A family of constant-length substitutions on the p-adic integers

Eric Rowland[*]
Université de Liège, Belgium
rowland@lacim.ca

Reem Yassawi[†]
Trent University, Peterborough, Canada
ryassawi@trentu.ca

March 31, 2014

Abstract

We set the stage for studying some substitution subshifts defined on an infinite alphabet. We consider sequences of p-adic integers that project modulo p^α to a p-automatic sequence for every $\alpha \geq 1$. Examples include algebraic sequences of integers, which satisfy this property for any prime p, and some cocycle sequences, which we show satisfy this property for a fixed p. By considering the shift-orbit closure of such a sequence in $\mathbb{Z}_p^{\mathbb{N}}$, we describe how this subshift is a letter-to-letter coding of a subshift generated by a constant-length substitution defined on a closed subset of $\mathbb{Z}_p$.

1 Introduction

1.1 Overview

A *substitution* (or *morphism*) on an alphabet $\mathcal{A}$ is a map $\theta : \mathcal{A} \to \mathcal{A}^*$, extended to $\mathcal{A}^{\mathbb{N}}$ using concatenation. The substitution is *length-k* (or *k-uniform*) if for each $a \in \mathcal{A}$, $\theta(a)$ is of length k. The (extensive) literature on substitutions has traditionally focused on the case where $\mathcal{A}$ is finite. Some exceptions include recent work, for example in [Fer06] and [Mau06]. Substitutions on a countably infinite alphabet have been used to describe lexicographically least sequences on $\mathbb{N}$ avoiding certain

[*]Supported by a Marie Curie Actions COFUND fellowship.
[†]Partially supported by an NSERC grant.

patterns [GPS09, RS12], and they have been used in the combinatorics literature to enumerate permutations avoiding patterns [Wes96].

In this article we present another approach, which generates a family of constant-length substitutions on an uncountable alphabet. There are two main advantages to this approach. One is that our fixed points, and their corresponding subshifts, live in a compact metric space, which means that we have all the tools of classical measurable and topological dynamics at our disposal. A second advantage is that we define a context for investigating empirical observations concerning convergence of certain sequences in $\mathbb{Z}_p$.

Our motivation comes from the following classical results. Let $(a(n))_{n\geq 0}$ be an *automatic sequence* (see Definition 2.1). Cobham's theorem (Theorem 2.5) characterizes an automatic sequence as being the coding, under a letter-to-letter map, of a fixed point of a constant-length substitution. Christol's theorem (Theorem 2.9) characterizes p-automatic sequences for prime p; they are precisely the sequences whose generating function is algebraic over a finite field of characteristic p.

What if $(a(n))_{n\geq 0}$ is an algebraic sequence of integers (or, more generally, p-adic integers in $\mathbb{Z}_p$), whose generating function is the root of a polynomial with coefficients from $\mathbb{Z}[x]$ (or $\mathbb{Z}_p[x]$)? Certainly, $(a(n) \bmod p)_{n\geq 0}$ is p-automatic, since projecting modulo p a polynomial for which $\sum_{n\geq 0} a(n)x^n$ is a root yields a polynomial for which $\sum_{n\geq 0}(a(n) \bmod p)x^n$ is a root. But what if one projects modulo p^α for $\alpha > 1$? The following generalization of Christol's theorem, proved by Christol [Chr74] and generalized to the multivariate setting by Denef and Lipshitz [DL87], answers this question.

Theorem 1.1. *Let $a(n)_{n\geq 0}$ be a sequence of p-adic integers such that $\sum_{n\geq 0} a(n)x^n$ is algebraic over $\mathbb{Z}_p(x)$, and let $\alpha \geq 1$. Then $(a(n) \bmod p^\alpha)_{n\geq 0}$ is p-automatic.*

Thus if $a(n)_{n\geq 0}$ is a sequence of integers which is algebraic over $\mathbb{Z}_p[x]$, then Theorem 1.1, combined with Cobham's theorem, tells us that modulo p^α, this sequence is the coding of a fixed point of θ_α a substitution of length p. If $a(n)_{n\geq 0}$ is algebraic over $\mathbb{Z}[x]$, then this is true for any prime p. Examples include notable combinatorial sequences such as the sequences of Catalan numbers and central trinomial coefficients. For a prime power p^α, one can explicitly compute an automaton for $(a(n) \bmod p^\alpha)_{n\geq 0}$ [RY14]. On the other hand, there are sequences of integers which are not algebraic, but which project modulo p^α, for a single prime p, to a p-automatic sequence. These arise from cocycle maps between two substitutions of the same constant length which have the same incidence matrices. We discuss them in Section 3.

In this article we formalize the idea of studying sequences from both families as fixed points of length-p substitutions which are obtained as "inverse limit substitutions". We also consider the shift-orbit closure of an algebraic sequence as the inverse limit of the corresponding codings of the substitution subshifts generated by $(\theta_\alpha)_{\alpha\geq 1}$. The algebraic sequence itself will be the coding of a fixed point of a length-p substitution.

Many questions that have been investigated for substitution systems on a finite alphabet can now be asked. What are the fixed points of these substitutions? What are the shift-invariant measures, of the defined subshifts? What spectral properties do these subshifts have? We can also ask questions about the inverse limit of a sequence of finite automata. What properties does it have? What languages does it generate? We leave these questions to future papers. For the time being we lay the groundwork.

2 Finite automata and automatic sequences

We first give the formal definition of a finite automaton.

Definition 2.1. Let $k \geq 2$. A *k-deterministic finite automaton with output* (k-DFAO) is a 6-tuple $(\mathcal{S}, \Sigma_k, \delta, s_0, \mathcal{A}, \tau)$, where $\mathcal{S}$ is a finite set of "states", $s_0 \in \mathcal{S}$ is the *initial state*, $\Sigma_k = \{0, 1, \dots, k-1\}$, $\mathcal{A}$ is a finite alphabet, $\tau : \mathcal{S} \to \mathcal{A}$ is the *output function*, and $\delta : \mathcal{S} \times \Sigma_k \to \mathcal{S}$ is the *transition function*.

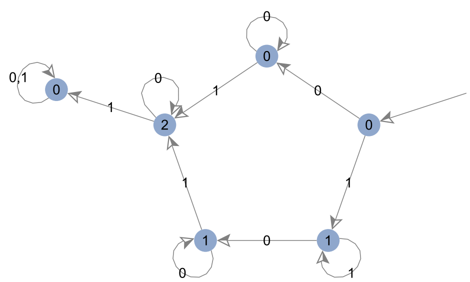
We shall write $(\mathcal{S}, \Sigma_k, \delta, s_0)$ to denote the automaton $(\mathcal{S}, \Sigma_k, \delta, s_0, \mathcal{S}, \mathrm{id})$ whose output alphabet is its set of states. The function δ extends in a natural way to the domain $\mathcal{S} \times \Sigma_k^+$, where Σ_k^+ is the set of nonempty words on the alphabet Σ_k. Namely, define $\delta(s, n_\ell \cdots n_1 n_0) := \delta(\delta(s, n_0), n_l \cdots n_1)$ recursively. Given a natural number n and an integer $k \geq 2$, we write $(n)_k = n_\ell \cdots n_1 n_0 \in \Sigma_k^*$ for the standard base-k representation of n where $n = n_0 + n_1 k + \cdots + n_\ell k^\ell$ and $n_\ell \neq 0$. We can feed $(n)_k$, beginning with the least significant digit n_0, into an automaton as follows. (Recall that the standard base-k representation of 0 is the empty word.)

Definition 2.2. A sequence $(a(n))_{n\geq 0}$ of elements in $\mathcal{A}$ is *k-automatic* if there is a k-DFAO $\mathcal{M} = (\mathcal{S}, \Sigma_k, \delta, s_0, \mathcal{A}, \tau)$ such that $a(n) = \tau(\delta(s_0, (n)_k))$ for each $n \geq 0$.

For ease of notation we will write $a(n)_{n\geq 0}$ for $(a(n))_{n\geq 0}$. When we obtain an automatic sequence $a(n)_{n\geq 0}$ by feeding into an automaton the base-k representation of n starting with the least significant digit as in Definition 2.2, we say that the automaton generates $a(n)_{n\geq 0}$ *in reverse reading*. Alternatively, we can feed each $(n)_k = n_\ell \cdots n_1 n_0$ into an automaton beginning with the most significant digit n_ℓ, in which case we say that the sequence we obtain is generated in *direct reading*. A sequence is automatic in reverse reading if and only if it is automatic in direct reading, although the automata that generate a given sequence in these two ways are generally different. When we need to, we will denote the automaton that generates

$a(n)_{n\geq 0}$ in reverse reading by $\mathcal{M}_{\mathrm{R}}$, and the automaton that generates by $a(n)_{n\geq 0}$ in direct reading by $\mathcal{M}_{\mathrm{D}}$.

Example 2.3. Consider the following automaton for $k = 2$. Each of the six states is represented by a vertex, labeled with its output under τ. Edges between vertices illustrate δ. The unlabeled edge points to the initial state.



The 2-automatic sequence produced by this automaton in reverse reading is

$$a(n)_{n\geq 0} = 0, 1, 2, 1, 2, 2, 0, 1, 2, 2, 0, 2, 0, 0, 0, 1, \ldots.$$

The 2-automatic sequence produced by this automaton in direct reading is

$$a(n)_{n\geq 0} = 0, 1, 1, 1, 1, 2, 1, 1, 1, 2, 2, 0, 1, 2, 1, 1, \ldots.$$

2.1 Cobham's characterization of automatic sequences

Definition 2.4. Let $\mathcal{A}$ be an alphabet. A *substitution* is a map $\theta : \mathcal{A} \to \mathcal{A}^*$. The map θ extends to a map $\theta : \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^* \cup \mathcal{A}^{\mathbb{N}}$ by concatenation: $\theta(a(n)_{n\geq 0}) := \theta(a(0)) \cdots \theta(a(k)) \cdots$. If there is some k with $|\theta(a)| = k$ for each $a \in \mathcal{A}$, then we say that θ is a *length-k* substitution. A *fixed point* of θ is a sequence $a(n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}}$ such that $\theta(a(n)_{n\geq 0}) = a(n)_{n\geq 0}$. Let $\tau : \mathcal{A} \to \mathcal{B}$ be a map; it induces a *letter-to-letter projection* $\tau : \mathcal{A}^{\mathbb{N}} \to \mathcal{B}^{\mathbb{N}}$.

Cobham's theorem [Cob72] gives us the relationship between k-automatic sequences and fixed points of length-k substitutions:

Theorem 2.5. *A sequence is k-automatic if and only if it is the image, under a letter-to-letter projection, of a fixed point of a length-k substitution.*

If a sequence $a(n)_{n\geq 0}$ is generated in direct reading by $\mathcal{M} = (\mathcal{S}, \Sigma_k, \delta, s_0, \mathcal{A}, \tau)$, then the transition map δ gives us the substitution described by Cobham's theorem. Namely, let $\theta(s) := \delta(s, 0)\delta(s, 1) \cdots \delta(s, k-1)$ for each state $s \in \mathcal{S}$. If $\delta(s_0, 0) \neq s_0$, introduce a new letter s_0' and let $\theta(s_0') := s_0'\delta(s_0, 1) \cdots \delta(s_0, k-1)$.

Example 2.6. Consider the length-2 substitution on the alphabet $\{s_0', s_0, s_1, s_2, s_3, s_4, s_5\}$ defined by

$$\begin{aligned} \theta(s_0') &= s_0' s_2 & \theta(s_3) &= s_3 s_5 \\ \theta(s_0) &= s_1 s_2 & \theta(s_4) &= s_4 s_3 \\ \theta(s_1) &= s_1 s_3 & \theta(s_5) &= s_5 s_5. \\ \theta(s_2) &= s_4 s_2 \end{aligned}$$

The sequence $\theta^\omega(s_0') = s_0', s_2, s_4, s_2, s_4, s_3, s_4, s_2, \ldots$ is a fixed point of θ. Let τ be the letter-to-letter projection defined by

$$\begin{aligned} \tau(s_0') = \tau(s_0) = \tau(s_1) = \tau(s_5) &= 0 \\ \tau(s_2) = \tau(s_4) &= 1 \\ \tau(s_3) &= 2. \end{aligned}$$

Then the sequence

$$\tau(\theta^\omega(s_0')) = 0, 1, 1, 1, 1, 2, 1, 1, 1, 2, 2, 0, 1, 2, 1, 1, \ldots$$

is the sequence produced by the automaton in Example 2.3 in direct reading.

2.2 The k-kernel characterization of an automatic sequence

Definition 2.7. The *k-kernel* of a sequence $a(n)_{n\geq 0}$ is the collection of sequences

$$\ker_k(a(n)_{n\geq 0}) := \{a(k^e n + j)_{n\geq 0} : e \geq 0,\ 0 \leq j \leq k^e - 1\}.$$

A proof of the following classical result can be found in [Eil74, Proposition V.3.3] and [AS03, Theorem 6.6.2].

Theorem 2.8. *Let $a(n)_{n\geq 0}$ be a sequence of elements from a finite alphabet $\mathcal{A}$. Then the k-kernel of $a(n)_{n\geq 0}$ is finite if and only if $a(n)_{n\geq 0}$ is k-automatic.*

If $a(n)_{n\geq 0}$ is k-automatic, then relationships between the elements of its k-kernel can be explicitly read off from any automaton that computes $a(n)_{n\geq 0}$ in reverse reading. If a sequence a is k-automatic, there is a minimal finite automaton $\mathcal{M}$ generating a in reverse reading. Formally $\mathcal{M}$ is *minimal* if it has the smallest number of states amongst all finite automata that generate a in reverse reading. Given an automatic sequence a, we can generate a minimal DFAO $\mathcal{M}_{\mathrm{R}}$ as follows. Each element in $\ker_k(a(n)_{n\geq 0})$ defines a state, and if $b(n)_{n\geq 0} \in \ker_k(a(n)_{n\geq 0})$ we define $\tau(b(n)_{n\geq 0}) = b(0)$ to be the projection onto the first term. Also, $\delta(b(n)_{n\geq 0}, i) := b(kn+i)_{n\geq 0}$ for $0 \leq i \leq k-1$. In some situations we will need to assume that we have a minimal $\mathcal{M}_{\mathrm{R}}$ that generates $a(n)_{n\geq 0}$.

2.3 Christol's Theorem

While Theorems 2.5 and 2.8 characterize k-automatic sequences for all $k \geq 2$, Christol's theorem characterizes p-automatic sequences for prime p. By taking a sufficiently large finite field of characteristic p, we may assume (by choosing an arbitrary embedding) that the output alphabet is a subset of this field.

Theorem 2.9 (Christol et al. [CKMFR80]). *Let $a(n)_{n\geq 0}$ be a sequence of elements in $\mathbb{F}_{p^\alpha}$. Then $\sum_{n\geq 0} a(n)x^n$ is algebraic over $\mathbb{F}_{p^\alpha}(x)$ if and only if $a(n)_{n\geq 0}$ is p-automatic.*

2.4 Dynamical systems generated by automatic sequences

We will consider two kinds of alphabets $\mathcal{A}$, each with a topology. If $\mathcal{A}$ is finite, we endow it with the discrete topology. If $\mathcal{A}$ is a closed subset of the p-adic integers $\mathbb{Z}_p$, it inherits the subspace topology from the standard compact topology on $\mathbb{Z}_p$, which comes from looking at $\mathbb{Z}_p$ as the countable product of the p-element set $\{0, 1, \ldots, p-1\}$ with the discrete topology. In both cases $\mathcal{A}^{\mathbb{N}}$, with the product topology, is a compact metric space. Let $\sigma : \mathcal{A}^{\mathbb{N}} \to \mathcal{A}^{\mathbb{N}}$ denote the *shift map* $\sigma(a(n)_{n\geq 0}) := a(n+1)_{n\geq 0}$. The shift map is a continuous mapping on $\mathcal{A}^{\mathbb{N}}$. If $a = a(n)_{n\geq 0} \in \mathcal{A}^{\mathbb{N}}$, define $X_a := \overline{\{\sigma^n(a) : n \in \mathbb{N}\}}$, the closure of the shift orbit of a in $\mathcal{A}^{\mathbb{N}}$. The dynamical system (X_a, σ) is called the (one-sided) subshift associated with a.

Given a natural number α, let $\pi_\alpha : \mathbb{Z}_p \to \mathbb{Z}/(p^\alpha\mathbb{Z})$ be the natural projection map. We will also use π_α to denote the projection map $\pi_\alpha : \mathbb{Z}/(p^{\alpha+1}\mathbb{Z}) \to \mathbb{Z}/(p^\alpha\mathbb{Z})$. We can compose these maps π_α to obtain projection maps $\pi_{\alpha,\beta} : \mathbb{Z}/(p^{\alpha+1}\mathbb{Z}) \to \mathbb{Z}/(p^\beta\mathbb{Z})$ for $\alpha \geq \beta \geq 1$ (where $\pi_{\alpha,\alpha} = \pi_\alpha$). These projection maps extend termwise to sequences. Suppose that $a \in \mathbb{Z}_p^{\mathbb{N}}$ and $a_\alpha := \pi_\alpha(a)$. We work in the category of symbolic dynamical systems, where $\Phi : (X, S) \to (Y, T)$ is a substitution if $\Phi : X \to Y$ is a continuous map such that $\Phi \circ S = T \circ \Phi$. The projection maps π_α, and $\pi_{\alpha,\beta}$ can be used to define an inverse limit. The proof of the following lemma is straightforward.

Lemma 2.10. *Let $a \in \mathbb{Z}_p^{\mathbb{N}}$. Then (X_a, σ) is the inverse limit of $(X_{a_\alpha}, \sigma)_{\alpha\geq 1}$.*

The next lemma gives us the commutative diagram in Figure 1.

Lemma 2.11. *Suppose that $a \in \mathbb{Z}_p^{\mathbb{N}}$ is such that for each $\alpha \geq 1$, a_α is p-automatic. Let $(\mathcal{S}_\alpha, \Sigma_p, \delta_\alpha, s_0, \mathbb{Z}/(p^\alpha\mathbb{Z}), \tau_\alpha)$ be the minimal automaton generating a_α in reverse reading. Let u_α be the sequence generated in reverse reading by $\mathcal{M}_\alpha = (\mathcal{S}_\alpha, \Sigma_p, \delta_\alpha, s_0)$. Then there are projection maps $\pi^*_\alpha : X_{u_{\alpha+1}} \to X_{u_\alpha}$ such that $\pi^*_\alpha(u_{\alpha+1}) = u_\alpha$, $\pi_\alpha \circ \sigma = \sigma \circ \pi_\alpha$ and $\tau_\alpha \circ \pi^*_\alpha = \pi_\alpha \circ \tau_{\alpha+1}$.*

$$
\begin{array}{ccccccccccc}
\cdots & \xleftarrow{\pi^*_{\alpha-1}} & (X_{u_\alpha},\sigma) & \xleftarrow{\pi^*_{\alpha}} & (X_{u_{\alpha+1}},\sigma) & \xleftarrow{\pi^*_{\alpha+1}} & (X_{u_{\alpha+2}},\sigma) & \xleftarrow{\pi^*_{\alpha+2}} & \cdots & & \\
 & & \downarrow\tau_\alpha & & \downarrow\tau_{\alpha+1} & & \downarrow\tau_{\alpha+2} & & & & \\
\cdots & \xleftarrow{\pi_{\alpha-1}} & (X_{a_\alpha},\sigma) & \xleftarrow{\pi_{\alpha}} & (X_{a_{\alpha+1}},\sigma) & \xleftarrow{\pi_{\alpha+1}} & (X_{a_{\alpha+2}},\sigma) & \xleftarrow{\pi_{\alpha+2}} & \cdots & & (X_a,\sigma)
\end{array}
$$

Figure 1: The commuting diagram according to Lemma 2.11.

Proof. It is important that we work with a minimal automaton that generates a_α in reverse reading. Note that this implies that $\mathcal{M}_\alpha$ is minimal, since $|\ker_p(a_\alpha)| = |\ker_p(u_\alpha)|$. What we shall show is that there is a graph epimorphism $\pi^*_\alpha : \mathcal{M}_{\alpha+1} \to \mathcal{M}_\alpha$ which maps the initial state of $\mathcal{M}_{\alpha+1}$ to the initial state of $\mathcal{M}_\alpha$. Define π^*_α to map s_0 to s_0 and $\delta_{\alpha+1}(s_0, i)$ to $\delta_\alpha(s_0, i)$. We extend this map recursively. If $\delta_{\alpha+1}(s, n_\ell \cdots n_0) = \delta_{\alpha+1}(s, m_{\ell'} \cdots m_0)$, this means then that $n_\ell \cdots n_0$ and $m_{\ell'} \cdots m_0$ define the same sequence $v(n)_{n\geq 0}$ in $\ker_p(u_{\alpha+1})$. The sequence $(v(n) \bmod p^\alpha)_{n\geq 0}$ belongs to $\ker_p(u_\alpha)$, and since $(\mathcal{S}_\alpha, \Sigma_p, \delta_\alpha, s_0)$ is minimal, then $\delta_\alpha(\pi^*_\alpha(s), n_\ell \cdots n_0) = \delta_\alpha(\pi^*_\alpha(s), m_{\ell'} \cdots m_0)$. In other words, π^*_α is well-defined.

By definition, $\pi^*_\alpha(u_{\alpha+1}(n)) = \pi^*_\alpha(\delta(s_0, (n)_p)) = u_\alpha(n)$. Thus the map $\pi^*_\alpha : \mathcal{S}_{\alpha+1} \to \mathcal{S}_\alpha$ induces a map on $\pi^*_\alpha : \mathcal{S}^{\mathbb{N}}_{\alpha+1} \to \mathcal{S}^{\mathbb{N}}_\alpha$ such that $\pi^*_\alpha(u_{\alpha+1}) = u_\alpha$. From this it follows that we can extend π^*_α to a shift-commuting map $\pi^*_\alpha : (X_{u_{\alpha+1}}, \sigma) \to (X_{u_\alpha}, \sigma)$, and that $\tau_\alpha \circ \pi^*_\alpha = \pi_\alpha \circ \tau_{\alpha+1}$ on $X_{u_{\alpha+1}}$. □

Since each u_α is a k-automatic sequence where the letter-to-letter projection is the identity map, the sequence u_α is a fixed point of a length-p substitution θ_α. The next lemma tells us that the projection map π^*_α also commutes with the substitution θ_α.

Lemma 2.12. *For each $\alpha \geq 1$, $\theta_\alpha \circ \pi^*_\alpha = \pi^*_\alpha \circ \theta_{\alpha+1}$ on $X_{u_{\alpha+1}}$.*

Proof. The claim is true for the $\theta_{\alpha+1}$-fixed point $u_{\alpha+1}$. On the shift-orbit of $u_{\alpha+1}$, we have

$$
\begin{aligned}
\theta_\alpha \circ \pi^*_\alpha(\sigma^k u_{\alpha+1}) &= \theta_\alpha \circ \sigma^k(\pi^*_\alpha u_{\alpha+1}) = \sigma^{kp}(\pi^*_\alpha u_{\alpha+1}) \\
&= \pi^*_\alpha(\sigma^{kp}(u_{\alpha+1})) = \pi^*_\alpha \circ \theta_{\alpha+1}(\sigma^k u_{\alpha+1}) \,.
\end{aligned}
$$

Since both θ_α and π^*_α are continuous, we can extend this identity to $X_{u_{\alpha+1}}$. □

Figure 2 summarizes the content of Lemma 2.12.

Remark 2.13. Note that Lemmas 2.11 and 2.12 would both follow from the existence of the inverse limit of the systems (X_{u_α}, σ), although we will not focus on this approach.

$$
\begin{array}{ccccccccc}
\cdots & \xleftarrow{\pi^*_{\alpha-1}} & X_{u_\alpha} & \xleftarrow{\pi^*_{\alpha}} & X_{u_{\alpha+1}} & \xleftarrow{\pi^*_{\alpha+1}} & X_{u_{\alpha+2}} & \xleftarrow{\pi^*_{\alpha+2}} & \cdots \\
 & & \downarrow\theta_\alpha & & \downarrow\theta_{\alpha+1} & & \downarrow\theta_{\alpha+2} & & \\
\cdots & \xleftarrow{\pi^*_{\alpha-1}} & X_{u_\alpha} & \xleftarrow{\pi^*_{\alpha}} & X_{u_{\alpha+1}} & \xleftarrow{\pi^*_{\alpha+1}} & X_{u_{\alpha+2}} & \xleftarrow{\pi^*_{\alpha+2}} & \cdots
\end{array}
$$

Figure 2: The commuting diagram according to Lemma 2.12.

We now want to define a subshift (X_u, σ), which projects via π^*_α to (X_{u_α}, σ), and a map $\tau : (X_u, \sigma) \to (X_a, \sigma)$ such that when these two items are added to Figure 1 as follows, the diagram still commutes.

$$
\begin{array}{ccccccccccc}
\cdots & \xleftarrow{\pi^*_{\alpha-1}} & (X_{u_\alpha}, \sigma) & \xleftarrow{\pi^*_{\alpha}} & (X_{u_{\alpha+1}}, \sigma) & \xleftarrow{\pi^*_{\alpha+1}} & (X_{u_{\alpha+2}}, \sigma) & \xleftarrow{\pi^*_{\alpha+2}} & \cdots & (X_u, \sigma) \\
 & & \downarrow\tau_\alpha & & \downarrow\tau_{\alpha+1} & & \downarrow\tau_{\alpha+2} & & & \downarrow\tau \\
\cdots & \xleftarrow{\pi_{\alpha-1}} & (X_{a_\alpha}, \sigma) & \xleftarrow{\pi_{\alpha}} & (X_{a_{\alpha+1}}, \sigma) & \xleftarrow{\pi_{\alpha+1}} & (X_{a_{\alpha+2}}, \sigma) & \xleftarrow{\pi_{\alpha+2}} & \cdots & (X_a, \sigma)
\end{array}
$$

We will use the following general notation. Suppose we have spaces X_α and maps $\pi_\alpha : X_{\alpha+1} \to X_\alpha$. If $x = (x_\alpha)_{\alpha \geq 1}$ with $x_\alpha \in X_\alpha$ and $\pi_\alpha(x_{\alpha+1}) = x_\alpha$ for each α, we shall write $x = \varprojlim x_\alpha$. Define

$$X_u := \left\{ x = \varprojlim x_\alpha : x_\alpha \in X_{u_\alpha} \text{ for each } \alpha \text{ and } \pi^*_\alpha(x_{\alpha+1}) = x_\alpha \text{ for each } \alpha \geq 1 \right\}.$$

Note that X_u lives in $\prod_{j=1}^{\infty} \mathcal{S}_\alpha^{\mathbb{N}}$, which can be endowed with the (metrizable) product topology when each of $\mathcal{S}_\alpha^{\mathbb{N}}$ is regarded as the previously described metric space. Define $u := \varprojlim u_\alpha \in X_u$.

Lemma 2.14. *X_u is closed in $\prod_{j=1}^{\infty} \mathcal{S}_\alpha^{\mathbb{N}}$ and $X_u = \overline{\{\sigma^n(x_u)\}}$. The sequence of maps $(\tau_\alpha)_{\alpha \geq 0}$ induce a projection map $\tau : X_u \to X_a$ such that $\tau \circ \sigma = \sigma \circ \tau$.*

Proof. Suppose that $x(n)_{n \geq 0}$ is a sequence in X_u and $x(n) \to x$. Let $x(n) = \varprojlim_\alpha x_\alpha(n)$ and $x = \varprojlim x_\alpha$. For each α, $x_\alpha(n) \to x_\alpha$ as $n \to \infty$. Since X_{u_α} is closed, $x_\alpha \in X_{u_\alpha}$. This shows that X_u is closed. We have just seen that each x_α is in the shift-orbit closure of u_α. Note that since π^*_α commutes with the shift, if $\sigma^{n_k^{(\alpha+1)}}(u_{\alpha+1}) \to x_{\alpha+1}$, then $\sigma^{n_k^{(\alpha+1)}}(u_\alpha) \to x_\alpha$. Now a diagonal argument on the sequences $\{(n_k^\alpha)_{k \geq 0} : \alpha \geq 1\}$ will give us a sequence n_k such that $d(\sigma^{n_k} u, x) \to 0$. The existence of the shift-commuting τ is guaranteed by the commutativity of Figure 1. $\square$

2.5 Substitutions on infinite alphabets

The commutative diagram of Figure 2 will allow us to define a substitution θ on X_u, for which u is a fixed point. We can do this in two equivalent ways: either by defining θ on X_u, or by defining a substitution on the appropriate alphabet. We investigate both approaches. Note that if $x = \varprojlim x_\alpha$ is an element of X_u, then

Lemma 2.12 tells us that $y = \varprojlim \theta_\alpha(x_\alpha)$ is also an element of X_u. Thus the map $\theta : X_u \to X_u$ given by

$$\theta(\varprojlim x_\alpha) := \varprojlim \theta_\alpha(x_\alpha)$$

is well-defined. Note also that since $u = \varprojlim u_\alpha$ and $\theta_\alpha(u_\alpha) = u_\alpha$ for each α, then $\theta(u) = u$. Finally we observe that

$$\theta(\sigma(\varprojlim x_\alpha)) = \theta(\varprojlim \sigma(x_\alpha)) = \varprojlim \theta_\alpha \sigma(x_\alpha) = \varprojlim \sigma^p \theta_\alpha(x_\alpha) = \sigma^p \theta(\varprojlim(x_\alpha)),$$

further suggesting that θ is defined by a substitution. To see this we will define the alphabet on which θ will act. Let

$$\mathcal{S} := \left\{ s = \varprojlim s_\alpha : s_\alpha \in \mathcal{S}_\alpha \text{ and } \pi_{\alpha^*}(s_{\alpha+1}) = s_\alpha \text{ for each } \alpha \geq 1 \right\}.$$

Note that by the definition of the projection maps π^*_α in Lemma 2.11, we have $\pi^*_\alpha : \mathcal{S}_{\alpha+1} \to \mathcal{S}_\alpha$ and $\pi^*_\alpha(\theta_{\alpha+1}(s)) = \theta_\alpha(\pi^*_\alpha(s))$, i.e. Figure 2 is also valid on the level of the alphabets $(\mathcal{S}_\alpha)_{\alpha\geq 1}$. So we can define, for $s = (s_\alpha)_{\alpha \geq 1} \in \mathcal{S}$,

$$\theta(s) := \varprojlim \theta_\alpha(s_\alpha).$$

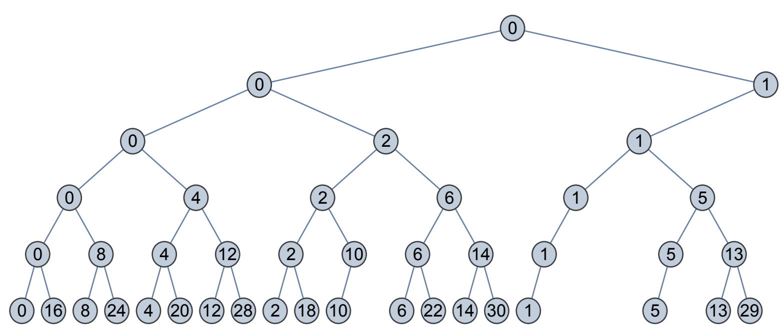
Note that for each $s \in \mathcal{S}$, $\theta(s)$ is a word of length p in $\mathcal{S}^+$. As in Definition 2.4, the substitution $\theta : \mathcal{S} \to \mathcal{S}^p$ extends to the substitution $\theta : X_u \to X_u$.

Now $\mathcal{S}$ projects, via τ, to a subset of $\mathbb{Z}_p$. The following set-up describes this subset of $\mathbb{Z}_p$ using a tree.

Definition 2.15. Let p be prime, and let $a = a(n)_{n\geq 0}$ be a sequence of p-adic integers. We define T_a, the *tree associated to* a. This will be a directed subtree of the infinite complete p-ary tree C_p, as follows. From each vertex of the complete p-ary tree, there are p edges labelled 0 to $p-1$. The level-α vertices of T_a will be all vertices v in C_p such that there is a path with edges labelled $n_0, \ldots, n_{\alpha-1}$ from the root to v at level α, whenever there is some j with $(a(j) \bmod p^\alpha)_p = n_{\alpha-1} \cdots n_1 n_0$.

Example 2.16. If $a(n) = n$, then T_a is the complete p-ary tree.

Example 2.17. Let $a(n) = \frac{1}{n+1}\binom{2n}{n}$ be the nth Catalan number. The sequence $a(n)_{n\geq 0} = 1, 1, 2, 5, 14, 42, 132, 429, \ldots$ is ubiquitous in combinatorial settings. Levels 0 through 5 of the tree T_a for $p = 2$ are as follows.



The sequence $0, (a(n) \bmod 4)_{n\geq 1}$ is produced by the automaton in Example 2.3 in reverse reading. In particular, $a(n) \not\equiv 3 \mod 4$ for all $n \geq 0$, as shown by Eu, Liu, and Yeh [ELY08]. Therefore T_a has only 3 vertices on level $\alpha = 2$. Liu and Yeh [LY10] showed that $a(n) \not\equiv 9 \mod 16$ for all $n \geq 0$, so T_a has only 11 vertices on level $\alpha = 4$. The forbidden residues modulo 2^α for $5 \leq \alpha \leq 9$ were identified by the authors [RY14, Section 3.1]. The behavior as $\alpha \to \infty$ of the number of vertices on level α in T_a is not known.

Given an infinite rooted tree T, let $\mathcal{P}_T$ be the set of all infinite paths in T which begin at the root and visit each vertex at most once. We can define a metric on $\mathcal{P}_T$: two paths x and y are $\frac{1}{p^\alpha}$ close if they agree on the finite initial segment from the root to level α. The following lemma can be stated more generally, for a sequence of n-adic integers that do not necessarily project to automatic sequences.

Lemma 2.18. *Let $a = a(n)_{n\geq 0}$ be a sequence of p-adic integers, and let T_a be the tree associated to a. Then there exists a continuous surjective map $f : X_a \to \mathcal{P}_{T_a}$.*

Proof. Take $x = x(n)_{n\geq 0} \in X_a$ and project it to $x(0) \in \mathbb{Z}_p$. Since there is a sequence $\sigma^{n_k}(a)$ converging to x, then $a(n_k) \to x(0)$ as $k \to \infty$. Thus for each α there exists a $K = K(\alpha)$ such that if $k > K(\alpha)$ then $a(n_k) \equiv x(0) \mod p^\alpha$; this means that for any α we can associate $x(0)$ to a unique path from the root of T_a to level α of T_a, which is extended to a unique path from the root of T_a to level $\alpha+1$ of T_a. Now let $\alpha \to \infty$: this gives us the desired infinite path that we shall denote by $f(x)$. If x is close to y in X_a, then $x(0)$ is close to $y(0)$ in $\mathbb{Z}_p$, which means that for some large α, $x(0) \equiv y(0) \mod p^\alpha$. Thus $f(x)$ and $f(y)$ agree on the first α levels of T_a, i.e. f is continuous.

To show that f is surjective, take a path P in $\mathcal{P}_{T_a}$; let P_α denote the truncation of P to the finite path ending at level α. By definition, there is an n_α such that $f(\sigma^{n_\alpha}(a))$ agrees with P_α. By compactness of X_a, we can take a convergent subsequence of $(\sigma^{n_\alpha}(a))_\alpha$, a subsequence which converges to x. We now see that $f(x) = P$. □

2.6 The inverse limits of the machines $\mathcal{M}_\alpha$

Suppose that $a = a(n)_{n\geq 0} \in \mathbb{Z}_p^n$ is such that for each $\alpha \geq 1$, $(a(n) \bmod p^\alpha)_{n\geq 0}$ is p-automatic and generated in reverse reading by the machine $\mathcal{M}_\alpha = (\mathcal{S}_\alpha, \Sigma_p, \delta_\alpha, s_0, \mathbb{Z}/(p^\alpha\mathbb{Z}), \tau_\alpha)$. As in Lemma 2.11, we assume that each of these machines is minimal. We now describe the object that serves as the "inverse limit" of the machines $(\mathcal{S}_\alpha, \Sigma_p, \delta_\alpha, s_0, \mathbb{Z}/(p^\alpha\mathbb{Z}), \tau_\alpha)$.

Let C_p be the complete p-ary tree with root s_0. We label the p edges emanating from each vertex 0 to $p-1$ (from left to right). We label the vertices as follows. For a vertex v, if the path from the root s_0 to v is labelled $n_0 n_1 \cdots n_k$ where $n_k \neq 0$ and

$(n)_p = n_0 n_1 \cdots n_k$, then we label v with $a(n)$. If the edge pointing to v has label 0, then label v by the label of its parent. If $\mathcal{M}_\alpha = (V_\alpha, E_\alpha)$, and $\mathcal{M} = (V, E)$ is C_p, but labelled as just described, we can define $\pi_\alpha : V \to V_\alpha$ as follows. The root in V is sent to the initial state in $\mathcal{M}_\alpha$. For a non-root vertex $v \in V$, if $\pi_\alpha(w)$ has already been defined and there is an edge with source w and range v, labelled i, define $\pi_\alpha(v) = \delta_\alpha(\pi_\alpha(w), i)$. Note that we can also define maps $\pi_\alpha : V_{j+1} \to V_j$ in the same way. Since the machines are minimal, the map $\pi_\alpha : V_{j+1} \to V_j$ is well-defined (see the proof of Lemma 2.11).

3 Examples

In this section we give examples of integer sequences, arising in two quite different settings, that produce the subshifts that we have defined. The first examples are algebraic sequences, which produce subshifts living in $\mathbb{Z}_p^{\mathbb{N}}$ for any p. The second set of examples define a subshift living in $\mathbb{Z}_p^{\mathbb{N}}$ for only on p.

3.1 Algebraic sequences

We consider some famous algebraic sequences of integers, and the substitutions they induce. More generally, sequences that are diagonals of certain rational power series can similarly be used to define substitutions.

If a substitution θ on a finite alphabet defines the subshift (X, σ), then the set $\mathcal{L} = \bigcap_{n\geq 0} \theta^n(X)$ often contains only the periodic points of θ, of which there are finitely many. When we consider substitutions on an uncountable alphabet, $\mathcal{L}$ can possibly be much larger.

1. Let $a(n) = n$. The generating function $y = \sum_{n\geq 0} nx^n$ satisfies $(1-x)^2 y - x = 0$; it is rational and hence algebraic. Let $p = 2$. The sequence $a(n)_{n\geq 0}$ is the fixed point of the substitution $\theta(m) = (2m)\,(2m+1)$. We claim that (X_a, σ) is topologically conjugate to $(\mathbb{Z}_2, +1)$. The conjugacy $\varphi : (X_a, \sigma) \to (\mathbb{Z}_2, +1)$ will be defined on the σ-orbit of a as $\varphi(\sigma^n(a)) := n$. If φ and φ^{-1} are uniformly continuous and bijective on $\{\sigma^n(a) : n \geq 0\}$, then φ extends to a conjugacy between (X_a, σ) and $(\mathbb{Z}_2, +1)$. Since φ is just the projection map to the first term, φ is uniformly continuous, and it is clearly a bijection between the σ-orbit of a and $\mathbb{N}$. Note that a is the unique point in $\bigcap_{n\geq 0} \theta^n(X_a)$. For every other prime p we have an analogous situation.

2. The Fibonacci sequence $F(n)_{n\geq 0} = 0, 1, 1, 2, 3, 5, 8, 13, \ldots$ satisfies a linear recurrence with constant coefficients and hence has a rational generating function. For every prime p, the sequence $F(n)_{n\geq 0}$ is the coding of fixed point of a length-p substitution. Note that because of the repeated 1, the letter-to-letter

coding τ for which $F = \tau(u)$ is not the identity map. For example, let $p = 2$, consider the alphabet $\mathcal{S} = \mathbb{N} \cup \{s\}$, and define θ on $\mathcal{S}^*$ by

$$\begin{aligned} \theta(0) &= 0\, s \\ \theta(s) &= 1\, 2 \\ \theta(F(m)) &= F(2m)\, F(2m+1) \quad \text{for } m \geq 2. \end{aligned}$$

Then $u = 0, s, 1, 2, 3, 5, 8, 13, \dots$ is a fixed point of θ. Letting $\tau(s) = 1$ and $\tau(m) = m$ for all $m \in \mathbb{N}$ gives $F = \tau(u)$. The following graphic shows the hundred least significant binary digits of $F(2^n)$ for each $0 \leq n \leq 20$, where 0 is represented by a white cell and 1 is represented by a black cell.

Therefore $F(2^n)_{n\geq 0}$ appears to converge to a periodic orbit in the image of $\mathcal{L} = \bigcap_{n\geq 0} \theta^n(X_u)$ under τ.

3. Let $a(n)_{n\geq 0}$ be the sequence of Catalan numbers, mentioned in Example 2.17. The analogous graphic, showing binary digits of $a(2^n)$ for $0 \leq n \leq 20$, suggests that $a(2^n)_{n\geq 0}$ converges in $\mathbb{Z}_2$.

Let $\phi : \mathbb{Z}_2^{\mathbb{N}} \to \mathbb{Z}_2$ denote projection to the first term. Since

$$a(2^n) = \phi\sigma^{2^n}(a) = \phi\sigma^{2^n}(\tau(u)) = \phi\tau\sigma^{2^n}(u) = \phi\tau\theta^n(\sigma(u)),$$

and since any limit point of $(\theta^n(\sigma(u)))_{n\geq 0}$ belongs to $\mathcal{L} = \bigcap_{j=1}^{\infty} \theta^j(X_u)$, this suggests that $a(2^n)_{n\geq 0}$ converges to the projection of an element in $\mathcal{L}$. Other than θ-fixed and periodic points, what else lies in $\mathcal{L}$?

3.2 Cocycle sequences

Let

$$M_p := \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$$

be the p by p matrix all of whose entries are 1. This matrix is the incidence matrix for the substitution θ^* on the alphabet $\mathbb{Z}/(p\mathbb{Z})$ defined as $\theta^*(j) = 01\cdots(p-1)$ for each $j \in \mathbb{Z}/(p\mathbb{Z})$, whose fixed point is periodic.

Let θ be a substitution on $\mathbb{Z}/(p\mathbb{Z})$ whose incidence matrix is also M_p. Assume that θ is *aperiodic*: i.e. that it has a fixed point which is not periodic. (In the case where θ has the incidence matrix M_p, if it is aperiodic, then none of its fixed points are periodic.) Let us assume also that $\theta(0)$ starts with 0, and let $u = 0\cdots$ be the fixed point starting with 0. There is a (unique) measure μ such that (X_u, σ, μ) is a measure-preserving dynamical system. This measurable system has an *adic* representation (X_B, φ_θ) which we now describe. Here B is an infinite graph called a *Bratteli diagram*. We illustrate B in Figure 3 in the case where $p = 3$.

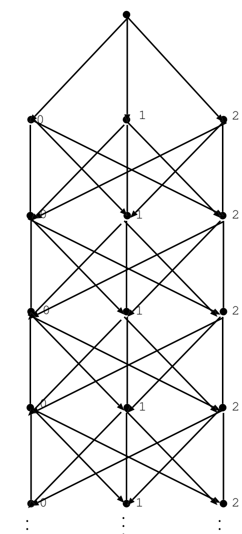


Figure 3: The Bratteli diagram associated to M_3.

Apart from the "root" vertex at the top of the diagram, there are p vertices at each level n, which we label $0, \ldots, p-1$, moving from left to right. The levels are indexed by increasing indices n as we move down in the diagram, $n = 0, 1, \ldots$; we do not think of the root vertex as occupying a level. We think of an edge as having its source at level $n-1$ and range level n for some $n \geq 1$. If θ is a substitution with incidence matrix M_p, it defines a linear order on the incoming edges to any vertex: if the vertex is labelled j and $\theta(j) := i_0 \cdots i_{p-1}$, then we give each edge with source i_k the label "k". Thus for any vertex v, the set of edges with range v are linearly ordered. Let X_B be the set of infinite paths in B. Such a path is labelled $x = x_0, x_1, \ldots$ where the edge x_i starts at vertex v_i and has range vertex v_{i+1}. The order placed on each set of edges with common range induces a partial order on X_B. In particular we can compare two infinite paths x and x' in B which eventually agree: if n is the smallest integer such that x and x' agree from level n onwards, so that the edges x_{n-1} and x'_{n-1} (between levels $n-1$ and levels n) have the same range at level n, and $x_{n-1} < x'_{n-1}$, then we write $x < x'$. If x is a non-maximal path (i.e. one if its edges is not maximally labelled) then it has a *successor* in this order. Namely, if n is the smallest integer such that x_n is not a maximal edge, then the successor of x agrees with x from level $n+1$, is minimal up to level n, and the edge between level n and level $n+1$ is the successor of x_n.

Thus any substitution θ with incidence matrix M_p defines a partial ordering of X_B and this determines an *adic* map $\varphi_\theta : X_B \to X_B$ where $\varphi_\theta(x)$ is defined to be the successor of x in the ordering determined by θ. Note that φ is not defined on the set of maximal paths, but this is a finite set (given measure 0 by the unique measure which is invariant under the tail equivalence relation) and if we needed to, we can arbitrarily define φ on this set, sending maximal paths to minimal paths.

If θ is not a periodic substitution, then (X_B, φ_θ) is measurably conjugate to (X_u, σ), with the minimal path through the vertices u_0 being mapped to u. If θ is a periodic substitution (as θ^* is), then (X_B, φ_θ) is (both topologically and measurably) conjugate to the p-adic odometer $(\mathbb{Z}_p, +1)$. In this latter case, if the finite path x has edges labelled by the base-p expansion of m, then $\varphi^n_{\theta^*}(x)$ is the finite path whose edges are labelled by the base-p expansion of $m+n$. We refer the reader to [VS08] and [DHS99] for the measurable and topological versions of these results.

Note that the ordering induced by θ^* on B has the special property that an edge labelled i has as source a vertex labelled i.

Suppose that θ is aperiodic and $\theta(0)$ starts with 0. Let 0^∞ denote the minimal path in B that runs through the vertices labelled 0, and let the sequence $s(n)_{n\geq 0}$ of natural numbers be defined by

$$\varphi^n_\theta(0^\infty) = \varphi^{s(n)}_{\theta^*}(0^\infty).$$

The sequence $s(n)_{n\geq 0}$ is called a *cocycle.*

Theorem 3.1. *For any $\alpha \geq 0$, $(s(n) \bmod p^\alpha)_{n\geq 0}$ is the fixed point of a length-p substitution θ_α.*

Proof. Note that if the finite path $x_0x_1\cdots x_k$ passes through the vertices $v_0v_1\cdots v_{k+1}$, and if $x_0x_1\cdots x_k = \varphi^n_\theta(00\cdots 0)$, then $s(n) = \sum_{j=0}^{k+1} p^j v_j$, and $s(n) \bmod p^\alpha = \sum_{j=0}^{\alpha-1} p^j v_j$. We use the $[\theta(a), j]$ to denote the j-th letter of $\theta(a)$.

Given $\alpha \geq 0$, we define a substitution θ_α on $\mathbb{Z}/(p^\alpha\mathbb{Z})$ of length p as follows. Given $j = j_0p^0 + j_1p^1 + \cdots + j_{\alpha-1}p^{\alpha-1} \in \mathbb{Z}/(p^\alpha\mathbb{Z})$, define $\theta_\alpha(j) = \theta(j_0) + p(j_0p^0 + j_1p^1 + \cdots + j_{\alpha-2}p^{\alpha-2})$, where here we are adding $p(j_0p^0 + j_1p^1 + \cdots + j_{\alpha-2}p^{\alpha-2})$ to each entry in the word $\theta(j_0)$. We claim that $(s(n) \bmod p^\alpha)_{n\geq 0}$ is a fixed point of θ_α.

To see this, we need to show that for each n, $\theta_\alpha(s(n) \bmod p^\alpha) = (s(pn), s(pn+1), \ldots, s(pn+p-1)) \bmod p^\alpha$. To get $s(pn+\ell) \bmod p^\alpha$, we need the first α vertices through which the path $\varphi_\theta^{pn+\ell}(00\cdots 0)$ runs. Suppose that the path $\varphi_\theta^{n}(00\cdots 0)$ passes through the vertices $v_0, v_1, \ldots, v_{\alpha-1}$, so that $s(n) \bmod p^\alpha = \sum_{j=0}^{\alpha-1} v_j p^j$. Then the path $\varphi_\theta^{pn+\ell}(00\cdots 0)$ starts at the vertex labelled $[\theta(v_0), \ell]$, followed by $v_0, \ldots, v_{\alpha-2}$ at levels $1, \ldots, \alpha-1$ of the diagram respectively. In other words, $s(pn+\ell) \bmod p^\alpha = [\theta(v_0), \ell] + p\sum_{j=0}^{\alpha-2} v_j p^j = [\theta_\alpha(s(n) \bmod p^\alpha), \ell]$, as desired. $\square$

Remark 3.2. Since cocycle sequences are bijections of $\mathbb{N}$, it is very easy to define the cocycle sequence as the fixed point of a length-p substitution on $\mathbb{N}$. For example, if $p = 2$ and $\phi(0) = 01$, $\phi(1) = 10$ is the Thue–Morse substitution, then it has as transition matrix M_2 and its cocycle sequence

$$s(n)_{n\geq 0} = 0, 1, 3, 2, 7, 6, 4, 5, 15, 14, 12, 13, 8, 9, 11, 10, \ldots$$

is the fixed point of the length-2 substitution θ on $\mathbb{N}$ defined by

$$\theta(m) = \begin{cases} (2m)\,(2m+1) & \text{if } m \text{ is even} \\ (2m+1)\,(2m) & \text{if } m \text{ is odd.} \end{cases}$$

In particular, $s(n)_{n\geq 0}$ projects modulo 2 to the Thue–Morse sequence. However, it can be shown that $s(n)_{n\geq 0}$ is 2-regular in the sense of Allouche and Shallit [AS92]; namely, we have the recurrence

$$\begin{aligned} s(4n) &= -2s(n) + 3s(2n) \\ s(4n+1) &= -2s(n) + 2s(2n) + s(2n+1) \\ s(4n+2) &= -2s(n) + 3s(2n+1) \\ s(4n+3) &= -2s(n) + s(2n) + 2s(2n+1). \end{aligned}$$

It follows that $(s(n) \bmod k)_{n\geq 0}$ is 2-automatic for every $k \geq 2$ [AS92, Corollary 2.4]. Therefore, by a theorem of Cobham, for a prime $p \neq 2$ the sequence $(s(n) \bmod p^\alpha)_{n\geq 0}$ is not p-automatic unless it is eventually periodic. Moreover, the generating function $\sum_{n\geq 0} s(n)x^n$ is not rational, so it follows from a result of Bézivin [Béz94, BCR13] that $s(n)_{n\geq 0}$ is not algebraic (nor is it the diagonal of a rational function), so it does not fall within the scope of Section 3.1.

Acknowledgments

The authors acknowledge the hospitality and support of LaCIM, Montréal.

References

[AS92] Jean-Paul Allouche and Jeffrey Shallit. The ring of k-regular sequences. *Theoretical Computer Science*, 98(2):163–197, 1992.

[AS03] Jean-Paul Allouche and Jeffrey Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, Cambridge, 2003.

[BCR13] Jason P. Bell, Michael Coons, and Eric Rowland. The rational–transcendental dichotomy of Mahler functions. *Journal of Integer Sequences*, 16:Article 13.2.10, 2013.

[Béz94] Jean-Paul Bézivin. Sur une classe d'equations fonctionnelles non linéaires. *Funkcialaj Ekvacioj*, 37:263–271, 1994.

[Chr74] Gilles Christol. Eléments analytiques uniformes et multiformes. *Séminaire Delange-Pisot-Poitou. Théorie des nombres*, 1(15):1–18, 1973-1974.

[CKMFR80] Gilles Christol, Teturo Kamae, Michel Mendès France, and Gérard Rauzy. Suites algébriques, automates et substitutions. *Bull. Soc. Math. France*, 108(4):401–419, 1980.

[Cob72] Alan Cobham. Uniform tag sequences. *Math. Systems Theory*, 6:164–192, 1972.

[DHS99] F. Durand, B. Host, and C. Skau. Substitutional dynamical systems, Bratteli diagrams and dimension groups. *Ergodic Theory Dynam. Systems*, 19(4):953–993, 1999.

[DL87] Jan Denef and Leonard Lipshitz. Algebraic power series and diagonals. *J. Number Theory*, 26(1):46–67, 1987.

[Eil74] Samuel Eilenberg. *Automata, languages, and machines. Vol. A*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York, 1974. Pure and Applied Mathematics, Vol. 58.

[ELY08] Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh. Catalan and Motzkin numbers modulo 4 and 8. *European Journal of Combinatorics*, 29(6):1449–1466, 2008.

[Fer06] Sébastien Ferenczi. Substitution dynamical systems on infinite alphabets. *Ann. Inst. Fourier (Grenoble)*, 56(7):2315–2342, 2006.

[GPS09] Mathieu Guay-Paquet and Jeffrey Shallit. Avoiding squares and overlaps over the natural numbers. *Discrete Mathematics*, 309:6245–6254, 2009.

[LY10] Shu-Chung Liu and Jean C.-C. Yeh. Catalan numbers modulo 2^k. *Journal of Integer Sequences*, 13:10.5.4, 2010.

[Mau06] Christian Mauduit. Propriétés arithmétiques des substitutions et automates infinis. *Ann. Inst. Fourier (Grenoble)*, 56(7):2525–2549, 2006.

[RS12] Eric Rowland and Jeffrey Shallit. Avoiding 3/2-powers over the natural numbers. *Discrete Mathematics*, 312:1282–1288, 2012.

[RY14] Eric Rowland and Reem Yassawi. Automatic congruences for diagonals of rational functions. *Journal de Théorie des Nombres de Bordeaux*, to appear, 2014. `http://arxiv.org/abs/1310.8635`.

[VS08] A. M. Vershik and B. Solomyak. The adic realization of the Morse transformation and the extension of its action to the solenoid. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 360(Teoriya Predstavlenii, Dinamicheskie Sitemy, Kombinatornye Metody. XVI):70–90, 296, 2008.

[Wes96] Julian West. Generating trees and forbidden subsequences. *Discrete Math.*, 157(1-3):363–374, 1996.