Perfect linear complexity profile and apwenian sequences

Jean-Paul Allouche
CNRS, IMJ-PRG
Sorbonne Université
4 Place Jussieu
F-75252 Paris Cedex 05, France
jean-paul.allouche@imj-prg.fr

Guo-Niu Han
CNRS, IRMA
Université de Strasbourg
7 rue René Descartes
F-67084 Strasbourg, France
guoniu.han@unistra.fr

Harald Niederreiter
Johann Radon Institute
RICAM, Austrian Academy of Sciences
Altenberger Straße 69
A-4040 Linz, Austria
ghnied@gmail.com

Abstract

Sequences with *perfect linear complexity profile* were defined more than thirty years ago in the study of measures of randomness for binary sequences. More recently *apwenian sequences*, first with values ± 1 , then with values in $\{0, 1\}$, were introduced in the study of Hankel determinants of automatic sequences. We explain that these two families of sequences are the same up to indexing, and give consequences and questions that this implies. We hope that this will help gathering two distinct communities of researchers.

Keywords: Perfect linear complexity profile, generalized Rueppel sequences, apwenian sequences, continued fractions with partial quotients with bounded degree, automatic sequences.

MSC: 11K45, 11K50, 11J70, 11B85, 11T71, 94A55.

1 Introduction

One of the intense pleasures in mathematical research is to discover a link between two fields that either did not seem immediately related or were studied from two distinct points of view unaware of each other. The first author, reading the paper [18], saw a relation satisfied by the so-called "0, 1-apwenian sequences" $(c_n)_{n\geq 0}$, namely

$$c_0 = 1, \text{ and } \forall n \geq 0,\ c_n \equiv c_{2n+1} + c_{2n+2} \mod 2.$$

This relation sounded familiar: he vaguely remembered a talk at SETA 98 where a similar relation was discussed, but was not able to find a reference, though he had the name of the third author in mind. And indeed the latter indicated to him two papers of his with a similar relation [35, 37]. The second of these papers contains *inter alia* a simple proof of a result initially due to Wang and Massey [48] which displays the following relation for the so-called "perfect linear complexity (PLCP) sequences" $(s_i)_{i\geq 1}$:

$$s_1 = 1, \text{ and } \forall i \geq 1,\ s_{2i+1} \equiv s_{2i} + s_i \mod 2.$$

Note that this relation already occurs in a 1977-paper of Baum and Sweet [7, p. 574]. As far as SETA 98 is concerned, the property above indeed occurs in [21, Theorem 3].

In this paper we propose to describe the links between PLCP sequences and apwenian sequences: unexpected connections arise from the two properties of $0,1$-sequences displayed above.

2 PLCP sequences

The construction of sequences with *perfect linear complexity profile* was motivated by the search for pseudorandom sequences having reasonable properties of unpredictability and randomness. We recall two definitions and some properties (see, e.g. [37]).

Definition 1. A sequence $(s_m)_{m\geq 1}$ of elements in a field F is called a k-th order *shift-register sequence* if there exist constants $a_0, a_1, \ldots, a_{k-1}$ in F such that, for all $i \geq 1$

$$s_{i+k} + a_{k-1}s_{i+k-1} + \ldots + a_1 s_{i+1} + a_0 s_i = 0.$$

Remark 2. Shift-register sequences are also called *sequences satisfying a linear recurrence relation.* Also note that $(s_m)_{m\geq 1}$ is a shift-register sequence if and only if the formal power series $\sum s_m X^m$ is rational (i.e., can be obtained as the quotient of two polynomials).

Definition 3. The *nth linear complexity* $L(n)$ of a sequence $(s_m)_{m\geq 1}$ of elements in a field F is defined as the least k such that $s_1, s_2, \ldots, s_n$ are the first n terms of a k-th order shift-register sequence. In the case where the first n terms of $(s_m)_{m\geq 1}$ are 0, $L(n)$ is defined by $L(n) = 0$. The sequence $(L(n))_{n\geq 1}$ is called the *linear complexity profile* of the sequence $(s_m)_{m\geq 1}$.

Remark 4. One clearly has $0 \leq L(n) \leq n$ and $L(n) \leq L(n+1)$.

Actually the linear complexity of a sequence $(s_m)_{n\geq 1}$ is related to the continued fraction expansion of the formal Laurent series $\sum_{m\geq 1} s_m t^{-m}$. Recall that formal Laurent series can be expanded into continued fractions similarly to the real case, where $\mathbb{R}$ is replaced by $F((t^{-1}))$ and $\mathbb{N}$ by $F[t]$. To the best of our knowledge E. Artin (in his thesis) was the first author who defined and studied these continued fractions (see [4]).

Theorem 5 (Theorem 1 in [37])**.** *Let P_j/Q_j be the convergents of the series $\sum_{m\geq 1} s_m t^{-m}$ and $L(n)$ be the nth linear complexity of the sequence $(s_m)_{m\geq 1}$. Then*

$$L(n) = \deg Q_j$$

where j is the unique integer defined by

$$\deg Q_{j-1} + \deg Q_j \leq n < \deg Q_j + \deg Q_{j+1}.$$

Now what is the "typical" linear complexity profile for a sequence? Since Rueppel [42] proved that the linear complexity profile of a random binary sequence is $n/2 + O(1)$ (where actually $0 \leq O(1) \leq 5/18$), "good" sequences are sequences whose linear complexity profile is as close to $n/2$ as possible. For more precise results on the typical linear complexity profile for sequences with values in a finite field, see [38]. Also see [44]. The following definition was adopted.

Definition 6. A sequence $(s_m)_{m\geq 1}$ of elements of a field F is said to have a *perfect linear complexity profile (PLCP)* if for all $n \geq 1$ one has $L(n) = \lfloor \frac{n+1}{2} \rfloor = \lceil \frac{n}{2} \rceil$. (Such sequences are also called 1-*perfect.*)

Theorem 7 (Theorem 2 in [37])**.** *The sequence $s_1, s_2, \ldots$ has a PLCP if and only if its generating function $\sum_{i\geq 1} s_i t^{-i}$ is irrational and has all partial quotients of degree 1 in its continued fraction expansion.*

3 Apwenian sequences

Let us first recall that the Hankel determinants of a sequence $(a_n)_{n\geq 0}$ are defined by

$$H_n = \begin{vmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_n & \cdots & a_{2n-2} \end{vmatrix} = (\det(a_{i+j}))_{0\leq i,j\leq n-1}.$$

Apwenian sequences (in fact, a variant, namely $\pm$-apwenian sequences) were defined in [14] and criteria for a sequence to be apwenian were given in [18]. The origin of these sequences is the paper [2] where —thanks to Z.-X. Wen and Z.-Y. Wen who discovered and studied sixteen simultaneous recurrence formulas— it was stated that the Hankel determinants of the ± 1 Thue-Morse sequence satisfy, for $n \geq 1$ the congruence $H_n/2^{n-1} \equiv 1 \bmod 2$, which of course implies that H_n is not zero. In particular, Hankel determinants were proved to form a 2-dimensional 2-automatic sequence and the frequency of nonzero Hankel determinants was proved to be positive, which implies properties of non-repetition in the Thue-Morse sequence and the existence of certain Padé approximants. The references in [18] point to several papers studying Hankel determinants related to automatic sequences, yielding in particular irrationality measures for certain automatic real numbers.

The case of $0, 1$-sequences was studied in [18], where the following definition was given

Definition 8. Let $\mathbf{c} = (c_n)_{n\geq 0}$ be a sequence with values in $\{0, 1\}$, such that $c_0 = 1$. Let $(H_n)_n$ be the sequence of its Hankel determinants. The sequence $\mathbf{c}$ is said to be *apwenian* if for all $n \geq 0$ one has

$$H_n \equiv 1 \bmod 2.$$

4 A single theorem gathering previous results

In this section we state a theorem that puts together seemingly unrelated previous results.

Theorem 9. *Let $(s_n)_{n\geq 1}$ be a sequence with values in $\mathbb{F}_2$ with $s_1 = 1$. Let $(c_n)_{n\geq 0}$ be the sequence defined by $c_n = s_{n+1}$ for all $n \geq 0$. (Note that $c_0 = 1$.) Then the following properties are equivalent.*

(i) *The sequence $(s_n)_{n\geq 1}$ has a perfect linear complexity profile (PLCP).*

(ii) *The Laurent series $\sum_{n\geq 1} s_n t^{-n}$ is irrational, and all the partial quotients in its (usual) continued fraction expansion have degree 1.*

(iii) *For all $n \geq 1$ one has $s_{2n+1} = s_{2n} + s_n$.*

(iv) *For all $n \geq 0$ one has $c_{2n+2} = c_{2n+1} + c_n$.*

(v) *The sequence $(c_n)_{n\geq 0}$ is apwenian.*

Proof. The equivalence of (iii) and (iv) is straightforward —just a shift of indices. But this is *the* point that links the two fields of perfect linear complexity and of nonzero Hankel determinants.

The equivalence of (i) and (ii) was proved in [37, Theorem 2] (also see Theorem 7).

The equivalence of (i) and (iii) was proved in [48], and a simpler proof was given in [37] using an idea of [45].

The equivalence of (iv) and (v) was proved in [18, Theorem 1.5].

□

Remark 10. Several examples of PLCP or apwenian sequences can be found in the literature. We cite some of them.

∗ The first two examples are the two (simple) Rueppel sequences. Recall that these sequences, say $(r_n)_{n\geq 1}$ and $(s_n)_{n\geq 1}$, are defined, respectively, as the characteristic function of the powers of 2 and the characteristic function of the integers of the form (2^k-1). Thus $r_1 = 1$, $r_{2n+1} = 0$ for all $n \geq 1$ and $r_{2n} = r_n$ for all $n \geq 1$. Similarly $s_1 = 1$, $s_{2n+1} = s_n$ for all $n \geq 1$, and $s_{2n} = 0$ for all $n \geq 1$.

∗ Generalized Rueppel sequences were defined in [35] (they are also mentioned, e.g., in [49]): first take a sequence of integers $n_1, n_2, \ldots$ such that $n_1 = 1$ and $n_{h+1} = 2n_h + c_h$ for $h = 1, 2, \ldots$ where $c_h \in \{0, 1\}$; the generalized Rueppel sequence (associated with $(c_h)_h$) is the sequence $(s_i)_{i\geq 1}$ defined by $s_i = 1$ if $i = n_h$ for some $h \geq 1$, and $s_i = 0$ otherwise.

∗ Sequences with PLCP constructed from curves over finite fields can be found in [50] (also see the references therein).

∗ Examples of apwenian and non-apwenian sequences can be found in [18].

Actually, generalized Rueppel sequences that are defined above can be viewed as generated by a map from the set of all binary sequences to the set of PLCP sequences or to the set of apwenian sequences after a small modification. There are at least three "natural" (or "simple"?) similar maps from the set of binary sequences to the set of apwenian sequences or to the set of PLCP sequences, namely $\varphi_i : \mathbf{b} = (b_n)_n \mapsto \mathbf{a} = (a_n)_n$ as follows.

∗ The first map φ_1 is the bijection defined by $\mathbf{a} = \varphi_1(\mathbf{b})$, with, in terms of (Artin) continued fraction

$$\sum_{n\geq 0} a_n t^{-n-1} = \cfrac{1}{t + b_0 + \cfrac{1}{t + b_1 + \cfrac{1}{t + b_2 + \cfrac{1}{\ddots}}}}$$

or in terms of Jacobi continued fraction

$$\sum_{n\geq 0} a_n x^n = \cfrac{1}{1 + b_0 x + \cfrac{x^2}{1 + b_1 x + \cfrac{x^2}{1 + b_2 x + \cfrac{x^2}{\ddots}}}}.$$

∗ The second map φ_2 is a bijection defined by $\mathbf{a} = \varphi_2(\mathbf{b})$ that arises from a "selector". Namely the apwenian sequence $\mathbf{a}$ satisfies

$$a_0 = 1, \quad a_n = a_{2n+1} + a_{2n+2} \ (n \geq 0).$$

Since $a_0 = 1 = a_1 + a_2$, we have two choices for a_1 and a_2: (i) $a_1 = 0$ and $a_2 = 1$, (ii) $a_1 = 1$ and $a_2 = 0$. We select one of the two choices according to the value of b_0. If $b_0 = 0$, we

select (i); if $b_0 = 1$, we select (ii). In general, if $a_n = 1$, we construct a_{2n+1} and a_{2n+2} in the same way. We define $a_{2n+1} = b_n$, $a_{2n+2} = 1 + b_n$. If $a_n = 0$, $0 = a_{2n+1} + a_{2n+2}$. We have two choices: (iii) $a_{2n+1} = a_{2n+2} = 0$, (iv) $a_{2n+1} = a_{2n+2} = 1$. We will select one of the two choices from the value of b_n. If $b_n = 0$, we select (iii); if $b_n = 1$, we select (iv). We define $a_{2n+1} = a_{2n+2} = b_n$. This process is of course exactly defining inductively the sequence by $a_0 = 1$ and, if a_n is known, then $a_{2n+1} = b_n$ and $a_{2n+2} = a_n + b_n$.

$*$ The third map $\mathbf{a} = \varphi_3(\mathbf{b})$ has been defined above (we replace (c_h) with (b_h) here). First we define a sequence $n_0, n_1, n_2, \ldots$ by

$$n_0 = 1, \quad n_{h+1} = 2n_h + b_h \ (h \geq 0).$$

Then we define $a_i = 1$ if $i = n_h$ for some $h \geq 0$, and $a_i = 0$ otherwise. Note that this map is not a bijection as the example of the fixed point of $1 \to 10$, $0 \to 11$ (the period-doubling sequence) shows: it is not a generalized Rueppel sequence because it contains infinitely many occurrences of the block 111.

An interesting question is to study which general properties of sequence $\mathbf{b}$ are kept when applying φ_i, $i = 1, 2, 3$. Typically, if $\mathbf{b}$ has some kind of "regularity", is it also the case for $\mathbf{a}$? See in particular Remark 13 below.

5 Automaticity of sequences having PLCP or being apwenian

Is it possible to impose extra "regularity" conditions on apwenian-PLCP sequences? In particular are there *automatic* or *morphic* such sequences? (For more about these notions, the reader can consult [3].) This question was asked in [18], where it was proved that the only $0,1$-apwenian sequence that is a fixed point of a uniform morphism is the period-doubling sequence (fixed point of $1 \to 10$, $0 \to 11$). Restricting to 2-uniform morphisms, one can ask a general question: are there $0,1$-apwenian sequences that are 2-automatic? We completely characterize these sequences below. To begin with, we give an easy lemma.

Lemma 1. *Let $f(t) = t + a_2t^2 + a_3t^3 + \ldots$ be a formal power series on $\mathbb{F}_2$. Then there exist unique formal power series $u \in 1 + \mathbb{F}_2[[t]]$ and $v \in t\mathbb{F}_2[[t]]$ such that $f(t) = v^2(t) + tu^2(t)$. Furthermore one has $u^2 = f'$ and $v^2 = f + tf'$.*

Proof. Decomposing f into its "even" and "odd" parts, we have (recall that $x^2 = x$ for any $x \in \mathbb{F}_2$):

$$\begin{aligned} f(t) &= (a_2t^2 + a_4t^4 + \ldots) + t(1 + a_3t^2 + a_5t^4 + \ldots) \\ &= (a_2t + a_4t^2 + \ldots)^2 + t(1 + a_3t + a_5t^2 + \ldots)^2. \end{aligned}$$

This gives the existence of $v(t) = a_2t + a_4t^2 + \ldots$ and $u(t) = 1 + a_3t + a_5t^2 + \ldots$ such that $f(t) = v^2(t) + tu^2(t)$. Conversely, if f as above satisfies $f(t) = v^2(t) + tu^2(t)$, then the uniqueness of u and v is a consequence of the uniqueness of the decomposition of f into its odd and even parts. The last assertion is clear (and it could be used to give another proof of the first assertion). □

We now address the question of which $0,1$-apwenian sequences are automatic. Our main tools are first the Christol (and Christol-Kamae-Mendès France-Rauzy) theorem (see, e.g., [3]) which asserts that a sequence $(c_n)_n$ with coefficients in $\mathbb{F}_2$ is 2-automatic if and only if the formal power series $\sum c_n t^n$ in $\mathbb{F}_2[[t]]$ is algebraic over $\mathbb{F}_2(t)$, and second the following result in [21]. (Note that this result is the same as [7, Theorem 1].)

Theorem 11 ([7, Theorem 1] and [21, Corollary 2])**.** *The binary series $f(t) = t + a_2t^2 + a_3t^3 + ...$ has a PLCP if and only if the series $u(t)$ and $v(t)$ defined above satisfy $v^2 + v = 1 + u + tu^2$. (In other words, v is the unique root in $t\mathbb{F}_2[[t]]$ of $v^2 + v = 1 + u + tu^2$.)*

It is easy to deduce from the result above the following theorem (which was proved in [7] and for which we give a proof in terms of u and v above, and a second proof using automatic sequences).

Theorem 12. *A series $f(t) = t + a_2t^2 + a_3t^3 + ... \in \mathbb{F}_2[[t]]$ which has a PLCP is algebraic over $\mathbb{F}_2(t)$ if and only if it can be written as $f(t) = v^2 + tu^2$, with u any series in $1 + t\mathbb{F}_2[[t]]$ algebraic over $\mathbb{F}_2(t)$ and v the root of $v^2 + v = 1 + u + tu^2$ lying in $t\mathbb{F}_2[[t]]$.*

Proof. The condition is clearly sufficient: if u is algebraic over $\mathbb{F}_2(t)$, then v is also algebraic since $v^2+v = 1+u+tu^2$, hence $f = v^2+tu^2$ is algebraic. Conversely it is easy to see that if $f = v^2+tu^2$ is algebraic over $\mathbb{F}_2(t)$, then so is f' (write a minimal algebraic equation for f and take the derivative: this yields that f' belongs to $\mathbb{F}_2(t)(f)$, hence is algebraic). But $f' = u^2$. Thus u is algebraic, hence v is algebraic. This implies the algebraicity of f. □

Second proof. It is also possible to prove Theorem 12 by looking at the 2-kernel of f. Recall that the 2-kernel of the sequence $(c_n)_{n\geq 0}$ is the set of subsequences

$$\{(c_{2^k n+j})_{n\geq 0},\ k \geq 0,\ j \in [0, 2^k - 1]\}.$$

Note that the 2-kernel of any sequence is the smallest set of subsequences of this sequence that is stable by the decimation operators (sometimes called Cartier operators):

$$T_0 :\ (z_n)_{n\geq 0} \to (z_{2n})_{n\geq 0} \text{ and } T_1 :\ (z_n)_{n\geq 0} \to (z_{2n+1})_{n\geq 0}.$$

Also recall that a sequence is 2-automatic if and only if its 2-kernel is a finite set (see, e.g., [3]) and that all computations on sequences below are done modulo 2.

Now, if $(a_n)_{n\geq 1}$ is a 2-automatic sequence with a PLCP, with $a_1 = 1$, one has that $(a_{2n+1})_{n\geq 0}$ is also 2-automatic. Define $(u_n)_{n\geq 0}$ and $(v_n)_{n\geq 0}$ by $u_n = a_{2n+1}$ for all $n \geq 0$ and $v_n = a_{2n}$ for all $n \geq 1$. Since (a_n) has a PLCP, we have that $v_n = a_{2n} = a_n + a_{2n+1} = a_n + u_n$ for all $n \geq 1$. Let $u := \sum_{n\geq 0} u_n t^n$ and $v := \sum_{n\geq 1} v_n t^n$. One has of course $\sum_{n\geq 1} a_n t^n = v^2 + tu^2$, and it is easy to see that $v^2 + v = 1 + u + tu^2$.

Conversely, let $(u_n)_{n\geq 0}$ be any 2-automatic sequence with $u_0 = 1$. Define the sequence $(a_n)_{n\geq 1}$ by $a_{2n+1} = u_n$ for all $n \geq 0$ and $a_{2n} = a_n + u_n$ for all $n \geq 1$. By construction the sequence $(a_n)_{n\geq 1}$ has a PLCP since $a_n + a_{2n} + a_{2n+1} = 0$ for all $n \geq 1$. Let us prove that it is 2-automatic. To make the manipulation of indices simpler, define $(\tilde{a}_n)_{n\geq 0}$ by $\tilde{a_0} = 0$ and $\tilde{a}_n = a_n$ for all $n \geq 1$. Furthermore let $(\delta_0(n))_{n\geq 0}$ be defined by $\delta_0(0) = 1$ and $\delta_0(n) = 0$ for all $n \geq 1$. The sequence $(a_n)_{n\geq 1}$ is 2-automatic if and only if $(\tilde{a}_n)_{n\geq 1}$ is 2-automatic. Since $(a_{2n+1})_{n\geq 0} = (\tilde{a}_{2n+1})_{n\geq 0}$ is 2-automatic, its 2-kernel is finite. Let $\mathcal{H}$ be the $\mathbb{F}_2$-vector space spanned by this 2-kernel, the sequence $(\tilde{a}_n)_{n\geq 0}$ and the sequence $(\delta_0(n))_{n\geq 0}$. This vector space has finite dimension, hence is finite. We will prove that the 2-kernel of $(\tilde{a}_n)_{n\geq 0}$ is included in $\mathcal{H}$. It suffices to prove that $\mathcal{H}$ is stable by the operators T_0 and T_1. The images by T_0 and by T_1 of each element in the 2-kernel of $(\tilde{a}_{2n+1})_{n\geq 0}$, as well as the images of $(\delta_0(n))_{n\geq 0}$ clearly belong to $\mathcal{H}$. It thus suffices to prove that $T_0((\tilde{a}_n)_{n\geq 0})$ and $T_1((\tilde{a}_n)_{n\geq 0})$ belong to $\mathcal{H}$. But we have

$$T_0((\tilde{a}_n)_{n\geq 0}) = (\tilde{a}_{2n})_{n\geq 0} = (\tilde{a}_n)_{n\geq 0} + (\tilde{a}_{2n+1})_{n\geq 0} + (\delta_0(n))_{n\geq 0}$$

$$\text{and } \ T_1((\tilde{a}_n)_{n\geq 0}) = (\tilde{a}_{2n+1})_{n\geq 0}. \quad \square$$

Remark 13. We give examples of sequences that have PLCP and are 2-automatic.

∗ It is easy to see that the 2-kernel of each of the two Rueppel sequences is finite, hence these sequences are 2-automatic.

∗ The period-doubling sequence is an apwenian 2-automatic sequence as shown in [18]. Shifting the indices, this is the sequence $(z_n)_{n\geq 1}$ with values in $\mathbb{F}_2$ defined by $z_1 = 1$, $z_{2n} = 1 + z_n$ for all $n \geq 1$, and $z_{2n+1} = 1$ for all $n \geq 1$. In particular $z_n + z_{2n} + z_{2n+1} = 0$ for all $n \geq 1$.

∗ Define the sequence $(w_n)_n$ with values in $\mathbb{F}_2$ by $w_1 = 1$, $w_{2n+1} = 1 + w_n$ for all $n \geq 1$, and $w_{2n} = 1$ for all $n \geq 1$. This sequence is 2-automatic (its 2-kernel is finite). Clearly $w_n + w_{2n} + w_{2n+1} = 0$ for all $n \geq 1$.

Now we give examples of precise questions about (non-)conservation of "regularity" properties alluded to at the end of Remark 10 above. Namely: *If the sequence* $\mathbf{b}$ *is eventually periodic (i.e., periodic from some index on), algebraic, automatic, regular, morphic, ..., what can be said about* $\mathbf{a} = \varphi_i(\mathbf{b})$*?* Here are some answers.

∗ If $\mathbf{b}$ is eventually periodic, then $\varphi_1(\mathbf{b})$ is quadratic. (This is well known and is the same as in the real case.)

∗ If $\mathbf{b}$ is the Thue-Morse sequence, then $\varphi_1(\mathbf{b})$ is 2-automatic, and its associated formal power series is algebraic of degree 4 [20]. Other examples of automatic sequences $\mathbf{b}$, for which $\varphi_1(\mathbf{b})$ is automatic can be found in [24, 25]. It would be interesting to find an example of a sequence $\mathbf{b}$ such that $\varphi_1(\mathbf{b})$ is not automatic. The examples we have in mind go the other way round: typically the Baum and Sweet sequence [6] is a 2-automatic sequence, but its associated power series has a continued fraction expansion —with partial quotients of degree 1 or 2— that is not automatic [33].

∗ As seen in Theorem 12 above, $\varphi_2(\mathbf{b})$ is 2-automatic if and only if $\mathbf{b}$ is 2-automatic.

∗ We have that $\varphi_3(\mathbf{b})$ is 2-automatic if and only if $\mathbf{b}$ is eventually periodic. Hint: the integers n_h are exactly the integers with base-2 expansions 1, $1b_0$, $1b_0b_1$, etc.; these expansions can be used to feed a direct automaton (i.e., an automaton reading the digits of the integers from left to right). J. Shallit (private communication) gave us the more formal proof below.

Theorem 14 (J. Shallit)**.** *Let* $(b_h)_h$ *a sequence with values in* $\{0, 1\}$*. The generalized Rueppel sequence* (s_i) *associated with* $(b_h)_h$ *is 2-automatic if and only if* (b_h) *is eventually periodic.*

Proof. We take the notation of Remark 10 for the generalized Rueppel sequences where $(c_h)_h$ is replaced with $(b_h)_h$. As noted above the set of all base-2 representations of the integers n_h is the set $\{1b_0b_1b_2 \dots b_i : i \geq 0\}$. In other words these representations are the prefixes of the infinite word $1b_0b_1b_2 \dots$ We know that a binary sequence is 2-automatic if and only if the base-2 representations of the indices where it is equal to 1 form a regular set (see, e.g., [3]). An easy classical result asserts that the set of all prefixes of an infinite word is regular if and only if that word is ultimately periodic. The result follows by combining these two claims. □

Remark 15. The linear complexity of sequences generated by *cellular automata* was studied by several authors (see, e.g., [15] and the references therein), but this is of course a quite different story.

6 More Laurent series with partial quotients of degree one

For a Laurent series, the equivalence between having partial quotients of degree 1 and having PLCP, namely properties (i) and (ii) in Theorem 9 above for the case $\mathbb{F}_2$, is actually true for any field (see [37]). It is thus interesting to find such series. Among many examples in the literature, we cite the following having all their partial quotients of degree 1. Mills and Robbins [32] give, for each prime $p \geq 3$, explicit examples of formal Laurent series with coefficients in $\mathbb{F}_p$ that are algebraic over $\mathbb{F}_p(X)$ and have all their partial quotients of degree 1. Examples in characteristic 2 can be found in [46, p. 290]. Quartic power series in $\mathbb{F}_3((X^{-1}))$ are given by Lasjaunias [22]. More general series are studied in [24] and [25] (where the name *flat series* is used for algebraic series having all partial quotients of degree 1). Hyperquadratic power series in $\mathbb{F}_3((X^{-1}))$ are given in [16], while [26] gives a large family of examples in odd characteristic and [23] addresses the case of characteristic 2.

Addressing the previous examples uses their *hyperquadraticity*: an irrational element α in $\mathbb{F}_q((X^{-1}))$ is called hyperquadratic if it satisfies an equation $\alpha = (A\alpha^r + B)/(C\alpha^r + D)$, where r is a power of the characteristic of $\mathbb{F}_q$ and the coefficients A, B, C, D belong to $\mathbb{F}_q[T]$. For examples of a different nature, van der Poorten and coauthors studied series related to the *folding lemma* for continued fractions and to paperfolding or to playing between finite fields and rational numbers, see [41] and [1]. We cannot resist to quote a remark of A. van der Poorten in [1]: *By the way, it is rather easy to see that, in the 'generic case', an infinite series has all its partial quotients linear — though it is debatable whether a series with coefficient* 0 *or* 1 *is 'generic'.* The results in [1] were generalized in [10], where a formal Laurent series is called *normal* if its continued fraction is not finite and all of the partial quotients, except perhaps the first, have degree 1. About continued fractions of Laurent series, and, in particular, *normal* series, one should certainly read the nice paper [40]. Interestingly enough a variation on the Thue-Morse power series, namely the Laurent series $t\prod_{k\geq 0}(1-t^{-2^k})$ has all its partial quotients of degree 1, see [5, Prop. 3.2].

Note that the seemingly simpler study of rational functions p/q having all partial quotients of degree one in their continued faction expansion, is far from being straightforward, see, e.g., [31, 36, 47, 8, 27, 13]; note that in [8], on p. 104, there is a link, for sequences with values in $\mathbb{F}_2$, between certain Hankel determinants being different from 0 and the relations $s_n + s_{2n} + s_{2n+1} = 0 \bmod 2$ for certain values of n.

Remark 16. We have seen that several terms are used to name Laurent series with all their partial quotients of degree 1, or the sequences of coefficients of these series: namely PLCP, flat, normal. Another term is used, e.g., in [8, 27]: the *orthogonal multiplicity* of a monic polynomial g is the number of polynomials f, coprime to g and of degree less than the degree of g, such that all the partial quotients of the continued fraction expansion of f/g are of degree 1: hence saying that there exists f such that f/g has all its partial quotients of degree 1, is the same as saying that g has positive orthogonal multiplicity. Let us also point out the name *badly approximable* for a series whose continued fraction expansion has partial quotients with bounded degree.

7 Miscellanea

In this section we propose a few questions that we could either not answer, or for which we only have partial results.

∗ We have seen conditions on a PLCP/apwenian sequence to be 2-automatic. The question of whether an apwenian sequence can be an iterative fixed point of a d-substitution was addressed

in [18]: *the only* $(0,1)$*-apwenian sequence which is a fixed point of some* d*-substitution is the period doubling sequence*. More generally, are there PLCP/apwenian sequences that are d-automatic for some d not a power of 2, or even morphic? Note that the authors of [18] conjecture that the *fixed points of substitutions of non-constant length on* $\{0,1\}$ *cannot be apwenian.*

$*$ Is there a "combinatorial" condition for PLCP on larger alphabets similar to the relation $s_{2n+1} + s_{2n} + s_n = 0$? And what would be apwenian sequences in this context? Note that a relation of the same kind appears in [28, p. 270].

$*$ Is the sequence of Hankel determinants of a d-automatic sequence a d-regular sequence or, once reduced modulo some k, a d-automatic sequence? Some positive results in this direction can be found in [19].

$*$ Can the previous question be addressed by studying in detail the way of transforming a Stieltjes, or a Jacobi, or a Hankel, continued fraction into a usual continued fraction?

$*$ The linear complexity profile of some automatic sequences was computed or evaluated in [29]. In particular, it is proved that several q-automatic sequences over $\mathbb{F}_q$, which are thus somehow "predictable", have their linear complexity $L(n)$ of order of magnitude n. For example, *the linear complexity profile of the Thue-Morse sequence is the sequence* $(2\lfloor\frac{n+2}{4}\rfloor)_{n\geq 1}$ *and the linear complexity profile of the period-doubling sequence is the sequence* $(\lfloor\frac{n+1}{2}\rfloor)_{n\geq 1}$ (by the way this sequence is the only binary sequence having both perfect linear complexity profile and *perfect lattice profile*, see [12] for more details). For which other automatic or morphic sequences is it possible to compute exactly their linear complexity profile?

$*$ How does linear complexity of a sequence compare with other "complexities"? The reader can, in particular, consult the following papers: [43] (where *inter alia* a measure called *rationality* that generalizes linear complexity is introduced); [11] (where "expansion" complexity is defined, and where a discussion about terminology takes place at the end of the paper about the terms "perfect linear complexity profile", "almost perfect linear complexity profile", "d-almost perfect complexity" and "d-perfect"); [30, 17] (where expansion complexity is studied, and also compared with linear complexity); [39] and [9] (where relations between correlation measure and Gowers norm, and linear complexity and correlation measure are discussed). A moral global view can be found in [34] (see in particular Sections 10.2 and 10.4).

Acknowledgments We thank J. Shallit for having provided a proof of Theorem 14. The first author would like to thank A. Lasjaunias and J.-Y. Yao for interesting discussions.

References

[1] J.-P. Allouche, M. Mendès France, A. J. van der Poorten, An infinite product with bounded partial quotients, *Acta Arith.* **59** (1991), 171–182.

[2] J.-P. Allouche, J. Peyrière, Z.-X. Wen, Z.-Y. Wen, Hankel determinants of the Thue-Morse sequence, *Ann. Inst. Fourier (Grenoble)* **48** (1998), 1–27.

[3] J.-P. Allouche, J. Shallit, *Automatic Sequences. Theory, Applications, Generalizations*, Cambridge University Press, Cambridge, 2003.

[4] E. Artin, Quadratische Körper im Gebiete der höheren Kongruenzen. I., *Math. Z.* **19** (1924), 153–206.

[5] D. Badziahin, E. Zorin, Thue-Morse constant is not badly approximable, *Int. Math. Res. Not.* **19** (2015), 9618–9637.

[6] L. E. Baum, M. M. Sweet, Continued fractions of algebraic power series in characteristic 2, *Ann. of Math.* **103** (1976), 593–610.

[7] L. E. Baum, M. M. Sweet, Badly approximable power series in characteristic 2, *Ann. of Math.* **105** (1977), 573–580.

[8] S. R. Blackburn, Orthogonal sequences of polynomials over arbitrary fields, *J. Number Theory* **68** (1998), 99–111.

[9] N. Brandstätter, A. Winterhof, Linear complexity profile of binary sequences with small correlation measure, *Period. Math. Hungar.* **52** (2006), 1–8.

[10] D. G. Cantor, On the continued fractions of quadratic surds, *Acta Arith.* **68** (1994), 295–305.

[11] C. Diem, On the use of expansion series for stream ciphers, *LMS J. Comput. Math.* **15** (2012), 326–340.

[12] G. Dorfer, W. Meidl, A. Winterhof, Counting functions and expected values for the lattice profile at n, *Finite Fields Appl.* **10** (2004), 636–652.

[13] C. Friesen, Rational functions over finite fields having continued fraction expansions with linear partial quotients, *J. Number Theory* **126** (2007), 185–192.

[14] H. Fu, G.-N. Han, Computer assisted proof for Apwenian sequences, ISSAC 2016 Conference, Waterloo, Ontario, Canada, 2016.

[15] A. Fúster-Sabater, Cellular automata in stream ciphers, in *Recent trends in cryptography*, pp. 1–20, Contemp. Math. **477**, Amer. Math. Soc., Providence, RI, 2009.

[16] D. Gómez-Pérez, A. Lasjaunias, Hyperquadratic power series in $\mathbb{F}_3((T^{-1}))$ with partial quotients of degree 1, *Ramanujan J.* **33** (2014), 219–226.

[17] D. Gómez-Pérez, L. Mérai, H. Niederreiter, On the expansion complexity of sequences over finite fields, *IEEE Trans. Inform. Theory* **64** (2018), 4228–4232.

[18] Y.-J. Guo, G.-N. Han, W. Wu, Criterions for apwenian sequences, Preprint, 2020, available at `https://arxiv.org/abs/2001.10246`.

[19] Y. Hu, G. Wei-Han, On the automaticity of the Hankel determinants of a family of automatic sequences, *Theoret. Comput. Sci.* **795** (2019), 154–164.

[20] Y. Hu, G. Wei-Han, On the algebraicity of Thue-Morse and period-doubling continued fractions, Preprint, 2020, Available at `https://arxiv.org/abs/2005.11937`.

[21] D. Kohel, S. Ling, C. Xing, Explicit sequence expansions, in *Sequences and their applications (Singapore, 1998)*, Springer Ser. Discrete Math. Theor. Comput. Sci., Springer, London, 1999, 308–317.

[22] A. Lasjaunias, Quartic power series in $\mathbb{F}_3((T^{-1}))$ with bounded partial quotients, *Acta Arith.* **95** (2000), 49–59.

[23] A. Lasjaunias, A note on hyperquadratic continued fractions in characteristic 2 with partial quotients of degree 1, *Acta Arith.* **178** (2017), 249–256.

[24] A. Lasjaunias, J.-J. Ruch, Algebraic and badly approximable power series over a finite field, *Finite Fields Appl.* **8** (2002), 91–107.

[25] A. Lasjaunias, J.-J. Ruch, Flat power series over a finite field, *J. Number Theory* **95** (2002), 268–288.

[26] A. Lasjaunias, J.-Y. Yao, Hyperquadratic continued fractions in odd characteristic with partial quotients of degree one, *J. Number Theory* **149** (2015), 259–284.

[27] A. G. B. Lauder, Polynomials with odd orthogonal multiplicity, *Finite Fields Appl.* **4** (1998), 453–464.

[28] A. G. B. Lauder, Continued fractions of Laurent series with partial quotients from a given set, *Acta Arith.* **90** (1999), 251–271.

[29] L. Mérai, A. Winterhof, On the Nth linear complexity of automatic sequences, *J. Number Theory* **187** (2018), 415–429.

[30] L. Mérai, H. Niederreiter, A. Winterhof, Expansion complexity and linear complexity of sequences over finite fields, *Cryptogr. Commun.* **9** (2017), 501–509.

[31] J. Mesirov, M. Sweet, Continued fraction expansions of rational expressions with irreducible denominators in characteristic 2, *J. Number Theory* **27** (1987), 144–148.

[32] W. Mills, D. Robbins, Continued fractions for certain algebraic power series, *J. Number Theory* **23** (1986), 388–404.

[33] M. Mkaouar, Sur le développement en fraction continue de la série de Baum et Sweet, *Bull. Soc. Math. France* **123** (1995), 361–374.

[34] G. L. Mullen, D. Panario (Eds.), *Handbook of Finite Fields*, Discrete Mathematics and its Applications (Boca Raton), CRC Press, Boca Raton, FL, 2013.

[35] H. Niederreiter, Continued fractions for formal power series, pseudorandom numbers, and linear complexity of sequences, Contributions to General Algebra **5** (Proc. Salzburg Conf., 1986), pp. 221–233, B. G. Teubner, Stuttgart, 1987.

[36] H. Niederreiter, Rational functions with partial quotients of small degree in their continued fraction expansion, *Monatsh. Math.* **103** (1987), 269–288.

[37] H. Niederreiter, Sequences with almost perfect linear complexity profile, in *Advances in Cryptology – EUROCRYPT '87*, Lecture Notes in Comput. Sci., **304**, pp. 37– 51, Springer, Berlin, 1988.

[38] H. Niederreiter, The probabilistic theory of linear complexity, in *Advances in Cryptology – EUROCRYPT'88 (Davos, 1988)*, pp. 191–209, Lecture Notes in Comput. Sci. **330**, Springer, Berlin, 1988.

[39] H. Niederreiter, J. Rivat, On the Gowers norm of pseudorandom binary sequences, *Bull. Aust. Math. Soc.* **79** (2009), 259–271.

[40] A. van der Poorten, Formal power series and their continued fraction expansion, in *Algorithmic number theory (Portland, OR, 1998)*, pp. 358–371, Lecture Notes in Comput. Sci. **1423**, Springer, Berlin, 1998.

[41] A. J. van der Poorten, J. Shallit, Folded continued fractions, *J. Number Theory* **40** (1992), 237–250.

[42] R. A. Rueppel, Linear complexity and random sequences, in *Advances in Cryptology - EUROCRYPT'85 (F. Pichler, ed.)*, Lecture Notes in Comput. Sci. **219**, pp. 167–188, Springer, Berlin, 1986.

[43] J. Shallit, Automaticity and rationality, in Descriptional complexity of automata, grammars and related structures (Magdeburg, 1999), *J. Autom. Lang. Comb.* **5** (2000), 255–268.

[44] L. Shen, J. Xu, H. Jing, On the largest degree of the partial quotients in continued fraction expansions over the field of formal Laurent series, *Int. J. Number Theory* **9** (2013), 1237–1247.

[45] Y. Taussat, *Approximation diophantienne dans un corps de séries formelles*, Thèse, Université Bordeaux I, 1986.

[46] D. S. Thakur, Diophantine approximation exponents and continued fractions for algebraic power series, *J. Number Theory* **79** (1999), 284–291.

[47] M. Wang, Linear complexity profiles and continued fractions, in *Advances in cryptology - EUROCRYPT'89 (Houthalen, 1989)*, Lecture Notes in Comput. Sci. **434**, pp. 571–585, Springer, Berlin, 1990.

[48] M.-Z. Wang, J. L. Massey, The characterization of all binary sequences with a perfect linear complexity profile, Paper presented at EUROCRYPT'86, Linköping, 1986.

[49] C. Xing, H. Niederreiter, Applications of algebraic curves to constructions of codes and almost perfect sequences, in *Finite fields and applications (Augsburg, 1999)*, pp. 475–489, Springer, Berlin, 2001.

[50] C. Xing, H. Niederreiter, K. Y. Lam, C. Ding, Constructions of sequences with almost perfect linear complexity profile from curves over finite fields, *Finite Fields Appl.* **5** (1999), 301–313.