduce that the PTM sequence is indeed 2-automatic. This is clear due to the fact that its kernel contains exactly two sequences, i.e., $\mathcal{K}(\mathbf{t}) = \{\mathbf{t}, 1 - \mathbf{t}\}$.

The PTM sequence has many interesting properties and applications in combinatorics, algebra, number theory, topology and even physics (see for example [3] and [4]).

Now, if $k = p$ is a prime number then Christol theorem says that the sequence $\mathbf{u}$ (with terms in a finite field $\mathbb{F}_p$) is $p$-automatic if and only if the formal power series $U(X) = \sum_{n=0}^{\infty} u_n X^n$ is algebraic over $\mathbb{F}_p(X)$. In this context one can ask the following general

**Problem 1.1.** Let $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ be a $p$-automatic sequence and let $U(X) = \sum_{n=0}^{\infty} u_n X^n \in \mathbb{F}_p[[X]]$ be a formal power series related to the sequence $\mathbf{u}$. Let us suppose that there exists a formal power series $V(X) = \sum_{n=0}^{\infty} v_n X^n \in \mathbb{F}_p[[X]]$ which is compositional inverse to the series $U$, i.e., $U(V(X)) = X$. What can be said about properties of the sequence $\mathbf{v} = (v_n)_{n\in\mathbb{N}}$?

It is well known that the formal power series $U(X) = \sum_{n=0}^{\infty} u_n X^n$ with coefficients in a commutative ring $R$ is invertible (in the sense of composition) if and only if $u_0 = 0$ and $u_1$ is an invertible element of $R$. Moreover, if $U(V(X)) = X$ then $V(U(V(X))) = V(X)$ and thus $H(V(X)) = 0$, where $H(X) = V(U(X)) - X$. Because $U$ is non-constant so is $V$ and thus $H \equiv 0$ which is equivalent with the equality $V(U(X)) = X$. In particular, the above problem does not make sense for all $p$-automatic sequences. However, we observe that if the problem above is correctly stated then the sequence $\mathbf{v}$ is $p$-automatic which is an immediate consequence of Christol theorem [4, Theorem 12.2.5]. Indeed, if $H \in \mathbb{F}_p(X)[Y]$ is a non-zero polynomial with the root $U(X)$, i.e., $H(X, U(X)) = 0$, then we clearly have $H(V(X), X) = 0$ and thus we get $p$-automaticity of the sequence of coefficients of $V$.

As we were unable to prove any general result for Problem 1.1 we concentrate on this problem in case of $\mathbf{u} = \mathbf{t}$, where $\mathbf{t}$ is the PTM sequence. To be more precise, let us consider the increasing sequence $(o_n)_{n\in\mathbb{N}}$ satisfying the equality

$$\mathcal{O} := \{m : t_m = 1\} = \{o_n : n \in \mathbb{N}_+\},$$

i.e., $o_n$ is the $n$th element of the set $\mathcal{O}$ of so called "odious" numbers. A positive integer is an odious number if the number of 1's in its binary expansion is odd. This is equivalent with the identity $t_n = s_2(n) \pmod 2 = 1$. It is interesting that the sequence $(o_n)_{n\in\mathbb{N}}$ satisfies $o_1 = 1, o_2 = 2, o_3 = 4$ and for $n \geq 1$ we have the following recurrence relations

$$\begin{aligned}
o_{4n} &= o_n - 3o_{n+1} + 3o_{2n+1} \\
o_{4n+1} &= -2o_{n+1} + 3o_{2n+1} \\
o_{4n+2} &= -o_n - 9o_{n+1} - o_{2n} + 8o_{2n+1} \\
o_{4n+3} &= -\frac{5}{3}o_n - 11o_{n+1} - \frac{5}{3}o_{2n} + 10o_{2n+1}.
\end{aligned}$$

(One can prove that in fact $o_n = 2n - 1 - t_{n-1}$ for $n \geq 1$). In particular, we deduce that the sequence $(o_n)_{n\in\mathbb{N}}$ is 2-*regular* [2,5]. The concept of a $k$-regular sequence is a generalization of $k$-automatic sequences to the case of infinite alphabets. More precisely, we say that the sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}}$ taking values in a $\mathbb{Z}$-module $R$ is a $k$-regular sequence if there exist a finite number of sequences over $R$, say $\{(r_{j,n})_{n\in\mathbb{N}} : j = 1, 2, \ldots, m\}$ such that for each integer $i \in \mathbb{N}$ and $b \in \{0, 1, \ldots, k^i - 1\}$ we have

$$u_{k^i n+b} = \sum_{j=1}^{m} b_j r_{j,n}$$

for some $b_1, \ldots, b_m \in \mathbb{Z}$ and each $n \in \mathbb{N}$. In other words, the $\mathbb{Z}$-module $R$ is finitely generated. Let us also note that the integer sequence, say $(e_n)_{n\in\mathbb{N}}$, enumerating the set

$$\mathcal{E} := \{m \in \mathbb{N} : t_m = 0\}$$

of "evil" numbers, satisfies the same recurrence relation as the sequence $(o_n)_{n\in\mathbb{N}}$ (with different initial conditions of course). In particular this sequence is 2-regular too. Let us note that several examples of $q$-automatic sequences which is characteristic function of an $k$-regular increasing sequence of integers are given in [6, p. 99–105].

2-regularity of the sequences related to the sets $\mathcal{O}$ and $\mathcal{E}$ is interesting. Indeed, this is a strong property due to the fact that for a general 2-automatic sequence $\mathbf{u}$ the sequence related to the set $\{m : u_m = 1\}$ need not be $k$-regular for any $k$. Indeed, let us consider the characteristic sequence of powers of 2, i.e., the sequence $\mathbf{u} = (u_n)_{n\in\mathbb{N}_+}$ satisfying $u_1 = 1$ and

$$u_{2n} = u_n, \qquad u_{2n+1} = 0$$

for $n \geq 1$. We then have an obvious equality $\{m : u_m = 1\} = \{2^n : n \in \mathbb{N}\}$. It is also clear that the sequence $(2^n)_{n\in\mathbb{N}}$ is not $k$-regular due to the fact that $k$-regular sequences grow polynomially fast (see [4, Theorem 16.3.1]).

Because $t_0 = 0, t_1 = 1$ we note that for the formal power series

$$F(X) = X + X^2 + X^4 + X^6 + X^7 + \cdots = \sum_{n=1}^{\infty} t_n X^n \in \mathbb{F}_2[[X]]$$