Chaos, Solitons and Fractals 99 (2017) 263–269

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

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

Morphisms on infinite alphabets, countable states automata and regular sequences[☆]

Jie-Meng Zhang[a], Jin Chen[b], Ying-Jun Guo[b,*], Zhi-Xiong Wen[a]

[a] School of Science, Wuhan Institute of Technology, Wuhan, 430205, PR China
[b] Institute of Applied Mathematics, College of Science, Huazhong Agricultural University, Wuhan, 430070, PR China

ARTICLE INFO

Article history:
Received 12 October 2016
Revised 8 April 2017
Accepted 8 April 2017

MSC:
Primary: 11B85
Secondary: 68R15
11B37

Keywords:
Automatic sequence
Regular sequence
Morphism
Countable states automaton
Linear recurrence

ABSTRACT

In this paper, we prove that a class of regular sequences can be viewed as projections of fixed points of uniform morphisms on a countable alphabet, and also can be generated by countable states automata. Moreover, we prove that the regularity of some regular sequences is invariant under some codings.

1. Introduction

Morphisms on a finite alphabet are widely studied in many fields, such as finite automata, symbolic dynamics, formal languages, number theory and also in physics in relation to quasicrystals (see [1,4,8,11,15]). A morphism is a map $\sigma\colon \Sigma^* \to \Sigma^*$ satisfying $\sigma(uv) = \sigma(u)\sigma(v)$ for all words $u, v \in \Sigma^*$, where Σ^* is the free moniod generated by a finite alphabet Σ (with ϵ as the neutral element). Naturally, the morphism can be extended to $\Sigma^{\mathbb{N}}$, which is the set of infinite sequences. The morphisms of constant length are called *uniform morphisms* and the sequence $u = u(0)u(1)u(2)\cdots \in \Sigma^{\mathbb{N}}$ satisfying $\sigma(u) = u$ is a *fixed point* of σ.

In [12], Cobham showed that a sequence is a fixed point of a uniform morphism (under a coding) if and only if it is an automatic sequence. Recall that a sequence is *automatic* if it can be generated by a finite state automaton. Moreover, a sequence $\{u(n)\}_{n\geq 0}$ is k-automatic if and only if its k-kernel is finite, where the *k-kernel* is defined by the set of subsequences,

$$\left\{\{u(k^i n + j)\}_{n\geq 0} : i \geq 0, 0 \leq j \leq k^i - 1\right\}.$$

However, the range of automatic sequences is necessarily finite. To overcome this limit, Allouche and Shallit [2] introduced a more general class of regular sequences which take on their values in a (possibly infinite) Noetherian ring R. Formally, a sequence is *k-regular* if the module generated by its k-kernel is finitely generated.

Many regular sequences have been found and studied in [3,5,6,21,26,28]. If a sequence $\{u(n)\}_{n\geq 0}$ takes finitely many values, Allouche and Shallit showed that it is regular if and only if it is automatic in [2]. Hence, we always assume that regular sequences take on infinitely many values. If a sequence $\{u(n)\}_{n\geq 0}$ is an unbounded integer regular sequence, Allouche and Shallit [2] proved that there exists $c_1 \geq 0$ such that $u(n) = O(n^{c_1})$ for all n and Bell et. al[7]. showed that there exists $c_2 \geq 0$ such that $|u(n)| > c_2 \log n$ infinitely often. Recently, Charlier et. al. characterized the regular sequences by counting the paths in the corresponding graph with finite vertices in [10].

Despite all this, there are no descriptions for regular sequences by automata. Note that automatic sequences can be generated by finite state automata, it is a natural question that can regular sequences be generated by automata with countable states, or morphisms on a countable alphabet?

☆ This research is supported by the Fundamental Research Funds for the Central Universities (Project No. 2662015QC010, 2662015QC016), and the National Natural Science Foundation of China (Grant Nos. 11501228, 11371156, 11431007, 11626110).

* Corresponding author.
E-mail addresses: zhangfanqie@hust.edu.cn (J.-M. Zhang), cj@mail.hzau.edu.cn (J. Chen), guoyingjun2005@126.com (Y.-J. Guo), zhi-xiong.wen@hust.edu.cn (Z.-X. Wen).

http://dx.doi.org/10.1016/j.chaos.2017.04.018
0960-0779/

Morphisms on infinite alphabets and countable states automata have been studied by many authors. In [22], Mauduit concerned the arithmetical and statistical properties of sequences generated by deterministic countable states automata or morphisms on a countable alphabet. Meanwhile, Ferenczi [14] studied morphism dynamical systems on infinite alphabets and Le Gonidec [17–20] studied complexity function for some q^∞-automatic sequences. More about morphisms on infinite alphabets and countable states automata, please see in [9,16,25]

In the present paper, we focus on morphisms on a countable alphabet and automata with countable states. We find a class of automata with countable states which can generate regular sequences. That is to say, a class of regular sequences can be generated by countable states automata.

This paper is organized as follows. In Section 2, we give some notations and definitions. In Section 3, we introduce a class of morphisms on infinite alphabets and countable states automata, which are called to be m-periodic k-uniform morphism and m-periodic k-DCAO, respectively. We prove that all the infinite sequences generated by them are k-regular. In Section 4, we consider the codings generated by the sequences satisfying a linear recurrence. Under some conditions, we show that the regularity is invariant under these codings. In the last section, we outline some generalizations.

2. Preliminary

Let $\mathbb{N}^{\geq 2}$ be the set of integers greater than 2. For every integer $b \in \mathbb{N}^{\geq 2}$, we define a nonempty alphabet $\Sigma_b := \{0, 1, \cdots, b-1\}$ and a countable alphabet $\Sigma_\infty := \mathbb{N} = \{0, 1, \cdots, n, \cdots\}$. For $b \in \mathbb{N}^{\geq 2} \cup \{\infty\}$, let Σ_b^* denote the set of all finite words on Σ_b. If $w \in \Sigma_b^*$, then its length is denoted by $|w|$. If $|w| = 0$, then we call w is the empty word, denoted by ϵ. Let Σ_b^k denote the set of words of length k on Σ_b, i.e., $\Sigma_b^* = \bigcup_{k\geq 0} \Sigma_b^k$. Let $u = u(0)u(1)\cdots u(m)$ and $v = v(0)v(1)\cdots v(n) \in \Sigma_b^*$. The word $uv := u(0)u(1)\cdots u(m)v(0)v(1)\cdots v(n)$ denotes their *concatenation*. If $|u| \geq 1$ (resp. $|v| \geq 1$), then u (resp. v) is a *prefix* (resp. *suffix*) of uv. Clearly, the set Σ_b^* together with the concatenation forms a monoid, where the empty word ϵ plays the role of the neutral element.

If $b \in \mathbb{N}^{\geq 2}$, then every non-negative integer n has a unique representation of the form $n = \sum_{i=0}^{\ell} n_i b^i$ with $n_\ell \neq 0$ and $n_i \in \Sigma_b$. We call $n_\ell n_{\ell-1} \cdots n_0$ its *canonical representation in base b*, denoted by $(n)_b$. If $\ell \geq |(n)_b|$, denote $(n)_b^\ell = 0^i (n)_b$ with $i = \ell - |(n)_b|$. If $(n)_b = n_\ell n_{\ell-1} \cdots n_0$, then the *base-b sum of digits function* is defined by $s_b(n) := \sum_{i=0}^{\ell} n_i$. If $b \in \mathbb{N}^{\geq 2}$ and $\mathbf{w} = w_\ell w_{\ell-1} \cdots w_0$, then $[\mathbf{w}]_b := \sum_{i=0}^{\ell} w_i \cdot b^i$. We denote by $\mathrm{rem}_b(n) := r$ if $n \equiv r \pmod b$ $(0 \leq r \leq b-1)$.

In this paper, unless otherwise stated, all alphabets under consideration are countable.

2.1. Morphisms on countable alphabets

Let Σ and Δ be two alphabets. A *morphism (or substitution)* is a map σ from Σ^* to Δ^* satisfying $\sigma(uv) = \sigma(u)\sigma(v)$ for all words $u, v \in \Sigma^*$. In the whole paper, we use the term "morphism".

Note that $\sigma(\epsilon) = \epsilon$. If $\Sigma = \Delta$, then we can iterate the application of σ. That is, $\sigma^i(a) = \sigma(\sigma^{i-1}(a))$ for all $i \geq 1$ and $\sigma^0(a) = a$.

Let σ be a morphism defined on $\Sigma = \{q_0, q_1, \cdots, q_n, \cdots\}$. If $\sigma(q_i) = q_{i_1} q_{i_2} \cdots q_{i_{t_i}}$ with $i_j = a_j i + b_j$ and $a_j, b_j \in \mathbb{Z}$, for every $i \geq 0$, then σ is called a *linear morphism*. If there exists some integer $k \geq 1$ such that $|\sigma(a)| = k$ for all $a \in \Sigma$, then σ is called a *k-uniform morphism (or k-constant length morphism)*. A 1-uniform morphism is called a *coding*. If there exists a finite or infinite word $w \in \Sigma^{\mathbb{N}}$ such that $\sigma(w) = w$, then the word w is a *fixed point* of σ. In fact, if $\sigma(a) = aw$ for some letter $a \in \Sigma$ and nonempty $w \in \Sigma^*$, then the sequence of words a, $\sigma(a)$, $\sigma^2(a)$, $\cdots$ converges to the infinite word

$$\sigma^\infty(a) := aw\sigma(w)\sigma^2(w)\cdots,$$

where the limit is defined by the metric $d(u, v) = 2^{-\min\{i: u(i) \neq v(i)\}}$ for $u = u(0)u(1)\cdots$ and $v = v(0)v(1)\cdots$. Clearly, $\sigma^\infty(a)$ is a fixed point of σ. Hence, for every morphism σ on the alphabet Σ, we always assume that there exists a letter $a \in \Sigma$ such that $\sigma(a) = aw$ with a nonempty word $w \in \Sigma^*$.

Example 1. Let $\Sigma = \Sigma_\infty := \{0, 1, \cdots, n, \cdots\}$. Define a 2-uniform morphism $\sigma(i) = i(i+1)$ for all $i \geq 0$, then $\sigma^\infty(i) = i(i+1)(i+1)(i+2)\cdots$ is a fixed point of σ. In particular, the fixed point $\sigma^\infty(0) = 01121223\cdots$ is the sequence of base-2 sum of digits function $\{s_2(n)\}_{n\geq 0}$.

Example 2 (The drunken man morphism)**.** Let $\Sigma = \{\iota\} \cup \mathbb{Z}$. Define a 2-uniform morphism $\sigma(\iota) = \iota 1$ and $\sigma(i) = (i-1)(i+1)$ for all $i \in \mathbb{Z}$, then the infinite word $\sigma^\infty(\iota) = \iota 102(-1)113(-2)0020224\cdots$ is the only non-empty fixed point of σ.

Example 3 (Infinibonacci morphism)**.** Let $\Sigma = \mathbb{N}$. Define a 2-uniform morphism $\sigma(i) = 0(i+1)$ for all $i \geq 0$, then $\sigma^\infty(0) = 0102010301020104\cdots$ is a fixed point of σ.

2.2. Deterministic infinite states automata

A *deterministic countable automaton (DCA)* is a directed graph $M = (Q, \Sigma, \delta, q_0, F)$, where Q is a countable set of states, $q_0 \in Q$ is the initial state, Σ is the finite input alphabet, $F \subseteq Q$ is the set of accepting states, δ: $Q \times \Sigma \to Q$ is the transition function. The transition function δ can be extended to $Q \times \Sigma^*$ by $\delta(q, \epsilon) = q$ and $\delta(q, wa) = \delta(\delta(q, w), a)$ for all $q \in Q$, $a \in \Sigma$ and $w \in \Sigma^*$.

Similarly, a *deterministic countable states automaton with output (DCAO)* is defined to be a 6-tuple $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$, where Q, Σ, δ, q_0 are as in the definition of DCA as above, Δ is the output alphabet and $\tau: Q \to \Delta$ is the output function. In particular, if the input alphabet $\Sigma = \Sigma_k$ for some $k \in \mathbb{N}^{\geq 2}$, then the automaton M is always called to be a k-DCAO.

Let $\{u(n)\}_{n\geq 0} = u(0)u(1)u(2)\cdots$ be a sequence on the alphabet Δ. The sequence $\{u(n)\}_{n\geq 0}$ is called to be *k-automatic*, if the sequence can be generated by a k-DCAO, that is, there exists a k-DCAO $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ such that $u(n) = \tau(\delta(q_0, w))$ for all $n \geq 0$, $w \in \Sigma_k^*$ and $[w]_k = n$.

By the choice of DCAO M satisfying $\delta(q_0, 0) = q_0$, our machine M always computes the same $u(n)$ even if the input one has leading zeros. Hence, unless otherwise stated, all DCAOs satisfy $\delta(q_0, 0) = q_0$ and $u(n) = \tau(\delta(q_0, (n)_k))$ for all $n \geq 0$.

Example 4. Let $Q = \{q_0, q_1, q_2, \cdots\}$, $\Delta = \mathbb{N}$, $\delta(q_i, 0) = q_i$, $\delta(q_i, 1) = q_{i+1}$ and $\tau(q_i) = i$ for all $i \geq 0$. Then, the sequence of base-2 sum of digits function $\{s_2(n)\}_{n\geq 0}$ is 2-automatic. It can be generated by a 2-DCAO in Fig. 1.

Example 5. Let $Q = \{q_0\} \cup \mathbb{Z}$, $\Delta = \{\iota\} \cup \mathbb{Z}$, $\delta(q_0, 0) = q_0$, $\delta(q_0, 1) = 1$, $\delta(i, 0) = i-1$, $\delta(i, 1) = i+1$, $\tau(q_0) = \iota$ and $\tau(i) = i$ for all $i \in \mathbb{Z}$. Then, the sequence defined in Example 2 can be generated by a 2-DCAO in Fig. 2.

Example 6. Let $Q = \{q_0, q_1, q_2, \cdots\}$, $\Delta = \mathbb{N}$, $\delta(q_i, 0) = q_0$, $\delta(q_i, 1) = q_{i+1}$ and $\tau(q_i) = i$ for all $i \geq 0$. Then, the sequence defined in Example 3 is 2-automatic. It can be generated by a 2-DCAO in Fig. 3.

By the definitions of k-uniform morphism and k-DCAO, we note that each sequence $\mathbf{u} = \{u(n)\}_{n\geq 0}$ can be generated by a k-uniform morphism or a k-DCAO for every $k \in \mathbb{N}^{\geq 2}$. In fact, let $\sigma : \mathbb{N} \to \mathbb{N}^*$ be a k-uniform morphism defined by $\sigma(i) = (ki)(ki+1)\cdots(ki+k-$