Linear Algebra and its Applications 500 (2016) 119–148

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Asymptotic properties of free monoid morphisms

Émilie Charlier *, Julien Leroy, Michel Rigo

Université de Liège, Institut de mathématique, Allée de la découverte 12 (B37), 4000 Liège, Belgium

ARTICLE INFO

Article history:
Received 1 July 2015
Accepted 20 February 2016
Available online 21 March 2016
Submitted by R. Brualdi

MSC:
68R15
11B85

Keywords:
Free monoid
Morphism
Asymptotics
Non-negative matrix
Numeration system
Algorithm

ABSTRACT

Motivated by applications in the theory of numeration systems and recognizable sets of integers, this paper deals with morphic words when erasing morphisms are taken into account. Cobham showed that if an infinite word $\mathbf{w} = g(f^\omega(a))$ is the image of a fixed point of a morphism f under another morphism g, then there exist a non-erasing morphism σ and a coding τ such that $\mathbf{w} = \tau(\sigma^\omega(b))$.

Based on the Perron theorem about asymptotic properties of powers of non-negative matrices, our main contribution is an in-depth study of the growth type of iterated morphisms when one replaces erasing morphisms with non-erasing ones. We also explicitly provide an algorithm computing σ and τ from f and g.

1. Introduction

Infinite words, i.e., infinite sequences of symbols from a finite set usually called alphabet, form a classical object of study. They have an important representation power: they are a natural way to code elements of an infinite set using finitely many symbols, e.g., the coding of an orbit in a discrete dynamical system or the characteristic sequence of

* Corresponding author.
E-mail addresses: echarlier@ulg.ac.be (É. Charlier), J.Leroy@ulg.ac.be (J. Leroy), M.Rigo@ulg.ac.be (M. Rigo).

http://dx.doi.org/10.1016/j.laa.2016.02.030

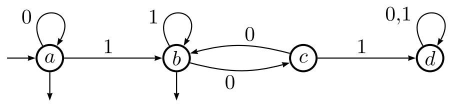

Fig. 1. The Baum–Sweet set is 2-recognizable.

a set of integers. A rich family of infinite words, with a simple algorithmic description, is made of the words obtained by iterating a morphism [5]. The necessary background about words is given in Section 3.1.

In relation with numeration systems, recognizable sets of integers are well studied. For instance, see [3]. Let $k \geq 2$ be an integer. A set $X \subseteq \mathbb{N}$ is said to be *k-recognizable* if the set of base-k expansions of the elements in X is accepted by a finite automaton. Characteristic sequences of k-recognizable sets have been characterized by Cobham [10]. They are the images of a fixed point of a k-uniform morphism under a coding (also called letter-to-letter morphism). We let A^* denote the set of finite words over the alphabet A. This set, equipped with a product which is the usual concatenation of words, is a monoid. A *morphism* $f\colon A^* \to B^*$ satisfies, for all $u, v \in A^*$, $f(uv) = f(u)f(v)$. A morphism is *k-uniform* if the image of every letter is a word of length k. A 1-uniform morphism is a *coding*. As an example of recognizable set, the Baum–Sweet set S is defined as follows [1]. The integer n belongs to S if and only if the base-2 expansion of n contains no block of consecutive 0's of odd length. The set S is 2-recognizable, the deterministic automaton depicted in Fig. 1 recognizes the base-2 expansions of the elements in S (read most significant digit first). The characteristic sequence $\mathbf{x}$ of S starts with $1101100101001001\cdots$. It is the image of the infinite word $abcbbdcbcbddbdcb\cdots$ under the coding $\tau\colon a, b \mapsto 1, c, d \mapsto 0$. Moreover, the latter infinite word is a fixed point of the 2-uniform morphism $\sigma\colon a \mapsto ab, b \mapsto cb, c \mapsto bd, d \mapsto dd$. We write $\mathbf{x} = \tau(\sigma^\omega(a))$. Indeed, to obtain $\mathbf{x}$, one iterates the morphism σ from a to get a sequence $(\sigma^n(a))_{n\geq 0}$ of finite words of increasing length whose first terms are: $a, ab, abcb, abcbbdcb, abcbbdcbcbddbdcb, \ldots$. This sequence converges to an infinite word which is a fixed point of σ. See, for instance, [4,32] for the definition of converging sequences of words. Note that there are infinitely many morphisms that can be used to generate the word $\mathbf{x}$. Take $\sigma'\colon a \mapsto abe, b \mapsto cefb, c \mapsto bfd, d \mapsto defd, e \mapsto ef, f \mapsto \varepsilon$ where ε is the empty word (the identity element for concatenation), i.e., the unique word of length 0. In that case, we say that σ' is *erasing*. Take $\tau'\colon a, b \mapsto 1, c, d \mapsto 0, e, f \mapsto \varepsilon$. One fixed point of σ' starts with $abecefbefbfdefcefb\cdots$ and the image by the erasing morphism τ' of this word again is $\mathbf{x}$. The general aim of this paper is to derive from erasing morphisms such as σ' and τ' new non-erasing morphisms (where images of all letters have positive length) such as σ and τ that produce the same infinite word $\mathbf{x}$ and to retrieve some kind of canonical information (e.g., spectral radius, growth order) about $\mathbf{x}$ itself.

In the theory of integer base systems, we also recall another important theorem of Cobham [9]. Let $k, \ell \geq 2$ be two multiplicatively independent integers, i.e., they are such that $\log k/\log \ell$ is irrational. If a set $X \subseteq \mathbb{N}$ is both k-recognizable and ℓ-recognizable,