Characteristic Sequences of the Sets of Sums of Squares as Columns of Cellular Automata

Pierre-Adrien Tahay

Characteristic sequences of the sets of sums of squares as columns of cellular automata*

Pierre-Adrien Tahay[1]

FNSPE, Czech Technical University in Prague, Czech Republic
pierre.adrien.tahay@cvut.cz

Abstract. A classical result due to Lagrange states that any natural number can be written as a sum of four squares. Characterizations of integers that are a sum of two and three squares were established by Fermat, Euler, Legendre and Gauss. In this paper we denote by s_1, s_2 and s_3 the characteristic functions of the integers which are respectively sums of one, two and three squares. We recall the already known results about the nonautomaticity of s_1 and about the 2-automaticity of s_3 and we prove the nonautomaticity of s_2. In the second part, we recall a construction of s_1 as a column of a cellular automaton and we give a construction for s_3 as an immediate application of a result of Rowland and Yassawi about the construction of p-automatic sequences when p is a prime number [17]. Finally we show that s_2 is also constructible as a column of a cellular automaton and we provide an explicit construction.

Keywords: sum of squares · cellular automata · automatic sequences · nonautomatic sequences.

1 Introduction

For all integers $k \geq 1$ we define:

$$s_k(n) = \begin{cases} \mathtt{1} \text{ if } n \text{ is a sum of } k \text{ squares,} \\ \mathtt{0} \text{ otherwise.} \end{cases}$$

The function s_1 is simply the characteristic sequence of the squares and by definition $s_1(n) = 1$ if and only if there exists $m \in \mathbb{Z}$ such that $n = m^2$.

We refer to [9] to have a complete survey of the representations of integers as sums of squares.

In 1632, Girard conjectured that an odd prime number is the sum of two squares if and only if it is of the form $4k + 1$. Fermat proved this result in 1654 and the complete characterisation for all integers was obtained by Euler in the following century and gives this expression for s_2:

$$s_2(n) = \begin{cases} \mathtt{1} \text{ if all prime divisors } q \equiv 3 \pmod 4 \text{ of } n \text{ occur in } n \text{ to an even power} \\ \mathtt{0} \text{ otherwise.} \end{cases}$$

* The research received funding from the Ministry of Education, Youth and Sports of the Czech Republic through the CAAS Project .

The binary sequence $(s_2(n))_{n\geq 0} = 1,1,1,0,1,1,0,0,1,1,1,0,0,1,0,0,1,\ldots$ is equal to 1 at integers of the sequence A001481 in the OEIS [18] and 0 otherwise.

The fact that every natural number can be written as a sum of four squares was already conjectured by Bachet. The first proof of this result is due to Lagrange in 1770 and so, we have trivially for all $k \geq 4$ $s_k(n) = 1$ for all $n \geq 0$.

The most difficult case is the three squares theorem. It was established by Legendre in 1798 and Gauss in 1801. Their results give this expression for s_3:

$$s_3(n) = \begin{cases} \mathtt{0} \text{ if } n = 4^a(8m+7) \text{ with nonnegative integers } a \text{ and } m \\ \mathtt{1} \text{ otherwise.} \end{cases}$$

The binary sequence $(s_3(n))_{n\geq 0} = 1,1,1,1,1,1,1,0,1,1,1,1,1,1,1,0,1,\ldots$ is equal to 0 at integers of the sequence A004215 in the OEIS [18] and 1 otherwise.

By Langrange's four-square theorem, only the functions s_1, s_2 and s_3 are of interest to be studied. The results about the nonautomaticity of s_1 are already well-known [15,16] and a first construction in a column of a cellular automaton has been obtained by Mazoyer and Terrier in 1999 [14] and a second by Delacourt, Poupet, Sablik and Theyssier in 2011 [6] which the author has generalized with Marcovici and Stoll in 2018 to any polynomial $P \in \mathbb{Q}(X)$ with $P(\mathbb{N}) \subset \mathbb{N}$ [13]. The automaticity of s_3 is due to Cobham [4] and we use this fact to build a cellular automaton with a method developed by Rowland and Yassawi [17]. Our main results concern the study of s_2.

2 Preliminaries

In this section we fix some notations and we recall some basic results of the theory of finite automata and automatic sequences on the one hand and cellular automata on the other hand.

2.1 Words and morphisms

An *alphabet* $\mathcal{A}$ is a finite set of symbols called *letters*. The set $\mathcal{A}^*$ refers to the set of finite words over $\mathcal{A}$ which is the free monoid having neutral element the empty word ε. The *length* of a word $w = a_0 a_1 \cdots a_{n-1}$, with $a_i \in \mathcal{A}$ is the integer $|w| = n$. For an integer $k \geq 2$, we denote by Σ_k the alphabet $\{0, 1, \ldots, k-1\}$. For all $n \in \mathbb{N}$ we denote by $(n)_k$ the standard base-k representation of n. For two alphabets $\mathcal{A}$ and $\mathcal{B}$, a morphism is a map $h : \mathcal{A}^* \longrightarrow \mathcal{B}^*$ such that for all words $x, y \in \Sigma^*$ we have $h(xy) = h(x)h(y)$. If $\mathcal{A} = \mathcal{B}$ we can iterate the morphism h. For all $a \in \mathcal{A}$ we define $h^0(a) = a$ and $h^i(a) = h(h^{i-1}(a))$. For a morphism $h : \mathcal{A} \longrightarrow \mathcal{A}$, if there is an integer k such that $|h(a)| = k$ for all $a \in \mathcal{A}$, we said that h is *k-uniform*. A 1-uniform morphism is called a *coding*. We can naturally extend the notion of morphism to infinite words. We said that a k-uniform morphism $h : \mathcal{A} \longrightarrow \mathcal{A}$ is *prolongable* if there exists a letter $a \in \mathcal{A}$ and a word $w \in \mathcal{A}^* \backslash \{\varepsilon\}$ such that $h(a) = aw$. In this case, the sequence $(h^n(a))_{n\geq 0}$ converges to the infinite word $h^\omega(a) = awh(w)h^2(w)\cdots$. Moreover, $h^\omega(a)$ is the unique fixed point of h which starts with a.

2.2 Finite automata and automatic sequences

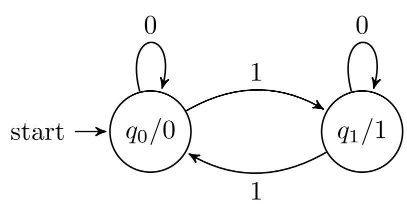
We refer to the book of Allouche and Shallit [1, Sections 5, 6] for a complete survey of the theory of finite automata and automatic sequences.

Definition 1. *A* deterministic finite automaton with output *(DFAO) is a 6-tuple* $(\mathcal{Q}, \Sigma_k, \delta, q_0, \mathcal{A}, \omega)$ *where* $\mathcal{Q}$ *is a finite set of* states, $q_0 \in \mathcal{Q}$ *is the initial state,* $\omega : \mathcal{Q} \longrightarrow \mathcal{A}$ *is the output function, and* $\delta : \mathcal{Q} \times \Sigma_k \longrightarrow \mathcal{Q}$ *is the transition function. We say that a sequence* $(u_n)_{n\geq 0}$ *of elements in* $\mathcal{A}$ *is* k*-automatic if there exists a DFAO* $(\mathcal{Q}, \Sigma_k, \delta, q_0, \mathcal{A}, \omega)$ *such that* $u_n = \omega(\delta(q_0, (n)_k))$ *for all* $n \geq 0$.

Example 1. One of the most famous automatic sequences is the Prouhet–Thue–Morse or Thue–Morse sequence. There are several equivalent ways to define it. For example, if we denote by $(t_n)_{n\geq 0}$ this sequence, it can be defined by:

$$t_n = \begin{cases} 0, & \text{if the number of 1's in } (n)_2 \text{ is even;} \\ 1, & \text{otherwise.} \end{cases}$$

With this definition it is clear that $(t_n)_{n\geq 0}$ is a 2-automatic sequence generated by the following finite automaton:



To prove that a sequence is automatic, the most natural way is to explicitly construct a finite automaton that generates it, as for the Thue–Morse sequence in the previous example. However, in practice there are many criteria for proving the automaticity of a sequence without building the automaton. One of the interests is that they make it possible to show, on the contrary, that a sequence is nonautomatic using the contraposition of one of these criteria. In a recent paper, Allouche, Shallit and Yassawi give a large number of methods to prove that a sequence is nonautomatic [2]. See also the paper of Coons [5] where the author establishes the nonautomaticity of several number theoretic functions with various criteria.

2.3 Cellular automata

Cellular automata are another model for calculation. They were first introduced in the 1940s by von Neumann to study a self-reproduction phenomenon. In 1966, Burks took up and completed von Neumann's work posthumously [20]. Cellular automata were definitely popularised by the famous Conway's Game of Life in 1970.

These objects have a very different nature from deterministic finite automata, but rather surprisingly, Rowland and Yassawi established in 2015 a very strong link between both [17]. Before recalling their result, let us begin with some definitions.

Definition 2. *Let $\mathcal{A}$ be a finite set (typically an alphabet) endowed with the discrete topology and let $\sigma : (\mathcal{A}^d)^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ be the shift map. A* cellular automaton with memory d *is a continuous map $\Phi : (\mathcal{A}^d)^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ with respect to the product topology such that $\sigma \circ \Phi = \Phi \circ \sigma$. The case $d = 1$ is the classical definition of a cellular automaton.*

By the Curtis–Hedlund–Lyndon theorem [10] the previous definition is equivalent to the following

Definition 3. *A map $\Phi : (\mathcal{A}^d)^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ is a* cellular automaton *if and only if there is a local rule $\phi : (\mathcal{A}^d)^{l+r+1} \longrightarrow \mathcal{A}$ for some $l \geq 0$ (left radius of ϕ) and some $r \geq 0$ (right radius of ϕ), such that for all $R \in (\mathcal{A}^{\mathbb{Z}})^d$ and for all $m \in \mathbb{Z}$,*

$$(\Phi(R))(m) = \phi(R(m-l), R(m-l+1), \ldots, R(m+r)).$$

Now, we define the *spacetime diagram* of a cellular automaton.

Definition 4. *If $\Phi : (\mathcal{A}^d)^{\mathbb{Z}} \longrightarrow \mathcal{A}^{\mathbb{Z}}$ is a cellular automaton with memory d, a* spacetime diagram *for Φ with initial conditions $R_0, \ldots, R_{d-1}$ is the sequence $(R_n)_{n \geq 0}$ defined recursively by $R_n = \Phi(R_{n-d}, \ldots, R_{n-1})$ for $n \geq d$.*

Consequently, for a cellular automaton with memory d, each row is determined by the previous d rows.

Definition 5. *Let q be a power of a prime number. We denote by $\mathbb{F}_q$ the finite field with q elements. In the special case where $\mathcal{A} = \mathbb{F}_q$ we say that a cellular automaton $\Phi : (\mathbb{F}_q^d)^{\mathbb{Z}} \longrightarrow \mathbb{F}_q^{\mathbb{Z}}$ with memory d is* linear *if Φ is a $\mathbb{F}_q$-linear map.*

By the Curtis–Hedlund–Lyndon theorem, a cellular automaton Φ with memory d is linear if and only if there exist coefficients $f_{j,i} \in \mathbb{F}_q$ for $-l \leq j \leq r$ and $0 \leq i \leq d-1$ such that $(\Phi(R_0, \ldots, R_{d-1}))(m) = \sum_{i=0}^{d-1} \sum_{j=-l}^{r} f_{j,i} R_i(m+j)$ for all $R_0, \ldots, R_{d-1} \in \mathbb{F}_q^{\mathbb{Z}}$ and $m \in \mathbb{Z}$.

We can now recall the result of Rowland and Yassawi.

Theorem 1 (Rowland and Yassawi [17]). *Let p a prime number and q a power of p. A sequence of elements in $\mathbb{F}_q$ is p-automatic if and only if it is a column of a spacetime diagram of a linear cellular automaton with memory over $\mathbb{F}_q$ whose initial conditions are eventually periodic in both directions.*

Remark 1. The fact that every column of a linear cellular automaton on $\mathbb{F}_q$ is p-automatic was already known since 1993 by Dumas and Litow [12]. Rowland and Yassawi established the converse by giving a complete characterization of p-automatic sequences. Moreover their proof is constructive and they give a method to have an explicit cellular automaton which generates a given p-automatic sequence. The main ingredient is Christol's theorem.

Theorem 2 (Christol [3]). *Let q a power of a prime number. A sequence $(u_n)_{n\geq 0}$ of elements in $\mathbb{F}_q$ is q-automatic if and only if its generating series $F(t) = \sum_{n\geq 0} u_n t^n$ is algebraic over $\mathbb{F}_q(t)$.*

The principle of the method of Rowland and Yassawi is to find a polynomial $P \in \mathbb{F}_q(t,x)$ such that $P(t, F(t)) = 0$ whose existence is guaranteed by Christol's theorem and define the rule of the linear cellular automaton we compute the giving p-automatic sequence from the coefficients of P. Several examples of explicit constructions of classical automatic sequences in columns of cellular automata are given directly in the article of Rowland and Yassawi or in the thesis manuscript of the author [17,19].

One of the first results about the construction of sequences as a column of cellular automata was obtained by Fischer in 1965, who built the characteristic sequence of prime numbers, using a cellular automaton with more than 30,000 states [8]. In 1997, Korec improved this result with another cellular automaton with only 11 states [11]. The geometric construction of increasing functions and closure properties has been established by Mazoyer and Terrier in 1999, which they call Fischer constructible [14]. Their constructions use signals which are a way to transmit information by connecting two cells in a cellular automaton. Marcovici, Stoll and Tahay use these kinds of contructions with signals to provide other constructions for characteristic sequences of polynomials and a construction for the Fibonacci word which is an emblematic nonautomatic sequence and which is also a Sturmian word. Recently, Dolce and Tahay extended the construction for the Fibonacci word to all Sturmian words with a quadratic slope also using signals [7]. We refer to [7,14] for a formal definition of signals in cellular automata and examples of contructions.

3 Study of the automaticity of s_1, s_2 and s_3

3.1 Nonautomaticity of the characteristic sequence of the squares

We recall the following well-known result

Proposition 1 ([15,16]). *The sequence $(s_1(n))_{n\geq 0}$ is nonautomatic.*

Ritchie proved in 1963 the fact that $(s_1(n))_{n\geq 0}$ is not a 2-automatic sequence [16]. Minsky and Papert gave an elegant criterion that can be applied to prove that $(s_1(n))_{n\geq 0}$ is not k-automatic for any $k \geq 2$.

Proposition 2 (Minsky and Papert [15]). *Let $f : \mathbb{N} \to \mathbb{N}$ be an increasing function and we define the set $\pi_f(x) = \#\{n : f(n) \leq x\}$. If the two conditions:*

1. $\lim_{x\to\infty} \frac{\pi_f(x)}{x} = 0$ *2.* $\lim_{n\to\infty} \frac{f(n+1)}{f(n)} = 1$

are satisfied, then the sequence $u = \mathbf{1}_{f(\mathbb{N})}$ *is nonautomatic.*

Proposition 1 can be directly deduced from Minsky and Papert's criterion with $f : \mathbb{N} \to \mathbb{N},\ x \to x^2$.

Remark 2. We can find an alternative proof of Proposition 1 using language theory, see Example 5.5.1 and Example 8.6.2 in [1].

Remark 3. Let $\mathcal{A} = \{a, b, c\}$. Let $f : \mathcal{A}^* \longrightarrow \mathcal{A}^*$ and $\pi : \mathcal{A} \longrightarrow \{0, 1\}$ be the morphism and the coding defined respectively by:

$$f : \begin{cases} \mathtt{a} \mapsto \mathtt{abcc} \\ \mathtt{b} \mapsto \mathtt{bcc} \\ \mathtt{c} \mapsto \mathtt{c} \end{cases} \qquad \text{and} \qquad \pi : \begin{cases} \mathtt{a}, \mathtt{b} \mapsto \mathtt{1} \\ \mathtt{c} \mapsto \mathtt{0} \end{cases}$$

It is clear that $\pi(f^\omega(a))$ generates $(s_1(n))_{n\geq 0}$. So, $(s_1(n))_{n\geq 0}$ is a morphic sequence which is nonautomatic.

3.2 Automaticity of $(s_3(n))_{n\geq 0}$

We give the result on the automaticity of $(s_3(n))_{n\geq 0}$ already known by Cobham [4] before the one on $(s_2(n))_{n\geq 0}$ which we establish in the next section.

Proposition 3 ([4]). *The sequence* $(s_3(n))_{n\geq 0}$ *is* 2*-automatic.*

Proof. Using the observation that a number is not representable in the form $4^a(8m + 7)$ if and only if its binary representation does not terminate with three successive 1's followed by an even number of 0's, Cobham gives an explicit construction of a finite automaton with 6 states, that generates the sequence $(s_3)_{n\geq 0}$ (see Fig. 1 where we use the convention to start with the most significant digit of the binary expansion to read the automaton).

Cobham deduces from the finite automaton generating $(s_3(n))_{n\geq 0}$ in Fig. 1 the following 2-uniform morphism g and the coding σ that generates the characteristic sequence of the set of sums of three squares.

Let $\mathcal{A} = \{a, b, c, d, e, f\}$. Let $g : \mathcal{A}^* \longrightarrow \mathcal{A}^*$ and $\sigma : \mathcal{A} \longrightarrow \{0, 1\}$ be the morphism and the coding defined by:

$$g : \begin{cases} \mathtt{a} \mapsto \mathtt{ab} \\ \mathtt{b} \mapsto \mathtt{ac} \\ \mathtt{c} \mapsto \mathtt{ad} \\ \mathtt{d} \mapsto \mathtt{ed} \\ \mathtt{e} \mapsto \mathtt{fb} \\ \mathtt{f} \mapsto \mathtt{eb} \end{cases} \qquad \text{and} \qquad \sigma : \begin{cases} \mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{e} \mapsto \mathtt{1} \\ \mathtt{d}, \mathtt{f} \mapsto \mathtt{0} \end{cases}$$

Then $\sigma(g^\omega(a))$ generates $(s_3(n))_{n\geq 0}$.

We give now an alternative proof using Christol's theorem which will be used to build $(s_3(n))_{n\geq 0}$ in a column of a cellular automaton in section 4.3.

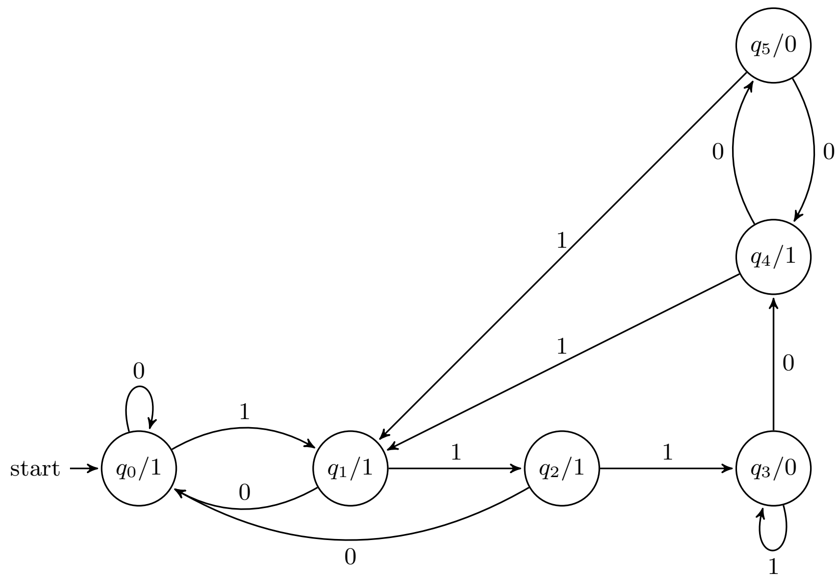


Fig. 1. A finite automaton generating the sequence $(s_3(n))_{n\geq 0}$

Proposition 4. *Let* $F(t) = \sum_{n\geq 0} s_3(n)t^n$ *be the generating series of* $(s_3(n))_{n\geq 0}$ *defined on* $\mathbb{F}_2(t)$. *Let* $P \in \mathbb{F}_2(t,x)$ *be the polynomial* $P(t,x) = t + t^2 + t^3 + t^5 + t^6 + (1+t^8)x + (1+t^8)x^4$. *Then* $P(t, F(t)) = 0$.

Proof. With Legendre and Gauss characterization for $(s_3(n))_{n\geq 0}$ we have clearly for all $n \geq 0$, $s_3(4n) = s_3(n)$, $s_3(8n+7) = 0$ and $s_3(8n+1) = s_3(8n+2) = s_3(8n+3) = s_3(8n+5) = s_3(8n+6) = 1$. Then, we have

$$\begin{aligned}
F(t) &= \sum_{n\geq 0} s_3(n)t^n \\
&= \sum_{n\geq 0} s_3(4n)t^{4n} \\
&+ \sum_{n\geq 0} s_3(8n+1)t^{8n+1} + \sum_{n\geq 0} s_3(8n+2)t^{8n+2} + \sum_{n\geq 0} s_3(8n+3)t^{8n+3} \\
&+ \sum_{n\geq 0} s_3(8n+5)t^{8n+5} + \sum_{n\geq 0} s_3(8n+6)t^{8n+6} + \sum_{n\geq 0} s_3(8n+7)t^{8n+7} \\
&= \sum_{n\geq 0} s_3(n)t^{4n} + \sum_{n\geq 0} t^{8n+1} + \sum_{n\geq 0} t^{8n+2} + \sum_{n\geq 0} t^{8n+3} + \sum_{n\geq 0} t^{8n+5} + \sum_{n\geq 0} t^{8n+6} \\
&= (F(t))^4 + (t + t^2 + t^3 + t^5 + t^6)\frac{1}{1+t^8}
\end{aligned}$$

which shows the result.

Then, the generating series of $(s_3(n))_{n\geq 0}$ is algebraic over $\mathbb{F}_2(t)$, which proves its 2-automaticity by Christol's theorem.

3.3 Nonautomaticity of the set of sums of two squares

The nonautomaticity of $(s_1(n))_{n\geq 0}$ and the 2-automaticity of $(s_3(n))_{n\geq 0}$ are well known results with different ways to prove them. A natural question now is to study the automaticity or the nonautomaticity of $(s_2(n))_{n\geq 0}$. Let us begin by recalling this definition.

Definition 6. *A sequence $(a(n))_{n\geq 1}$ is called* multiplicative *if, for all integers $m, n \geq 1$ coprime we have $a(mn) = a(m)a(n)$.*

Proposition 5. *The sequence $(s_2(n))_{n\geq 1}$ is multiplicative.*

Proof. With the identity $(a^2+b^2)(c^2+d^2) = (ac+bd)^2 + (ad-bc)^2$ we have clearly for all $m, n \geq 0, s_2(mn) \geq s_2(m)s_2(n)$.

Now, let $m, n \geq 1$ coprime.

If $s_2(mn) = 0$ we have trivially $s_2(mn) \leq s_2(m)s_2(n)$. If $s_2(mn) = 1$, by the two squares theorem this means that every prime divisor $q \equiv 3 \pmod 4$ of mn occurs in mn to an even power. Because m and n are coprime, it is necessarily also the case for all prime divisors $q \equiv 3 \pmod 4$ respectively in m and in n and then $s_2(m) = s_2(n) = 1$, which completes the proof.

We give now a criterion of nonautomaticity for multiplicative sequences.

Theorem 3 ([2,21]). *Let $v > 1$ be an integer and f a multiplicative function. Assume that for some integer $h \geq 1$ there exist infinitely many primes q_1 such that $f(q_1^h) \equiv 0 \pmod v$. Furthermore assume that there exist relatively prime integers b and c such that for all primes $q_2 \equiv c \pmod b$ we have $f(q_2) \not\equiv 0 \pmod v$. Then the sequence $(f(n))_{n\geq 1} \pmod v$ is not k-automatic for any $k \geq 2$.*

Proposition 6. *The sequence $(s_2(n))_{n\geq 1}$ is not k-automatic for any $k \geq 2$.*

Proof. Because $(s_2(n))_{n\geq 1}$ is a binary sequence it equals $(s_2(n))_{n\geq 1} \pmod 2$. Moreover $(s_2(n))_{n\geq 1}$ is multiplicative by Proposition 5, and by the Girard-Fermat's theorem, an odd prime number p is a sum of two squares if and only if $p \equiv 1 \pmod 4$. So, we can apply the previous theorem to $(s_2(n))_{n\geq 1}$ with $v = 2, h = 1, c = 1, b = 4$, the infinity of prime numbers $q_1 \equiv 3 \pmod 4$ and for q_2 we can take the prime numbers such that $q_2 \equiv 1 \pmod 4$.

4 Construction by cellular automata of s_1, s_2 and s_3

In all our constructions, time axis is oriented upward.

4.1 Cellular automaton of the characteristic sequence of the square

Here, we recall the construction of Delacourt, Poupet, Sablik and Theyssier [6] to obtain the characteristic sequence of squares by the hitting of one single signal in column 0. Vertical signals are called *walls* and represented by a straight line. Signals of other slopes are represented by arrows (see Fig. 2). The two first lines are initial conditions. We start by sending a signal of slope 1 to the northeast-direction (blue arrow). Every time it meets the wall (represented by a vertical green straight line), the wall gets shifted by one cell to the right and the blue signal of slope 1 changes its direction to a blue signal of slope -1. When the blue signal of slope -1 meets the 0-column, then a 1 is marked, and a new signal of slope 1 is sent.

4.2 Cellular automaton of the set of sums of two squares

Theorem 4. *The sequence* $(s_2(n))_{n\geq 0}$ *can be obtained as a column of a cellular automaton.*

Proof. We will use a geometric construction similar to the one for $(s_1(n))_{n\geq 0}$. First, we use the fact that for all $n \geq 0$ "$s_1(n) = 1 \Rightarrow s_2(n) = 1$" which corresponds to the fact that if an integer is a square $n = m^2$, then it is a sum of two squares, $n = m^2 + 0^2$. So, we start by taking up the cellular automaton which builds the characteristic of squares.
Using exactly the same construction but with each cell shifted one line above, we obtain all the integers that are sums of two squares of the form $n^2 + 1^2$. We represent the signals of slope 1 and slope -1 by red arrows in Fig. 3. In comparison to Fig. 2, the green wall propagates vertically one more cell before being shifted to the right when it meets a red signal.
To finish the construction, we need to mark the columns whose horizontal coordinate is a perfect square. We use the general method developed by Marcovici, Stoll and Tahay to build the polynomial sequences (see [13, Proposition 2]). We define a new signal of slope 1 in the diagonal of the spacetime diagram (grey arrows on Fig. 3). If F denotes the cellular automaton of Fig. 2, we build the cellular automaton $\sigma \circ F$. The blue arrows of slope -1 in Fig. 2 become blue walls in the columns which correspond to the perfect squares in Fig. 3. When these blue walls meet the grey diagonal signals, they continue to spread through the columns.

Now, it just remains to define signals of slope $\frac{1}{2}$ from each perfect square m^2 in the colums 0 at the same time as the blue signals of slope 1 (black arrows to the northeast direction in Fig. 3). When one of these signals meets a blue wall in a column that is a perfect square n^2, it changes its direction and we define a signal of slope $-\frac{1}{2}$ which meets column 0 at the line $m^2 + n^2$.

4.3 Linear cellular automaton of the 2-automatic sequence $(s_3)_{n\geq 0}$

In this last case, we can use the result of Rowland and Yassawi [17] by using the polynomial obtained in Proposition 4.

Let $F(t) = \sum_{n\geq 0} s_3(n)t^n$ and $P(t,x) = t+t^2+t^3+t^5+t^6+(1+t^8)x+(1+t^8)x^4$ the polynomial such that $P(t,F(t)) = 0$. We encode the spacetime diagram of a cellular automaton by the series $\sum_{n\geq 0}\sum_{m\in\mathbb{Z}} a_{n,m}t^n x^m$, where $a_{n,m}$ represents the cell on row $n \in \mathbb{N}$ and column $m \in \mathbb{Z}$.

In order to apply the method of Rowland and Yassawi, we use the transformation $x \longrightarrow 1+t+tx$ to the polynomial $P(t,x)$. We have

$$\begin{aligned} P(t,1+t+tx) &= t+t^2+t^3+t^5+t^6+1+t+tx+t^8+t^9+t^9x \\ &\quad +1+t^4+t^4x^4+t^8+t^{12}+t^{12}x^4 \\ &= tx+t^2+t^3+t^4(1+x^4)+t^5+t^6+t^9(1+x)+t^{12}(1+x^4) \end{aligned}$$

We can divide all the coefficients in the last row by t and we define $Q(t,x) = x+t+t^2+(1+x^4)t^3+t^4+t^5+(1+x)t^8+(1+x^4)t^{11}$ and we have for $G(t) = \sum_{n\geq 1} s_3(n+1)t^n$, $F(t) = 1+t+tG(t)$ and $Q(t,G(t)) = 0$.

We denote by $Q^{(0,1)}$ the derivative function of Q with respect to its second argument. Now, we have $G(0) = 0$ and $Q^{(0,1)}(t,x) = 1+t^8$ which satisfies $Q^{(0,1)}(0,0) \neq 0$.

By Rowland and Yassawi theorem, the power series $\dfrac{Q^{(0,1)}(t,x)}{Q(t,x)} = \sum_{n\geq 0} R_n(x)t^n$ encodes a spacetime diagram of a cellular automaton where for all $n \in \mathbb{N}$ and all $m \in \mathbb{Z}$ the coefficient x^m in $R_n(x)$ represents the cell at row n and column m and where the sequence $(s_3(n+1))_{n\geq 1}$ occurs in the column -2.

By using the relation $Q^{(0,1)}(t,x) = Q(t,x)\sum_{n\geq 0} R_n(x)t^n$ and by collecting the terms by common powers of t, we deduce that $R_n(x)$ satisfies the recurrence :

$$\begin{aligned} R_n(x) =& \frac{1}{x}R_{n-1}(x) + \frac{1}{x}R_{n-2}(x) + \left(\frac{1}{x}+x^3\right)R_{n-3}(x) + \frac{1}{x}R_{n-4}(x) + \frac{1}{x}R_{n-5}(x) \\ &+ \left(\frac{1}{x}+1\right)R_{n-8}(x) + \left(\frac{1}{x}+x^3\right)R_{n-11}(x), \end{aligned}$$

for all $n \geq 12$. Let $R_{-1}(x) = s_3(0)x^{-2} = x^{-2}$ and $R_0(x) = s_3(1)x^{-2} = x^{-2}$. Then, we define a cellular automaton with memory 14 $(12+1+1)$, where the sequence $(s_3(n))_{n\geq 0}$ occurs in the column -2 which is highlighted in red in the spacetime diagram in Fig. 4.

5 Open questions

1. Let $s_{k,p}$ be the characteristic sequence of sums of k pth powers. For every p, does there exist K such that $s_{k,p}$ is automatic for $k \geq K$ and nonautomatic for $k < K$? Are all these sequences representable by CA? If not, minimal p and a value of k such that $s_{k,p}$ is not representable?
2. Is it possible to generate s_3 by using a geometric construction with signals similarly to that for s_1 and s_2?
3. Sequences s_1 and s_3 are both morphic sequences. Is it also the case for s_2? Because of the nonautomaticity of s_2, of course in this case the morphism would be nonuniform like for s_1.

Acknowledgment: The author thanks the anonymous referees for the many corrections and suggestions for open questions.

A Appendices

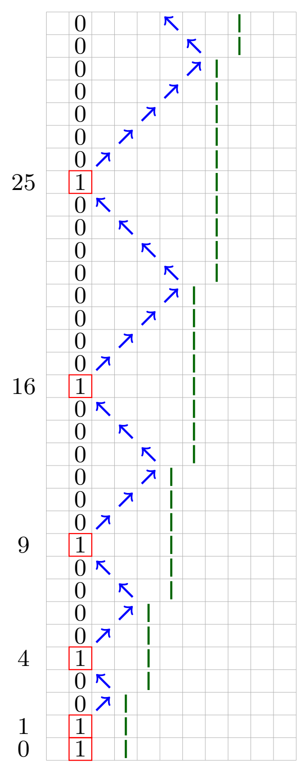


Fig. 2. Construction of $(s_1(n))_{n\geq 0}$ by a cellular automaton

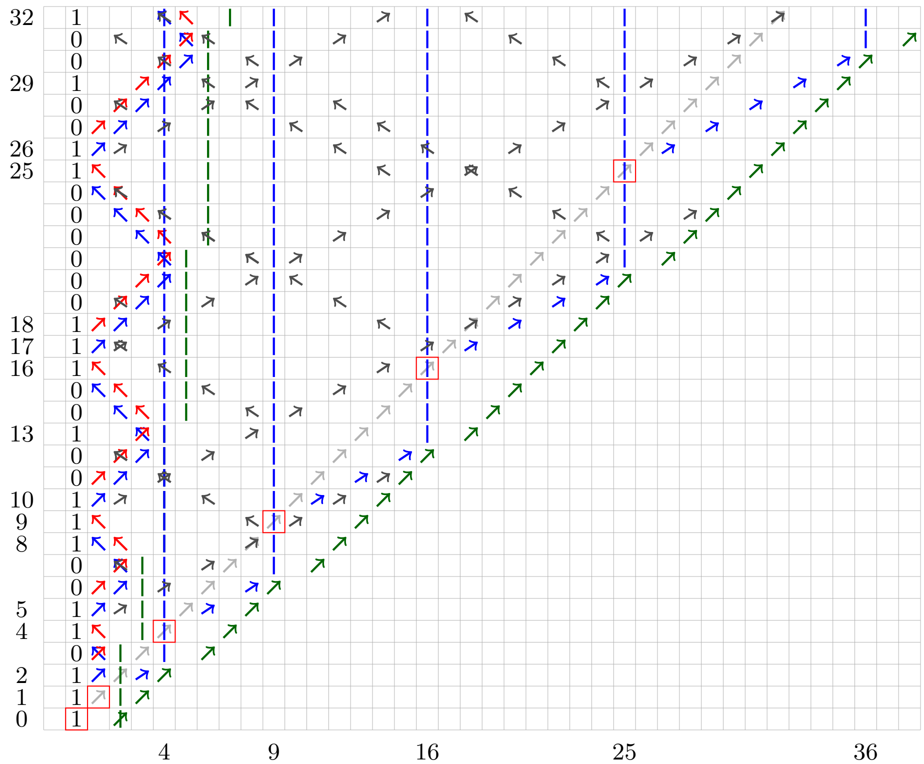


Fig. 3. Construction of $(s_2(n))_{n\geq 0}$ by a cellular automaton

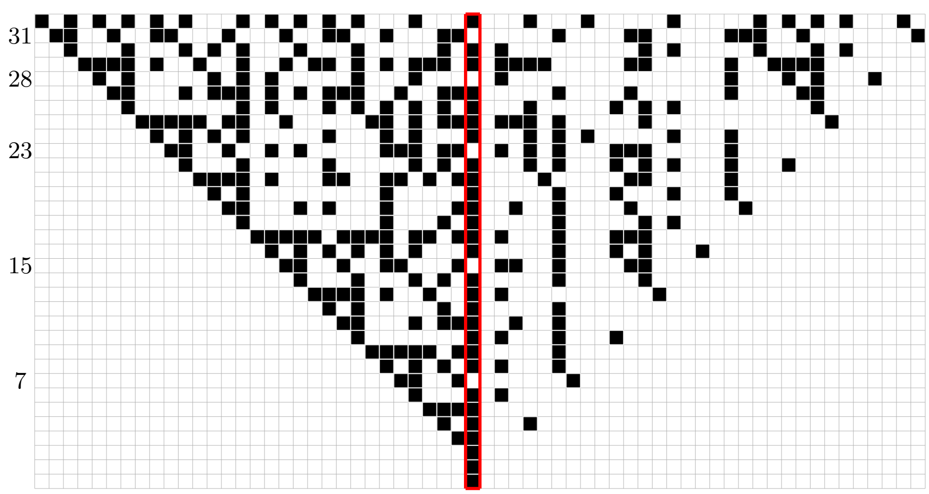


Fig. 4. Construction of $(s_3(n))_{n\geq 0}$ by a cellular automaton

References

1. J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations.* Cambridge University Press, Cambridge, 2003.
2. J.-P. Allouche, J. Shallit, and R. Yassawi. How to prove that a sequence is not automatic. *Expo. Math.*, 40:1–22, 2022.
3. G. Christol, T. Kamae, M. Mendès France, and G. Rauzy. Suites algébriques, automates et substitutions. *Bul. Soc. Math. France*, 79:401–419, 1980.
4. A. Cobham. Uniform tag sequences. *Math. Systems Theory*, 6:164–192, 1972.
5. M. Coons. (Non)automaticity of number theoretic functions. *J. de Theor. des Nr. de Bordeaux*, 22:339–352, 2010.
6. M. Delacourt, V. Poupet, M. Sablik, and G. Theyssier. Directional dynamics along arbitrary curves in cellular automata. *Theoret. Comput. Sci.*, 412:3800–3821, 2011.
7. F. Dolce and P.-A. Tahay. Column representation of sturmian words in cellular automata. *Lecture Notes in Comput. Sci.*, 13257:127–138, 2022.
8. P. C. Fischer. Generation of primes by a one–dimensional real–time iterative array. *Journal of the ACM*, 12:388–394, 1965.
9. E. Grosswald. *Representations of integers as sums of squares.* Springer New York, New York, NY, 1985.
10. G. A. Hedlund. Endomorphisms and automorphisms of the shift dynamical system. *Math. Systems Theory*, 3:320–375, 1969.
11. I. Korec. Real–time generation of primes by a one–dimensional cellular automaton with 11 states. In *Mathematical Foundations of Computer Science 1997*, volume 1295, pages 358–367. Berlin, Heidelberg, 1997.
12. B. Litow and Ph. Dumas. Additive cellular automata and algebraic series. *Theoret. Comput. Sci.*, 119:345–354, 1993.
13. I. Marcovici, T. Stoll, and P.-A. Tahay. Construction of some nonautomatic sequences by cellular automata. *Lecture Notes in Comput. Sci.*, 10875:113–126, 2018.
14. J. Mazoyer and V. Terrier. Signals in one–dimensional cellular automata. *Theoret. Comput. Sci.*, 217:53–80, 1999.
15. M. Minsky and S. Papert. Unrecognizable sets of numbers. *Journal of the ACM*, 13:281–286, 1966.
16. R. W. Ritchie. Finite automata and the set of squares. *Journal of the ACM*, 10:528–531, 1963.
17. E. Rowland and R. Yassawi. A characterization of p–automatic sequences as columns of linear cellular automata. *Adv. in Appl. Math.*, 63:68–89, 2015.
18. N.J.A. Sloane. On-line Encyclopedia of Integer Sequences. URL: `https://oeis.org/`.
19. P.-A. Tahay. *Colonnes dans les automates cellulaires et suites généralisées de Rudin—Shapiro.* PhD thesis, Université de Lorraine, 2020. `https://hal.univ-lorraine.fr/tel-03184510/document`.
20. J. von Neumann. *The theory of self reproducing automata.* University of Illinois Press, Cambridge, 1966.
21. S. Yazdani. Multiplicative functions and k–automatic sequences. *J. Théor. Nombres Bordeaux*, 13:651–658, 2001.