CrossMark

Self-similar groups, automatic sequences, and unitriangular representations

R. Grigorchuk[1] · Y. Leonov[2] · V. Nekrashevych[1] · V. Sushchansky[3]

Dedicated to L. A. Bokut on the occasion of his 80th birthday

Abstract We study natural linear representations of self-similar groups over finite fields. In particular, we show that if the group is generated by a finite automaton, then obtained matrices are automatic. This shows a new relation between two separate notions of automaticity: groups generated by automata and automatic sequences. We also show that if the group acts on the tree by p-adic automorphisms, then the corresponding linear representation is a representation by infinite triangular matrices. We relate this observation with the notion of height of an automorphism of a rooted tree due to L. Kaloujnine.

Contents

Communicated by Efim Zelmanov.

✉ V. Nekrashevych
  nekrash@math.tamu.edu

R. Grigorchuk
  grigorch@math.tamu.edu

[1] Texas A&M University, College Station, USA

[2] Odessa National Academy of Telecommunications, Odessa, Ukraine

[3] Silesian University of Technology, Gliwice, Poland

Springer

1 Introduction

Self-similar groups is an active topic of modern group theory. They initially appeared as interesting examples of groups with unusual properties (see [1,21,24,44]). The main techniques of the theory were developed for the study of these examples. Later a connection to dynamical systems was discovered (see [33,35]) via the notion of the iterated monodromy group. Many interesting problems were solved using self-similar groups (see [10,11,16,22]).

One of the ways to define self-similar groups is to say that they are groups generated by all states of an automaton (of Mealy type, also called a transducer, or sequential machine). Especially important case is when the group is generated by the states of a finite automaton. All examples mentioned above (including the iterated monodromy groups of expanding dynamical systems) are like that.

The main goal of this article is to indicate a new relation between self-similar groups and another classical notion of automaticity: automatic sequences and matrices. See the monographs [2,50] for theory of automatic sequences and applications.

More precisely, we are going to study natural linear representations of self-similar groups over finite fields, and show that matrices associated with elements of a group generated by a finite automaton are automatic.

There are several ways to define automatic sequences and matrices. One can use Moore automata, substitutions (e.g., Morse–Thue substitution leading to the famous Morse–Thue sequence), or Christol's characterization of automatic sequences in terms of algebraicity of the generating power series over a suitable finite field [2, Theorem 12.2.5]. The theory of automatic sequences is rich and is related to many topics in dynamical systems, ergodic theory, spectral theory of Schrödinger operators, number theory etc., see [2,50].

It is well known that linear groups (that is subgroups of groups of matrices $GL_N(\Bbbk)$, where $\Bbbk$ is a field) is quite a restrictive class of groups as the Tits alternative [48] holds

for them. Moreover, the group of (finite) upper triangular matrices is solvable, and the group of upper unitriangular matrices is nilpotent. In contrast, if one uses infinite triangular matrices over a finite field, one can get much more groups. In particular, every countable residually finite p-group can be embedded into the group of upper uni-triangular matrices over the finite field $\mathbb{F}_p$.

We will pay special attention to the case when the constructed representation is a representation by infinite unitriangular matrices. One of the results of our paper is showing that the natural (and optimal in certain sense) representation by uni-triangular matrices constructed in [29,31] leads to automatic matrices, if the group is generated by a finite automaton. In particular, the diagonals of these uni-triangular matrices are automatic sequences. We study them separately, in particular, by computing their generating series (as algebraic functions).

The roots of the subject of our paper go back to L. Kaloujnin's results on Sylow p-subgroups of the symmetric group [26–28], theory of groups generated by finite automata [1,17,19,25,45], theory of self-similar groups [7,35], and group actions on rooted trees [8,19,23].

Note that study of actions on rooted trees (every self-similar group is, by definition, an automorphism group of the rooted tree of finite words over a finite alphabet) is equivalent to the study of residually finite groups by geometric means, i.e., via representations of them in groups of automorphisms of rooted trees. The theory of actions on rooted trees is quite different from the Bass–Serre theory [41] of actions on (unrooted) trees, and uses different methods and tools. The important case is when a group is residually finite p-group (p prime), i.e., is approximated by finite p-groups. The class of residually finite p-groups contains groups with many remarkable properties. For instance, Golod's p-groups that were constructed in [20] based on Golod-Shafarevich theorem to answer a question of Burnside on the existence of a finitely generated infinite torsion group are residually p groups. Other important examples are the first self-similar groups mentioned at the beginning of this introduction.

At the end of the paper we study a notion of uniseriality which plays an important role in the study of actions of groups on finite p-groups [15,30]. Our analysis is based upon classical results of L. Kaloujnine on height of automorphisms of rooted trees [27,28]. Applications of uniseriality to Lie algebras associated with self-similar groups were, for instance, demonstrated in [5]. Proposition 5.29 gives a simple criterion of uniseriality of action of a group on rooted trees and allows to substitute lemma 5.2 from [5]. A number of examples is presented which demonstrate the basic notions, ideas, and results.

2 Groups acting on rooted trees

2.1 Rooted trees and their automorphisms

Let X be a finite alphabet. Denote by X^* the set of finite words over X, which we will view as the free monoid generated by X. It is a disjoint union of the sets X^n of words of *length n*. We denote the empty word, which is the only element of X^0, by ∅. We will write elements of X^n either as words $x_1 x_2 \cdots x_n$, or as n-tuples $(x_1, x_2, \ldots, x_n)$.

We consider X^* as the set of vertices of a rooted tree, defined as the right Cayley graph of the free monoid. Namely, two vertices of X^* are connected by an edge if and only if they are of the form v, vx for $v \in X^*$ and $x \in X$. The empty word is the root of the tree. For $v \in X^*$, we consider $vX^* = \{vw \ : \ w \in X^*\}$ as the set of vertices of a sub-tree with the root v.

Denote by $\operatorname{Aut}(X^*)$ the group of all automorphisms of the rooted tree X^*. Every element of $\operatorname{Aut}(X^*)$ preserves the levels X^n of the tree, and for every $g \in \operatorname{Aut}(X^*)$ beginning of length $n \leq m$ of the word $g(x_1 x_2 \cdots x_m)$ is equal to $g(x_1 x_2 \cdots x_n)$. It follows that for every $v \in X^*$ the transformation $\alpha_{g,v} : X \longrightarrow X$ defined by $g(vx) = g(v)\alpha_{g,v}(x)$ is a permutation, and that the action of g on X^* is determined by these permutations according to the rule

$$g(x_1 x_2 \cdots x_n) = \alpha_{g,\emptyset}(x_1)\alpha_{g,x_1}(x_2)\alpha_{g,x_1 x_2}(x_3)\cdots\alpha_{g,x_1 x_2\cdots x_{n-1}}(x_n). \tag{1}$$

The map $v \mapsto \alpha_{g,v}$ from X^* to the symmetric group $\operatorname{Symm}(X)$ is called the *portrait* of the automorphism g.

Equivalently, we can represent g by the sequence

$$\tau = [\tau_0, \tau_1, \tau_2, \ldots]$$

of maps $\tau_n : X^n \longrightarrow \operatorname{Symm}(X)$, where $\tau_n(v) = \alpha_{g,v}$. Such sequences are called, following L. Kaloujnine [26], *tableaux*, and are denoted $\tau(g)$.

If $[\tau_0, \tau_1, \ldots]$ and $[\sigma_0, \sigma_1, \ldots]$ are tableaux of elements $g_1, g_2 \in \operatorname{Aut}(X^*)$, respectively, then the tableau of their product $g_1 g_2$ is the sequence of functions

$$\tau(g_1 g_2) = [\tau_n(g(x_1 x_2 \cdots x_n)) \cdot \sigma_n(x_1 x_2 \cdots x_n)]_{i=0}^{\infty}. \tag{2}$$

The tableau of the inverse of the element g_1 is

$$\tau(g_1^{-1}) = \left[\tau_n(g^{-1}(x_1 x_2 \cdots x_n))^{-1}\right]_{n=0}^{\infty}. \tag{3}$$

Here, and in most of our paper, (except when we talk about bisets, i.e., about sets with left and right actions) group elements and permutations act from the left.

Denote by $X^{[n]}$ the finite sub-tree of X^* spanned by the set of vertices $\bigcup_{k=0}^{n} X^k$. The group $\operatorname{Aut}(X^*)$ acts on $X^{[n]}$, and the kernel of the action coincides with the kernel of the action on X^n. The quotient of $\operatorname{Aut}(X^*)$ by the kernel of the action is a finite group, which is naturally identified with the full automorphism group of the tree $X^{[n]}$. We will denote this finite group by $\operatorname{Aut}(X^{[n]})$.

The group $\operatorname{Aut}(X^*)$ is naturally isomorphic to the inverse limit of the groups $\operatorname{Aut}(X^{[n]})$ (with respect to the restriction maps). This shows that $\operatorname{Aut}(X^*)$ is a profinite group. The basis of neighborhoods of the identity of $\operatorname{Aut}(X^*)$ is the set of kernels of its action on the levels X^n of the tree X^*.

2.2 Self-similarity of Aut(X^*)

Let $g \in \text{Aut}(X^*)$, and $v \in X^*$. Then there exists an automorphism of X^*, denoted $g|_v$ such that

$$g(vw) = g(v)g|_v(w)$$

for all $w \in X^*$.

We call $g|_v$ the *section* of g at v. The sections obviously have the properties

$$g|_{v_1v_2} = g|_{v_1}|_{v_2}, \qquad (g_1g_2)|_v = g_1|_{g_2(v)}g_2|_v, \qquad (4)$$

for all $g, g_1, g_2 \in \text{Aut}(X^*)$ and $v, v_1, v_2 \in X^*$.

The portrait of the section $g|_v$ is obtained by restricting the portrait of g to the subtree vX^*, and then identifying vX^* with X^* by the map $vw \mapsto w$.

Definition 2.1 The set $g|_{X^*} = \{g|_v : v \in X^*\} \subset \text{Aut}(X^*)$ for $g \in \text{Aut}(X^*)$, is called the *set of states* of g. An automorphism $g \in \text{Aut}(X^*)$ is said to be *finite state* if $g|_{X^*}$ is finite.

It follows from (4) that

$$g^{-1}|_x = (g|_{X^*})^{-1}, \qquad (g_1g_2)|_x \subset g_1|_xg_2|_x,$$

which implies that the set of finite state elements of $\text{Aut}(X^*)$ is a group. We call it the *group of finite automata*, and denote it $\text{FAut}(X^*)$. This name comes from interpretation of elements of $\text{Aut}(X^*)$ with automata (transducers), see 3.1 below. Namely, the set of states of the automaton corresponding to g is $g|_{X^*}$. The element $g \in g|_{X^*}$ is the *initial state*. If the current state of the automaton is h, and when it reads a letter $x \in X$ on its input, then it outputs $h(x)$ and changes its current state to $h|_x$. It is easy to check that if we give the consecutive letters of a word $x_1x_2 \cdots x_n$ on input of the automaton with the initial state g, then we will get on output the word $g(x_1x_2 \cdots x_n) = g(x_1)g|_{x_1}(x_2)g|_{x_1x_2}(x_3) \ldots$, compare with (1).

Every element $g \in \text{Aut}(X^*)$ is uniquely determined by the permutation π it defines on the first level $X \subset X^*$ and the first level sections $g|_x$, $x \in X$. In fact, the map

$$g \mapsto \pi \cdot (g|_x)_{x \in X}$$

is an isomorphism of $\text{Aut}(X^*)$ with the wreath product $\text{Symm}(X) \ltimes \text{Aut}(X^*)^X = \text{Symm}(X) \wr \text{Aut}(X^*)$. We call the isomorphism

$$\phi : \text{Aut}(X^*) \longrightarrow \text{Symm}(X) \wr \text{Aut}(X^*) : g \mapsto \pi \cdot (g|_x)_{x \in X}$$

the *wreath recursion*.

For a fixed ordering $x_1, x_2, \ldots, x_d$ of the letters of X, the elements of $\text{Symm}(X) \wr \text{Aut}(X^*)$ are written as $\pi(g_1, g_2, \ldots, g_d)$, where $\pi \in \text{Symm}(d)$ and $g_i = g|_{x_i}$.

Definition 2.2 A subgroup $G \leq \text{Aut}(X^*)$ is said to be *self-similar* if $g|_x \in G$ for all $g \in G$ and $x \in X$.

In other words, a group $G \leq \text{Aut}(X^*)$ is self-similar if restriction of the wreath restriction to G is a homomorphism $\phi : G \longrightarrow \text{Symm}(X) \wr G$. Note that the wreath recursion is usually not an isomorphism (but is an embedding, since we assume that G acts faithfully on X^*).

Example 2.3 Let $X = \{0, 1\}$. Consider the automorphism of the tree X^* given by the rules

$$a(0w) = 1w, \qquad a(1w) = 0a(w).$$

These rules can be written using wreath recursion as $\phi(a) = \sigma(1, a)$, where $\sigma = (01)$ is the transposition. We will usually omit ϕ, and write

$$a = \sigma(1, a),$$

thus identifying $\text{Aut}(X^*)$ with $\text{Symm}(X) \wr \text{Aut}(X^*)$.

The automorphism a is called the (binary) *adding machine*, since it describes the process of adding one to a dyadic integer: $a(x_1 x_2 \cdots x_n) = y_1 y_2 \cdots y_n$ if and only if

$$(x_1 + 2x_2 + 2^2 x_3 + \cdots + 2^{n-1} x_n) + 1 = y_1 + 2y_2 + \cdots + 2^{n-1} y_n \qquad (\text{mod } 2^n).$$

The group generated by a (which is infinite cyclic) is self-similar, and is a subgroup of the group of finite automata.

Example 2.4 Consider the group G generated by the elements a, b, c, d that are defined inductively by the recursions

$$a = \sigma, \quad b = (a, c), \quad c = (a, d), \quad d = (1, b).$$

Here σ, as before, is the transposition (01), and when we omit either the element of $\text{Symm}(X)$ or the element of $\text{Aut}(X^*)^X$ when writing elements of $\text{Symm}(X) \wr \text{Aut}(X^*)$, we assume that it is equal to the identity element of the respective group.

The group G is then a self-similar subgroup of the group of finite automata. It is the Grigorchuk group, defined in [21].

2.3 Self-similarity bimodule

We can identify the letters $x \in X$ with transformations $v \mapsto xv$ of the set X^*. Then the identity $g(xv) = yh(v)$ for $x \in X$, $y = g(x)$, and $h = g|_x$ is written as equality of compositions of transformations:

$$g \cdot x = y \cdot h.$$

Consider the set $X \cdot G$ of compositions of the form $x \cdot g$, i.e., transformations $v \mapsto xg(v)$, $v \in X^*$. It is closed with respect to pre- and post-compositions with the elements of G:

$$(x \cdot g) \cdot h = x \cdot (gh), \qquad h \cdot (x \cdot g) = h(x) \cdot (h|_x g).$$

We get in this way a *biset*, i.e., a set with two commuting left and right actions of the group G.

Let $\Bbbk$ be a field, and let $\Bbbk[G]$ be the group ring over $\Bbbk$. Denote by Φ the vector space $\Bbbk^{X \cdot G}$ spanned by $X \cdot G$. Then the left and the right actions of G on $X \cdot G$ are extended by linearity to a structure of a $\Bbbk[G]$-bimodule on Φ. We will denote by ${}_G\Phi$ and Φ_G the space Φ seen as a left and a right $\Bbbk[G]$-module, respectively.

It follows directly from the definition of the right action of G on $X \cdot G$ that X (identified with $X \cdot 1$) is a free basis of Φ_G. The left action is not free in general, since it is possible to have $g(xv) = xv$ for all $v \in X^*$ and for a non-trivial element $g \in G$, which will imply, by definition of the left action, that $g \cdot x = x$.

For every element $a \in \Bbbk[G]$ the map $v \mapsto a \cdot v$ for $v \in \Phi$ is an endomorphism of Φ_G, denoted $\Xi(a)$. The map $\Xi : \Bbbk[G] \longrightarrow \mathrm{End}\,\Phi_G$ is obviously a homomorphism of $\Bbbk$-algebras.

After fixing a basis of the right module Φ_G (for example X), we can identify the algebra of endomorphisms $\mathrm{End}\,\Phi_G$ of the right $\Bbbk[G]$-module Φ_G with the algebra of $|X| \times |X|$ matrices over $\Bbbk[G]$. In this case the homomorphism $\Xi : \Bbbk[G] \longrightarrow \mathsf{M}_{|X|}(\Bbbk[G]) \cong \mathrm{End}\,\Phi_G$ is called the *matrix recursion* associated with the self-similar group G (and the basis of the right module).

More explicitly, if $B = \{e_1, \ldots, e_d\}$ is a basis of the right $\Bbbk[G]$-module Φ_G, then, for $a \in \Bbbk[G]$ the matrix $\Xi(a) = (a_{i,j})_{1 \le i, j \le d} \in \mathsf{M}_d(\Bbbk[G])$ is given by the condition

$$a \cdot e_j = \sum_{i=1}^{d} e_i \cdot a_{i,j}.$$

If we use the basis $\{x_1, x_2, \ldots, x_d\} = X$ of the right module Φ_G, then the matrix recursion Ξ is a direct rewriting of the wreath recursion $\phi : G \longrightarrow \mathrm{Symm}(X) \wr G$ in matrix terms. Namely, $\Xi(g)$ is the matrix with entries a_{ij}, $1 \le i, j \le d$, given by the rule

$$a_{ij} = \begin{cases} g|_{x_j} & \text{if } g(x_j) = x_i, \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Example 2.5 The adding machine recursion $a = \sigma(1, a)$ is defined in the terms of the bimodule as

$$a \cdot x_0 = x_1, \qquad a \cdot x_1 = x_0 \cdot a,$$

where x_0, x_1 are identified with the symbols 0, 1, respectively, from Example 2.3.

It follows that the recursion is written in matrix form as

$$\Xi(a) = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}.$$

The recursive definition of the generators a, b, c, d of the Grigorchuk group is written as

$$\Xi(a) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Xi(b) = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}, \quad \Xi(c) = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}, \quad \Xi(d) = \begin{pmatrix} 1 & 0 \\ 0 & b \end{pmatrix}.$$

When we change the basis of the right module Φ_G, we just conjugate the map Ξ by the transition matrix. Namely, if $\{x_1, \ldots, x_d\}$ and $\{y_1, \ldots, y_d\}$ are bases of the right module Φ_G, then we can write $y_j = \sum_{i=1}^{\infty} y_i \cdot b_{i,j}$ for $b_{i,j} \in \Bbbk[G]$. Then the matrix $T = (b_{i,j})_{1 \le i, j \le d}$ is the transition matrix from the basis $\{x_i\}_{1 \le i \le d}$ to the basis $\{y_i\}_{1 \le i \le d}$.

Example 2.6 Consider again the adding machine example. Let us take, instead of the standard basis $\{x_0, x_1\} = X$, the basis $y_0 = x_0 + x_1$, $y_1 = x_1$. (Here we replace the letters 0, 1 of the binary alphabet by x_0 and x_1, respectively, in order not to confuse them with elements $0, 1 \in \Bbbk[G]$.) Then the transition matrix to the new basis is $T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. It inverse is $T^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$. Consequently, the matrix recursion in the new basis is

$$a \mapsto T^{-1} \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix} T = \begin{pmatrix} a & a \\ 1-a & -a \end{pmatrix}.$$

This can be checked directly:

$$a \cdot y_0 = a \cdot (x_0 + x_1) = x_1 + x_0 \cdot a = y_1 + (y_0 - y_1) \cdot a = y_0 \cdot a + y_1 \cdot (1 - a)$$

and

$$a \cdot y_1 = a \cdot x_1 = x_0 \cdot a = y_0 \cdot a - y_1 \cdot a.$$

If we take the basis $\{y_0 = x_0, y_1 = x_1 \cdot a\}$, then matrix recursion becomes

$$a \mapsto \begin{pmatrix} 1 & 0 \\ 0 & a^{-1} \end{pmatrix} \cdot \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} = \begin{pmatrix} 0 & a^2 \\ a^{-1} & 0 \end{pmatrix}.$$

If the basis is a subset of $X \cdot G$, then the matrix recursion corresponds to a wreath recursion $G \mapsto \text{Symm}(X) \wr G$. For instance, in the last example the matrix recursion corresponds to the wreath recursion

$$a \mapsto \sigma(a^{-1}, a^2).$$

This wreath recursion describes the process of adding 1 to a dyadic numbers in the binary numeration system with digits 0 and 3. For more on changes of bases in the biset $X \cdot G$ and the corresponding transformations of the wreath recursion see [35,36].

If Φ_1 and Φ_2 are bimodules over a $\Bbbk$-algebra A, then their tensor product $\Phi_1 \otimes \Phi_2$ is the quotient of the $\Bbbk$-vector space spanned by $\Phi_1 \times \Phi_2$ by the sub-space generated by the elements of the form

$$(v_1, a \cdot v_2) - (v_1 \cdot a, v_2)$$

for $v_1 \in \Phi_1$, $v_2 \in \Phi_2$, $a \in A$. It is a $\Bbbk[G]$-bimodule with respect to the actions $a \cdot (v_1 \otimes v_2) = (a \cdot v_1) \otimes v_2$ and $(v_1 \otimes v_2) \cdot a = v_1 \otimes (v_2 \cdot a)$.

If Φ_2 is a left A-module, and Φ_1 is an A-bimodule, then the left module $\Phi_1 \otimes \Phi_2$ is defined in the same way.

Let Φ, as above, be the bimodule associated with a self-similar group G. Then X is a basis of the right $\Bbbk[G]$-module Φ_G, and the set

$$X^n = \{x_1 \otimes \cdots \otimes x_n \ : \ x_i \in X\}$$

is a basis of the right module $\Phi_G^{\otimes n}$, which is hence a free module. Note that $X^n \cdot G$ is the basis of $\Phi^{\otimes n}$ as a vector space over $\Bbbk$.

We identify $x_1 \otimes \cdots \otimes x_n$ with the word $x_1 \cdots x_n$. The left module structure on $\Phi^{\otimes n}$ is given by the rules similar to the definition of Φ:

$$g \cdot v = g(v) \cdot g|_v \tag{6}$$

for $v \in X^n$ and $g \in G$. In particular, up to an ordering of the basis X^n, the associated matrix recursion $\Xi^n : \Bbbk[G] \longrightarrow \mathsf{M}_{|X^n|}(\Bbbk[G])$ is obtained from the recursion $\Xi^{n-1} : \Bbbk[G] \longrightarrow \mathsf{M}_{|X^{n-1}|}(\Bbbk[G])$ by replacing every entry a_{ij} of the matrix $\Xi^{n-1}(a)$ by the matrix $\Xi(a_{ij})$.

Example 2.7 The matrix recursion $G \longrightarrow \mathsf{M}_4(\Bbbk[G])$ for the adding machine (in the standard basis X^2) is

$$a \mapsto \left(\begin{array}{cc|cc} 0 & 0 & 0 & a \\ 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right),$$

which is obtained by iterating the matrix recursion

$$a \mapsto \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}.$$

In this case the basis X^2 is ordered in the lexicographic order $x_0 x_0 < x_0 x_1 < x_1 x_0 < x_1 x_1$. But since a is the adding machine, and it describes adding 1 to a dyadic integer that is written is such a way that the less significant digits come before the more significant ones, it is more natural to order the basis in the *inverse* lexicographic order $x_0 x_0 < x_1 x_0 < x_0 x_1 < x_1 x_1$. In this case the matrix recursion becomes

$$a \mapsto \begin{pmatrix} 0 & 0 & 0 & a \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Proposition 2.8 *Let T be the transition matrix from a basis X of Φ_G to a basis Y. Suppose that all entries of T are elements of $\Bbbk$. Then the transition matrices T_n from the basis $X^{\otimes n}$ to $Y^{\otimes n}$ are equal to*

$$T_n = \underbrace{T \otimes T \otimes \cdots \otimes T}_{|X|},$$

where $\otimes$ is the Kronecker product of matrices.

Proof Let $T_{n-1} = (a_{u,v})_{u \in X^{\otimes(n-1)}, v \in Y^{\otimes(n-1)}}$, i.e.,

$$v = \sum_{u \in X^{\otimes(n-1)}} u \cdot a_{u,v}$$

for all $v \in Y^{\otimes(n-1)}$. Similarly, denote $T = (a_{x,y})_{x \in X, y \in Y}$. Then

$$y \otimes v = y \otimes \sum_{u \in X^{\otimes(n-1)}} u \cdot a_{u,v} = \left(\sum_{x \in X} x \cdot a_{x,y} \right) \otimes \sum_{u \in X^{\otimes(n-1)}} u \cdot a_{u,v}$$

$$= \sum_{u \otimes x \in X^{\otimes n}} x \cdot \otimes a_{x,y} \cdot u \cdot a_{u,v} = \sum_{u \otimes x \in X^{\otimes n}} x \otimes u \cdot a_{x,y} \cdot a_{u,v},$$

which shows that

$$a_{xu,yv} = a_{x,y} a_{u,v},$$

which agrees with the definition of the Kronecker product.   □

In other words, we can write

$$T_n = T^{(n)} \begin{pmatrix} T_{n-1} & 0 & \cdots & 0 \\ 0 & T_{n-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_{n-1} \end{pmatrix}, \tag{7}$$

where $T_1 = T$, and $T^{(n)}$ is the matrix T in which each entry a_{ij} is replaced by a_{ij} times the unit matrix of dimension $|X|^{n-1} \times |X|^{n-1}$. Here the rows and columns of T_n correspond to the elements of $X^{\otimes n}$ and $Y^{\otimes n}$, respectively, ordered lexicographically.

It is easy to see from the proof, that in the general case (when not all entries of T are elements of $\Bbbk$), the formula (7) remains to be true, if we replace $T^{(n)}$ by the image of T under the $(n-1)$st iteration of the matrix recursion (in the basis X).

Example 2.9 Let $\Bbbk = \mathbb{R}$, $X = \{x_0, x_1\}$. Consider a new basis of Φ_G

$$\left\{ y_0 = \frac{x_0 + x_1}{\sqrt{2}}, \; y_1 = \frac{x_0 - x_1}{\sqrt{2}} \right\}.$$

The transition matrix to the new basis is

$$\begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Then the transition matrix from $X^{\otimes n}$ to $Y^{\otimes n}$ satisfies the recursion

$$H_n = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} H_{n-1} & 0 \\ 0 & H_{n-1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{pmatrix}.$$

2.4 Inductive limit of $\Bbbk^{X^n}$

Let $\Bbbk^{X^n}$ be the vector space of functions $X^n \longrightarrow \Bbbk$. It is naturally isomorphic to the nth tensor power of $\Bbbk^X$. The isomorphism maps an elementary tensor $f_1 \otimes f_2 \otimes \cdots \otimes f_n$ to the function

$$f_1 \otimes f_2 \otimes \cdots \otimes f_n(x_1 x_2 \cdots x_n) = f_1(x_1) f_2(x_2) \cdots f_n(x_n).$$

More generally, we have natural isomorphisms $\Bbbk^{X^n} \otimes \Bbbk^{X^m} \cong \Bbbk^{X^{n+m}}$ defined by the equality

$$f_1 \otimes f_2(x_1 x_2 \cdots x_{n+m}) = f_1(x_1 x_2 \cdots x_n) f_2(x_{n+1} x_{n+2} \cdots x_{n+m}).$$

We denote by δ_v, for $v \in X^n$ the delta-function of v, i.e., the characteristic function of $\{v\}$. It is an element of $\Bbbk^{X^n}$. Note that

$$\delta_{x_1 x_2 \cdots x_n} = \delta_{x_1} \otimes \delta_{x_2} \otimes \cdots \otimes \delta_{x_n},$$

with respect to the above identification of $\Bbbk^{X^{n+m}}$ with $\Bbbk^{X^n} \otimes \Bbbk^{X^m}$.

Let $G \leq \mathrm{Aut}(X^*)$. Denote by π_n the natural permutational representation of G on $\Bbbk^{X^n}$ coming from the action G on X^n. It is given by the rule $\pi_n(\delta_v) = \delta_{g(v)}$, i.e., by

$$\pi_n(g)(f)(v) = f(g^{-1}(v)), \quad f \in \Bbbk^X, \quad v \in X^n.$$

Denote by V_n the vector space $\Bbbk^{X^n}$ seen as a left $\Bbbk[G]$-module of the representation π_n, and by $[\varepsilon]$ the left $\Bbbk[G]$-module of the trivial representation of G. More explicitly, it is a one-dimensional vector space over $\Bbbk$ spanned by an element ε, together with the left action of $\Bbbk[G]$ given by the rule

$$g \cdot \varepsilon = \varepsilon$$

for all $g \in G$. The following proposition is a direct corollary of (6).

Proposition 2.10 *The left module V_n is isomorphic to $\Phi^{\otimes n} \otimes [\varepsilon]$. The isomorphism is the $\Bbbk$-linear extension of the map $\delta_{x_1 x_2 \cdots x_n} \mapsto x_1 \otimes x_2 \otimes \cdots \otimes x_n \otimes \varepsilon$ for $x_i \in X$.*

Denote by $\mathbf{1}$ the function $\sum_{x \in X} \delta_x \in V_1$ taking constant value $1 \in \Bbbk$. We have then, for every $f \in V_n = \Bbbk^{X^n}$,

$$f \otimes \mathbf{1}(x_1 x_2 \cdots x_{n+1}) = f(x_1 x_2 \cdots x_n).$$

The following proposition is straightforward.

Proposition 2.11 *The map $\iota_n : v \mapsto v \otimes \mathbf{1} : V_n \longrightarrow V_{n+1}$ is an embedding of the left $\Bbbk[G]$-modules. In other words,*

$$\pi_{n+1}(g)(f \otimes \mathbf{1}) = \pi_n(g)(f) \otimes \mathbf{1}$$

for all $g \in G$ and $f \in V_n$.

The space $X^\omega = \{x_1 x_2 \cdots : x_i \in X\}$ has a natural topology of a direct (Tikhonoff) power of the discrete space X. A basis of this topology consists of the cylindrical sets $v X^\omega$, for $v \in X^*$.

Denote by $C(X^\omega, \Bbbk)$ the vector space of maps $f : X^\omega \longrightarrow \Bbbk$ such that $f^{-1}(a)$ is open and closed (*clopen*) for every $a \in \Bbbk$. In other words, $C(X^\omega, \Bbbk)$ is the space of all continuous maps $f : X^\omega \longrightarrow \Bbbk$, where $\Bbbk$ is taken with discrete topology. Note that the set of values of any element of $C(X^\omega, \Bbbk)$ is finite, since X^ω is compact.

For example, a map $f : X^\omega \longrightarrow \mathbb{R}$ belongs to $C(X^\omega, \mathbb{R})$ if and only it is continuous and has a finite set of values.

The group G acts naturally on X^ω by homeomorphisms, hence it also acts naturally on the space $C(X^\omega, \Bbbk)$ by the rule

$$g(\xi)(w) = \xi(g^{-1}(w))$$

for $g \in G$, $\xi \in C(X^\omega, \Bbbk)$, and $w \in X^\omega$.

For every $f \in V_n = \Bbbk^{X^n}$ consider the natural extension of $f : X^n \longrightarrow \Bbbk$ to a function on X^ω:

$$f(x_1 x_2 \cdots) = f(x_1 x_2 \cdots x_n).$$

For example, the delta-function δ_v is extended to the characteristic functions of the subset $v X^\omega$, which we will also denote δ_v.

It is easy to see that this defines an embedding of the $\Bbbk[G]$-modules $V_n \longrightarrow C(X^\omega, \Bbbk)$. Moreover, these embeddings agree with the embeddings $\iota_n : V_n \longrightarrow V_{n+1}$.

Denote by V_∞ the direct limit of the G-modules V_n with respect to the maps ι_n. We will denote by π_∞ the corresponding representation of G on V_∞.

Proposition 2.12 *The module V_∞ is naturally isomorphic to the left $\Bbbk[G]$-module $C(X^\omega, \Bbbk)$.*

Proof The set $\{f^{-1}(t) \ : \ t \in \Bbbk, f^{-1}(t) \neq \emptyset\}$ is a finite covering of X^ω by clopen disjoint sets. Every clopen set of X^ω is a finite union of cylindrical sets of the form vX^ω, for $v \in X^*$. Consequently, there exists n such that f is constant on every cylindrical set of the form vX^ω for $v \in X^n$. Then $f \in \Bbbk^{X^n}$ in the identification of $\Bbbk^{X^n}$ with a subspace of $C(X^\omega, \Bbbk)$, described above. It follows that the inductive limit of $\Bbbk^{X^n}$ coincides with $C(X^\omega, \Bbbk)$. We have already seen that the representations π_n agree with the representation of G on $C(X^\omega, \Bbbk)$, restricted to $V_n = \Bbbk^{X^n}$) which finishes the proof. $\qquad\square$

Let B be a basis of the $\Bbbk$-vector space $\Bbbk^X$ such that the constant one function $\mathbf{1}$ belongs to B. Then $\mathsf{B}^{\otimes n}$ is a basis of the $\Bbbk$-vector space $\Bbbk^{X^n} = V_n$, and we have $\iota_n(\mathsf{B}^{\otimes n}) \subset \mathsf{B}^{\otimes n+1}$. Then the inductive limit B_∞ of the bases $\mathsf{B}^{\otimes n}$ with respect to the maps ι_n is a basis of $C(X^\omega, \Bbbk) = V_\infty$. The elements of this basis are equal to functions of the form

$$f(x_1 x_2 \cdots) \mapsto f_1(x_1) f_2(x_2) \cdots,$$

where $f_i \in \mathsf{B}$ and all but a finite number of the functions f_i are equal to the constant one.

Example 2.13 Suppose that the field $\Bbbk \cong \mathbb{F}_q$ is finite, and let $X = \Bbbk$. Then the functions $e_k : X \longrightarrow \Bbbk : x \mapsto x^k$ for $k = 1, 2, \ldots, q-1$ together with the constant one function $\mathbf{1}$, formally denoted x^0, form a basis of V_1.

The corresponding basis of $C(X^\omega, \Bbbk)$ is equal to the set of all monomial functions

$$f(x_1, x_2, \ldots) = x_1^{k_1} x_2^{k_2} \cdots,$$

where all but a finite number of powers k_i are equal to zero.

Writing the elements of $C(X^\omega, \Bbbk)$ in this basis amounts to representing them as polynomials.

Example 2.14 Let $X = \{x_0, x_1\}$, char $\Bbbk \neq 2$, and let W be the basis of $\Bbbk^X$ consisting of the functions $y_0 = \delta_{x_0} + \delta_{x_1} = \mathbf{1}$ and $y_1 = \delta_{x_0} - \delta_{x_1}$. The corresponding basis W_∞ of $C(X^\omega, \Bbbk)$ is called the *Walsh basis*, see [51].

For $\Bbbk = \mathbb{C}$, the Walsh basis is an orthonormal set of complex-valued functions on X^ω with respect to the uniform Bernoulli measure on X^ω. This is a direct corollary of the fact that $\{y_0, y_1\}$ is orthonormal. Since W_∞ is a basis of the linear space of continuous functions $X^\omega \longrightarrow \mathbb{C}$ with finite sets of values, and this space is dense in the Hilbert space $L^2(X^\omega)$, the Walsh basis is an orthonormal basis of $L^2(X^\omega)$.

We can use Proposition 2.8 to find transition matrices from $\{\delta_v\}_{v \in X^n}$ to the basis $\mathsf{W}^{\otimes n}$ (just use the proposition for the case of the trivial group G). In the case of Walsh basis we get the matrices from Example 2.9, but without $1/\sqrt{2}$:

$$H_n = \begin{pmatrix} H_{n-1} & H_{n-1} \\ H_{n-1} & -H_{n-1} \end{pmatrix},$$
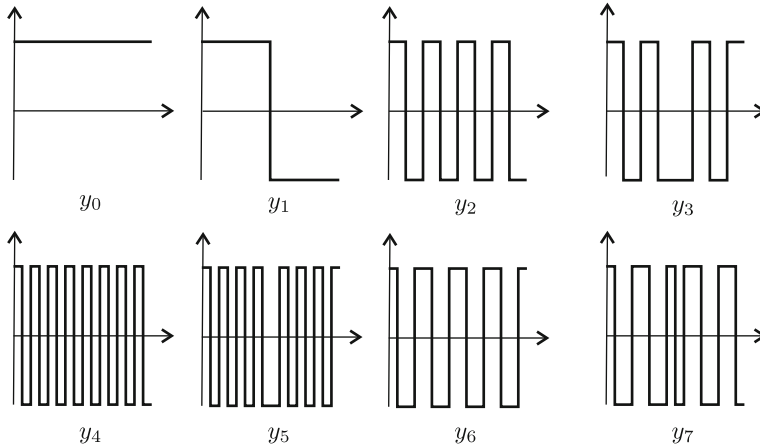
Fig. 1 Walsh basis

compare with Example 2.9. These matrices are examples of Hadamard matrices (i.e., matrices whose entries are +1 and −1 and whose rows are orthogonal) and were constructed for the first time by Sylvester [46]. They are also called *Walsh matrices*.

See Fig. 1, where graphs of the first eight elements of the Walsh basis are shown. Here we identify $\{0, 1\}^\omega$ with the unit interval [0, 1] via real binary numeration system.

Example 2.15 A related basis of $C(X^\omega, \Bbbk)$ is the *Haar basis*, which is constructed in the following way. Again, we assume that characteristic of $\Bbbk$ is different from 2, and $X = \{x_0, x_1\}$. Let $y_0 = \mathbf{1}$ and $y_1 = \delta_{x_0} - \delta_{x_1}$, as in the previous example. Let us construct an increasing sequence of bases Y_n of $\Bbbk^{X^n} < C(X^\omega, \Bbbk)$ in the following way. Let $\mathsf{Y}_0 = \{y_0\}$. Define then inductively:

$$\mathsf{Y}_{n+1} = \mathsf{Y}_n \cup \{\delta_v \otimes y_1 \; : \; v \in X^n\}.$$

Note that, since $\{\delta_v \; : \; v \in X^n\}$ is a basis of $\Bbbk^{X^n}$, the set $\{\delta_v \otimes y_0 \; : \; v \in X^n\} \cup \{\delta_v \otimes y_1 \; : \; v \in X^n\}$ is a basis of $\Bbbk^{X^{n+1}}$. But $\{\delta_v \otimes y_0 \; : \; v \in X^n\} = \{\delta_v \; : \; v \in X^n\}$ is a basis of $\Bbbk^{X^n}$, since $y_0 = \mathbf{1}$. Consequently, Y_{n+1} is a basis of $\Bbbk^{X^{n+1}}$. (Here everywhere $\Bbbk^{X^n}$ are identified with the corresponding subspaces of $C(X^\omega, \Bbbk)$.)

In the case $\Bbbk = \mathbb{C}$, and identification of $C(X^\omega, \mathbb{C})$ with a linear subspace of $L^2(X^\omega, \mu)$, where μ is the uniform Bernoulli measure on X^ω, it makes sense to normalize the elements of Y_n in order to make them of norm one. Since norm of δ_v is equal to $2^{-n/2}$, the recurrent definition of the basis in this case is

$$\mathsf{Y}_{n+1} = \mathsf{Y}_n \cup \left\{ 2^{-n/2} \otimes y_1 \; : \; v \in X^n \right\}.$$

It is easy to see that the union Y_∞ of the bases Y_n is an orthonormal basis of $L^2(X^\omega, \mu)$. It is called the *Haar basis*. See its use in the context of groups acting on rooted trees in [6].

3 Automata

3.1 Mealy and Moore automata

Definition 3.1 A *Mealy automaton* (or Mealy machine) is a tuple

$$\mathfrak{A} = (Q, X, Y, \pi, \tau, q_0),$$

where

- Q is the set of *states* of the automaton;
- X and Y are the *input* and *output alphabets* of the automaton;
- $\pi : Q \times X \longrightarrow Q$ is the *transition map*;
- $\tau : Q \times X \longrightarrow X$ is the *output map*;
- $q_0 \in Q$ is the *initial state*.

We always assume that X and Y are finite and have more than one element each.

We frequently assume that $X = Y$, and say that the automaton is *defined over the alphabet X*. The automaton is *finite* if the set Q is finite. In some cases, we do not assume that an initial state is chosen.

Let $\mathfrak{A} = (Q, X, Y, \pi, \tau, q_0)$ be a Mealy automaton. Let us extend the definition of the maps π and τ to maps $\pi : Q \times X^* \longrightarrow Q$ and $\tau : Q \times X^* \longrightarrow X$ by the inductive rules

$$\pi(q, xv) = \pi(\pi(q, x), v), \qquad \tau(q, xv) = \tau(\pi(q, x), v)).$$

We interpret the automaton $\mathfrak{A}$ as a machine, which being in a state $q \in Q$ and reading a letter $x \in X$, goes to the state $\pi(q, x)$, and gives the letter $\tau(q, x)$ on the output. If the machine starts at the state $q \in Q$, and reads a word v, then its final state will be $\pi(q, v)$, and its final letter on output will be $\tau(q, v)$.

Definition 3.2 The transformation $\mathfrak{A}_{q_0} : X^* \longrightarrow X^*$ or $\mathfrak{A}_{q_0} : X^\omega \longrightarrow X^\omega$ defined by a Mealy automaton $\mathfrak{A} = (Q, X, Y, \pi, \tau, q_0)$ is the map

$$\mathfrak{A}_{q_0}(x_0 x_1 x_2 \cdots) = \tau(q_0, x_0)\tau(q_1, x_1)\tau(q_2, x_2) \ldots, \tag{8}$$

where $q_{i+1} = \pi(q_i, x_i)$.

In other words, $\mathfrak{A}_{q_0}(v)$ is the word that the machine gives on output, when it reads the word v on input, if q_0 is its initial state.

Example 3.3 Let $G \leq \mathrm{Aut}(X^*)$ be a self-similar group. Consider the corresponding *full automaton* with the set of states $Q = G$, and output and transition functions defined by the rules:

$$\pi(g, x) = g|_x, \qquad \tau(g, x) = g(x).$$

It follows from (4) that if we choose $g \in G$ as the initial state, then the transformations of X^* and X^{ω} defined by this automaton coincides with the original transformations defined by $g \in \mathrm{Aut}(X^*)$.

This automaton is infinite, but if $G \leq \mathrm{FAut}(X^*)$, then for every $g \in G$, the set $\{g|_v : v \in X^*\}$ is a finite set, and we can take it as a set of states of a finite automaton defining the transformation g.

A special type of Mealy automata are the *Moore automata*. The definition of a Moore automaton is the same as Definition 3.1, except that the output function is a map $\tau : Q \longrightarrow X$, i.e., the output depends only on the state, and does not depend on the input letter.

Moore automata also act on words, essentially in the same way as Mealy automata. We can extend the definition of the transition function π to $Q \times X^*$ by the same formula as for the Mealy automata. Then the action of a Moore automaton with initial state q_0 on words is given by the rule

$$\mathfrak{A}_{q_0}(x_1 x_2 \cdots) = \tau(q_1)\tau(q_2)\ldots, \tag{9}$$

where $q_{i+1} = \pi(q_i, x_{i+1})$.

Even though the definition of a Moore automaton seems to be more restrictive than the definition of a Mealy automaton, the two notions are basically equivalent, as any Mealy automaton can be modeled by a Moore automaton. Hence, the set of maps defined by finite Mealy automata coincides with the set of maps defined by finite Moore automata.

Let $\mathfrak{A} = (Q, X, Y, \pi, \tau, q_0)$ be a Mealy automaton. Consider the Moore automaton $\mathfrak{A}'$ over the input and output alphabets X and Y, respectively, with the set of states $Q \times X \cup \{p_0\}$, where p_0 is an element not belonging to $Q \times X$, and with the transition and output maps π' and τ' given by the rules

$$\pi'(p_0, x) = (\pi(q_0, x), x), \quad \pi'((q, x_1), x_2)) = (\pi(q, x_2), x_2)),$$

and

$$\tau'(q, x_1) = \tau(q, x_1),$$

where $x, x_1, x_2 \in X$ (We define $\tau'(p_0)$ to be any letter, since it will never appear in the output). It is easy to check that the new Moore automaton with the initial state p_0 defines the same maps on X^* and X^{ω} as the original Mealy automaton $\mathfrak{A}$.

Therefore, we will not use Moore automata to define transformations of the sets of words. They will be used to define automatic sequences and matrices in Sect. 4. Traditionally, Mealy automata are used in theory of groups generated by automata (see [19]), while Moore automata are used for generation of sequences (even though the term "Moore automata" is not used in [2]).
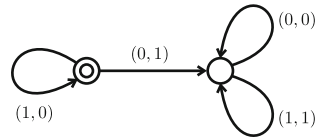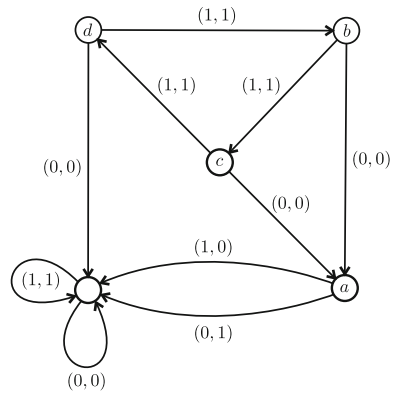
Fig. 2 The binary adding machine

Fig. 3 The generators of the Grigorchuk group

3.2 Diagrams of automata

The automata are usually represented as labeled graphs (called *Moore diagrams*). The set of vertices coincides with the set of states Q. For every $q \in Q$ and $x \in X$ there is an arrow from q to $\pi(q, x)$ labeled by $(x, \tau(q, x))$ in the case of Mealy automata, and just by x in the case of Moore automata. The initial state is marked, and the states are marked by the values of $\tau(q)$, if it is a Moore automaton.

Sometimes the arrows of diagrams of Mealy automata are just labeled by the input letters x, and the vertices are labeled by the corresponding transformation $x \mapsto \tau(q, x)$.

Consider a directed graph with one marked (initial) vertex, in which the edges are labeled by pairs $(x, y) \in X^2$. The necessary and sufficient condition for such a graph to represent a Mealy automaton is that for every vertex q and every letter $x \in X$ there exists a unique arrow starting at q and labeled by (x, y) for some $y \in X$. Then the image of a word $x_1 x_2 \cdots$ under the action of the automaton is calculated by finding the unique direct path $e_1, e_2, \ldots$ of arrows starting at the initial vertex, whose arrows are labeled by $(x_1, y_1), (x_2, y_2), \ldots$, respectively. Then $y_1 y_2 \cdots$ is the image of $x_1 x_2 \cdots$.

The diagram of the adding machine transformation (see Example 2.3) is shown on Fig. 2. We mark the initial state by a double circle.

Example 3.4 The generators a, b, c, d are defined by one automaton, shown on Fig. 3, for different choices of the initial state.

Fig. 4 Appending and erasing letters
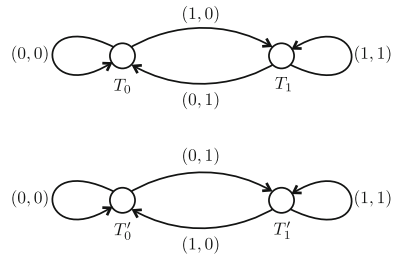
3.3 Non-deterministic automata

Let us generalize the notion of a Mealy automaton by allowing more general Moore diagrams.

Definition 3.5 A (*non-deterministic*) *synchronous* automaton $\mathfrak{A}$ over an alphabet X is an oriented graph whose arrows are labeled by pairs of letters $(x, y) \in X^2$. Such automaton is called ω-*deterministic* if for every infinite word $x_1 x_2 \cdots \in X^\omega$ and for every vertex (i.e., *state*) q of $\mathfrak{A}$ there exists at most one directed path starting in q which is labeled by $(x_1, y_1), (x_2, y_2), \ldots$ for some $y_i \in X$.

Note that in the above definition for a vertex state q of $\mathfrak{A}$ and a letter $x \in X$ there maybe several or no edges starting at q and labeled by (x, y) for $y \in X$. It means that the automaton $\mathfrak{A}$ may be *non-deterministic* on finite words and *partial*, i.e., that a state q transforms a finite word $v \in X^*$ into several different words, and may not accept some of the words on input.

If an automaton $\mathfrak{A}$ is ω-deterministic, then every its state q defines a map between closed subsets of X^ω, mapping $x_1 x_2 \cdots$ to $y_1 y_2 \cdots$, if there exists a directed path starting in q and labeled by $(x_1, y_1), (x_2, y_2), \ldots$.

Example 3.6 Let $X = \{0, 1\}$. The states T_0 and T_1 of the automaton shown on Fig. 4 define the transformations $x_1 x_2 \cdots \mapsto 0 x_1 x_2 \cdots$ and $x_1 x_2 \cdots \mapsto 1 x_1 x_2 \cdots$, respectively. The states T_0' and T_1' define the inverse transformations $0 x_1 x_2 \cdots \mapsto x_1 x_2 \cdots$ and $1 x_1 x_2 \cdots \mapsto x_1 x_2 \cdots$.

Note that the first automaton (defining the transformations T_0 and T_1) is deterministic. For example, the state T_0 acts on the finite words by transformations $x_1 x_2 \cdots x_n \mapsto 0 x_1 x_2 \cdots x_{n-1}$. The second automaton is partial and non-deterministic on finite words. For example, there are two arrows starting at T_0' labeled by $(0, 1)$ and $(0, 0)$, but no arrows labeled by $(1, y)$.

An *asynchronous* automaton is defined in the same way, but the labels are pairs of arbitrary words $(u, v) \in (X^*)^2$.

Definition 3.7 A homeomorphism $\phi : X^\omega \longrightarrow X^\omega$ is *synchronously* (resp. *asynchronously*) *automatic* if it is defined by a finite ω-deterministic synchronous (resp. asynchronous) automaton.

A criterion for a homeomorphism to be synchronously automatic is given in Proposition 4.23.

Asynchronously automatic homeomorphisms of X^ω are studied in [18,19]. It is shown there that the set of asynchronously automatic homeomorphisms of X^ω is a group, and that it does not depend on X (if $|X| > 1$). More precisely, it is proved that for any two finite alphabets X, Y (such that $|X|, |Y| > 1$) there exists a homeomorphism $X^\omega \longrightarrow Y^\omega$ conjugating the corresponding groups of asynchronously automatic homeomorphisms. Very little is known about this group, which is called in [19] the *group of rational homeomorphisms of the Cantor set*.

4 Automatic matrices

4.1 Automatic sequences

Here we review the basic definitions and facts about automatic sequences. More can be found in the monographs [2,50].

Let A be a finite alphabet, and let A^ω be the space of the right-infinite sequence of elements of A with the direct product topology.

Fix an integer $d \geq 2$, and consider the transformation $\Xi_d : A^\omega \longrightarrow (A^\omega)^d$, which we called the *stencil map*, defined by the rule

$$\Xi_d(a_0 a_1 a_2 \cdots) = (a_0 a_d a_{2d} \ldots, a_1 a_{d+1} a_{2d+1} \ldots, \ldots, a_{d-1} a_{2d-1} a_{3d-1} \cdots).$$

It is easy to see that Ξ_d is a homeomorphism. We denote the coordinates of $\Xi_d(w)$ by $\Xi_d(w)_i$, so that

$$\Xi_d(w) = (\Xi_d(w)_0, \Xi_d(w)_1, \ldots, \Xi_d(w)_{d-1}),$$

and call them *d-decimations* of the sequence w. *Repeated d-decimations* of w are all sequences that can be obtained from w by iterative application of the decimation procedure, i.e., all sequences of the form

$$\Xi_d(\Xi_d(\cdots \Xi_d(w)_{i_n} \cdots)_{i_2})_{i_1}.$$

Definition 4.1 A sequence $w \in A^\omega$ is *d-automatic* if the set of all repeated *d*-decimations of w (called the *kernel* of w in [2, Section 6.6]) is finite.

We say that a subset $Q \subset A^\omega$ is *d-decimation-closed* if for every $w \in Q$ all *d*-decimations of w belong to Q. The following is obvious.

Lemma 4.2 *A sequence is d-automatic if it belongs to a finite d-decimation-closed subset of A^ω.*

Classically, a sequence $w = a_0 a_1 \cdots$ is called *d*-automatic if there exists a Moore automaton $\mathfrak{A}$ with input alphabet $\{0, 1, \ldots, d - 1\}$ and output alphabet A such that if $n = i_0 + i_1 d + \cdots + i_m d^m$ is a base d expansion of n, then the output of $\mathfrak{A}$ after reading the word $i_0 i_1 \cdots i_m$ is a_n. An equivalent variant of the definition requires that a_n is the output of the automaton after reading $i_m i_{m-1} \cdots i_1 i_0$. One also may allow,

Fig. 5 Automaton generating
the Thue–Morse sequence
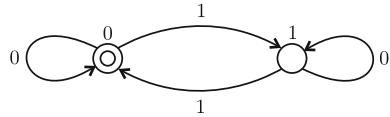
or not i_m to be equal to zero, and the numeration of the letters of the sequence w to start from 1. All these different definitions of automaticity of sequences are equivalent to each other, see [2, Section 5.2]. They are also equivalent to Definition 4.1, see [2, Theorem 6.6.2].

Example 4.3 The *Thue–Morse sequence* is the sequence $t_0 t_1 \cdots \in \{0, 1\}^\omega$, where t_n is the sum modulo 2 of the digits of n in the binary numeration system. The beginning of length 2^n of this sequence can be obtained from 0 by applying the substitution

$$0 \mapsto 01, \quad 1 \mapsto 10$$

n times:

$$0 \mapsto 01 \mapsto 0110 \mapsto 01101001 \mapsto 0110100110010110 \mapsto \cdots$$

It is easy to see that this sequence is generated by the automaton shown on Fig. 5. Here we label the vertices (the states) of the automaton by the corresponding values of the output function. The initial state is marked by a double circle. For more on properties of the Thue–Morse sequence, see [2, 5.1].

The last example can be naturally generalized to include all automatic sequences. Namely, a *k-uniform morphism* $\phi : X^* \longrightarrow Y^*$ is a morphism of monoids such that $|\phi(x)| = k$ for every $x \in X$. By a theorem of Combham (see [2, Theorem 6.3.2]) a sequence is k-automatic if and only if it is an image, under a coding (i.e., a 1-uniform morphism), of a fixed point of a k-uniform endomorphism $\phi : X^* \longrightarrow X^*$.

Example 4.4 Consider the alphabet $X = \{a, b, c\}$, and the morphism $\phi : X^* \longrightarrow X^*$ given by

$$\phi(a) = aca, \quad \phi(b) = d, \quad \phi(c) = b, \quad \phi(d) = c.$$

This substitution appears in the presentation [32] of the Grigorchuk group.

The fixed point of ϕ is obtained as the limit of $\phi^n(a)$, and starts with $acabacadacabaca \ldots$. The morphism ϕ is not uniform, but it is easy to see that the fixed point belongs to $\{ab, ac, ad\}^\infty$, and for the words $B = ab$, $C = ac$, $D = ad$ it acts on $\{B, C, D\}^\infty$ as a 2-uniform endomorphism:

$$\phi(B) = acad = CD, \quad \phi(C) = acab = CB, \quad \phi(D) = acac = CC.$$

It follows from Combham's theorem that the fixed point of ϕ is 2-automatic.

Let us show how to construct an automaton producing a sequence satisfying the conditions of Definition 4.1.

Suppose that $w_0 \in A^\omega$ is automatic, and let Q be a finite d-decimation-closed subset of A^ω that contains w_0 (for example, we can take Q to be equal to the set of all repeated d-decimations of w_0).

Consider a Moore automaton with the set of states Q, initial state w_0, input alphabet $\{0, 1, \ldots, d-1\}$, output alphabet A, transition function

$$\pi(w, i) = \Xi_d(w)_i,$$

output function

$$\tau(x_0 x_1 \cdots) = x_0.$$

We call the constructed Moore automaton

$$\mathfrak{A} = (Q, \{0, 1, \ldots, d-1\}, A, \pi, \tau, w_0)$$

the *automaton of* w_0.

Proposition 4.5 *Let $w_0 = a_0 a_1 \cdots \in A^\omega$ be an automatic sequence, and let $\mathfrak{A}$ be its automaton. Let n be a non-negative integer, and let $i_0, i_1, \ldots, i_m$ be a sequence of elements of the set $\{0, 1, \ldots, d-1\}$ such that $n = i_0 + i_1 d + i_2 d^2 + \cdots + i_m d^m$. Then $\tau(w_0, i_0 i_1 \cdots i_m) = a_n$, i.e., the output of $\mathfrak{A}$ after reading $i_0 i_1 \cdots i_m$ is a_n.*

Proof It follows from the definition of the automaton $\mathfrak{A}$ that

$$\pi(w, i_0 \cdots i_m) = \Xi_d(\cdots \Xi_d(\Xi_d(x_0 x_1 \cdots)_{i_0})_{i_1} \cdots)_{i_m} \qquad (10)$$

for all $w = x_0 x_1 \cdots \in Q$.

It also follows from the definition of the stencil map that the sequence (10) is equal to

$$x_n x_{n+d^{m+1}} x_{n+2d^{m+1}} x_{n+3d^{m+1}} \cdots,$$

where $n = i_0 + i_1 d + i_2 d^2 + \cdots + i_m d^m$. It follows that $\tau(w_0, i_0 i_1 \cdots i_m) = x_n$. $\qquad \square$

4.2 Automatic infinite matrices

The notion of automaticity of sequences can be generalized to matrices in a straightforward way (see [2, Chapter 14], where they are called *two-dimensional sequences*).

Let A be a finite alphabet, and let $A^{\omega \times \omega}$ be the space of all infinite to right and down two-dimensional matrices of elements of A, i.e., arrays of the form

$$a = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}. \qquad (11)$$

Fig. 6 The stencil

$$
\begin{array}{cccccccc}
a_{00} & \boxed{a_{01}} & a_{02} & \boxed{a_{03}} & a_{04} & \boxed{a_{05}} & \cdots & \square \\
a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & \cdots & \\
a_{20} & \boxed{a_{21}} & a_{22} & \boxed{a_{23}} & a_{24} & \boxed{a_{25}} & \cdots & \square \\
a_{30} & a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & \cdots & \\
a_{40} & \boxed{a_{41}} & a_{42} & \boxed{a_{43}} & a_{44} & \boxed{a_{45}} & \cdots & \square \\
a_{50} & a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & \cdots & \\
\vdots & \boxed{\vdots} & \vdots & \boxed{\vdots} & \vdots & \boxed{\vdots} & \ddots & \square
\end{array}
$$
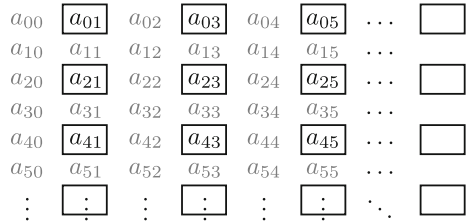
Fix an integer $d \geq 2$, and consider the map:

$$
\Xi_d : A^{\omega \times \omega} \longrightarrow (A^{\omega \times \omega})^{d \times d}
$$

from the set of infinite matrices to the set of $d \times d$ matrices whose entries are elements of $A^{\omega \times \omega}$. It maps the matrix a to the matrix

$$
\Xi(a) =
\begin{pmatrix}
\Xi_d(a)_{00} & \Xi_d(a)_{01} & \cdots & \Xi_d(a)_{0,d-1} \\
\Xi_d(a)_{10} & \Xi_d(a)_{11} & \cdots & \Xi_d(a)_{1,d-1} \\
\vdots & \vdots & \ddots & \vdots \\
\Xi_d(a)_{d-1,0} & \Xi_d(a)_{d-1,1} & \cdots & \Xi_d(a)_{d-1,d-1}
\end{pmatrix},
$$

where

$$
\Xi_d(a)_{ij} =
\begin{pmatrix}
a_{i,j} & a_{i,j+d} & a_{i,j+2d} & \cdots \\
a_{i+d,j} & a_{i+d,j+d} & a_{i+d,j+2d} & \cdots \\
a_{i+2d,j} & a_{i+2d,j+d} & a_{i+2d,j+2d} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix}.
$$

The entries of $\Xi_d(a)$ are called d-decimations of a. We call Ξ_d the *stencil map*, since entries of the matrix $\Xi_d(A)$ are obtained from the matrix A by selecting entries using a "stencil" consisting of a square grid of holes, see Fig. 6.

The definition of automaticity for matrices is then the same as for sequences.

Definition 4.6 A matrix $a \in A^{\omega \times \omega}$ is *d-automatic* ([d, d]-automatic in terminology of [2]) if the set of matrices that can be obtained from a by repeated d-decimations is finite.

One can also use stencils with a rectangular grid of holes, i.e., selecting the entries of a decimation with one step horizontally, and with a different step vertically. This will lead us to the notion of a [d_1, d_2]-automatic matrix, as in [2], but we do not use this notion in our paper.

An interpretation of automaticity of matrices via automata theory is also very similar to the interpretation for sequences. The only difference is that the input alphabet of the automaton is the direct product $\{0, 1, \ldots, d-1\} \times \{0, 1, \ldots, d-1\}$. If we want

to find an entry a_{n_1,n_2} of an automatic matrix defined by a Moore automaton $\mathfrak{A}$, then we represent the indices n_1 and n_2 in base d:

$$n_1 = i_0 + i_1 d + \cdots + i_m d^m, \qquad n_2 = j_0 + j_1 d + \cdots + j_m d^m$$

for $0 \leq i_s, j_t \leq d - 1$, and then feed the sequence $(i_0, j_0)(i_1, j_1) \cdots (i_m, j_m)$ to the automaton $\mathfrak{A}$. Its final output will be a_{n_1,n_2}.

We say that a matrix $a = (a_{ij})_{i \geq 0, j \geq 0}$ over a field $\Bbbk$ is *column-finite* if the number of non-zero entries in each column of a is finite. The set $\mathsf{M}_\infty(\Bbbk)$ of all column-finite matrices is an algebra isomorphic to the algebra of endomorphisms of the infinite-dimensional vector space $\Bbbk^\infty = \bigoplus_{\mathbb{N}} \Bbbk$. We denote by $\mathsf{M}_k(\Bbbk)$ the algebra of $k \times k$-matrices over $\Bbbk$.

Lemma 4.7 *The stencil map $\Xi_d : \mathsf{M}_\infty(\Bbbk) \longrightarrow \mathsf{M}_d(\mathsf{M}_\infty(\Bbbk))$ is an isomorphism of $\Bbbk$-algebras.*

Proof A direct corollary of the multiplication rule for matrices. $\qquad\square$

Denote by $\mathcal{A}_d(\Bbbk) \subset \mathsf{M}_\infty(\Bbbk)$ the set of all column-finite d-automatic matrices over $\Bbbk$.

Proposition 4.8 *Let $\Bbbk$ be a finite field. Then $\mathcal{A}_d(\Bbbk)$ is an algebra. The stencil map*

$$\Xi_d : \mathcal{A}_d(\Bbbk) \longrightarrow \mathsf{M}_d(\mathcal{A}_d(\Bbbk))$$

is an isomorphism of $\Bbbk$-algebras.

Proof Let A and B be d-automatic column-finite matrices. Let $\mathfrak{A}$ and $\mathfrak{B}$ be finite decimation-closed sets containing A and B, respectively. Then the set $\{a_1 A_1 + b_1 B_1 : A_1 \in \mathfrak{A}, B_1 \in \mathfrak{B}, a_1, b_1 \in \Bbbk\}$ is decimation-closed, and it contains any linear combination of A and B. The set is finite, which shows that any linear combination of A and B is automatic.

Let $\mathfrak{C}$ be the linear span of all products $A_1 B_1$ for $A_1 \in \mathfrak{A}, B_1 \in \mathfrak{B}$. Since the stencil map $\Xi_d : \mathsf{M}_\infty(\Bbbk) \longrightarrow \mathsf{M}_d(\mathsf{M}_\infty(\Bbbk))$ is an isomorphism of $\Bbbk$-algebras, the set $\mathfrak{C}$ is decimation-closed. It contains AB, hence AB is automatic.

The map $\Xi_d : \mathcal{A}_d(\Bbbk) \longrightarrow \mathsf{M}_d(\mathcal{A}_d(\Bbbk))$ is obviously a bijection. It is a homomorphism of $\Bbbk$-algebras, because it is a homomorphism on $\mathsf{M}_\infty(\Bbbk)$. $\qquad\square$

If we have a finite d-decimation-closed set of matrices $\mathfrak{A}$, then its elements are uniquely determined by the corresponding matrix recursion $\Xi_d : \mathfrak{A} \longrightarrow \mathsf{M}_d(\mathfrak{A})$ and by the top left entries a_{00} of each matrix $A \in \mathfrak{A}$. Namely, suppose that we want to find an entry $a_{m,n}$ of a matrix $A \in \mathfrak{A}$. Let $0 \leq i < d$ and $0 \leq j < d$ be the remainders of division of n and m by d. Then $a_{m,n}$ is equal to the entry $b_{\frac{m-i}{d}, \frac{n-j}{d}}$ of the matrix $B = \Xi_d(A)_{i,j}$. Repeating this procedure several times, we eventually will find a matrix $C = (c_{i,j})_{i,j=0}^\infty \in \mathfrak{A}$ such that $a_{m,n} = c_{0,0}$.

Example 4.9 A particular case of automatic matrices are triangular matrices of the form

$$
\begin{pmatrix}
a_{00} & a_{01} & a_{02} & \cdots \\
0 & a_{11} & a_{12} & \cdots \\
0 & 0 & a_{22} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
$$

where the diagonals $a_{0,i}, a_{1,i+1}, a_{2,i+2}, \ldots$ are eventually periodic, and only a finite number of them are non-zero. Note that the set of such uni-triangular matrices is a group.

The following subgroup of this group (of matrices over the field $\mathbb{F}_2$) was considered in [38]. Let $B_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$, $C_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $B_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}$, $C_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. It is shown in [38] that the infinite matrices

$$
F_1 = \begin{pmatrix}
I & B_1 & C_1 & O & O & \cdots \\
O & I & B_1 & C_1 & O & \cdots \\
O & O & I & B_1 & C_1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}, \quad
F_2 = \begin{pmatrix}
I & B_2 & C_2 & O & O & \cdots \\
O & I & B_2 & C_2 & O & \cdots \\
O & O & I & B_2 & C_2 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
$$

generate a free group of rank 2. Here I and O are the identity and zero matrices of size 3×3, respectively.

In Sect. 5 we will show that any residually finite p-group can be represented in triangular form. Groups generated by finite automata will be represented by automatic uni-triangular matrices. The next subsection is the first step in this direction.

4.3 Representation of automata groups by automatic matrices

Let B be a basis of $\Bbbk^X$ such that $\mathbf{1} \in \mathsf{B}$. Order the elements of B into a sequence $y_0 < y_1 < \cdots < y_{d-1}$, where $y_0 = \mathbf{1}$. Recall that the inductive limit B_∞ of the bases $\mathsf{B}^{\otimes n}$ of $\Bbbk^{X^n}$ with respect to the embeddings $f \mapsto f \otimes \mathbf{1}$ is a basis of $C(X^\omega, \Bbbk)$, whose elements are infinite tensor products $y_{i_0} \otimes y_{i_1} \otimes \cdots$, where all but a finite number of factors y_{i_k} are equal to $y_0 = \mathbf{1}$. In other words, B_∞ consists of functions of the form

$$
f(x_0 x_1 \cdots) = y_{i_0}(x_0) y_{i_1}(x_1) \cdots,
$$

where all but a finite number of factors on the right-hand side are equal to the constant one function.

We can order such products using the *inverse* lexicographic order, namely $y_{i_0} \otimes y_{i_1} \otimes \cdots < y_{j_0} \otimes y_{j_1} \otimes \cdots$ if and only if $y_{i_k} < y_{j_k}$, where k is the *largest* index such that $y_{i_k} \neq y_{j_k}$.

It is easy to see that the ordinal type of B_∞ is ω. Let $e_0 < e_1 < e_2 < \cdots$ be all elements of B_∞ taken in the defined order. It is checked directly that if $e_n = y_{i_0} \otimes y_{i_1} \otimes \ldots$, then $n = i_0 + i_1 \cdot d + i_2 \cdot d^2 + \cdots$, i.e., $\overline{\ldots i_2 i_1 i_0}$ is the base-d expansion of n (only a finite number of coefficients i_j are different from zero).

Definition 4.10 An ordered basis B of $\Bbbk^X$ is called *marked* if its minimal element is **1**.

The *self-similar basis* B_∞ of $C(X^\omega, \Bbbk)$ *associated with* B is the inverse lexicographically ordered set of functions of the form $e_1 \otimes e_2 \otimes \ldots$, where $e_i \in B$ and all but a finite set of elements e_i are equal to **1**, as it is described above.

Let $B = \{b_0, b_1, \ldots, b_{d-1}\}$ be an arbitrary (non necessarily marked) basis of the space $\Bbbk^X$. We define the associated *matrix recursion* for linear operators on $C(X^\omega, \Bbbk)$ in the usual way: given an operator a, define its image $\Xi_B(a) = (A_{i,j})_{i,j=0}^{d-1}$ in $M_d(C(X^\omega, \Bbbk))$ by the rule

$$a(b_j \otimes f) = \sum_{i=0}^{d-1} b_i \otimes A_{i,j}(f), \qquad f \in C(X^\omega, \Bbbk), \quad 0 \le j \le d-1.$$

If B is the basis $\{\delta_x\}_{x \in X}$, then the matrix recursion Ξ_B restricted to a self-similar group $G \le \mathrm{Aut}(X^*)$ coincides with the matrix recursion (5) coming directly from the wreath recursion.

Lemma 4.11 *Let B be a marked basis of $\Bbbk^X$. Then the matrix recursion Ξ_B coincides with the stencil map for the matrices of linear operators in the associated basis B_∞.*

Proof Let $A_{ij}, 0 \le i, j \le d-1$, be the entries of $\Xi_B(a)$. Let $n = i_0 + i_1 \cdot d + i_2 \cdot d^2 + \cdots$ be a non-negative integer written in base d. Then

$$a(b_{j+dn}) = a(b_j \otimes b_n) = \sum_{i=0}^{d-1} b_i \otimes A_{i,j}(b_n). \tag{12}$$

Let $a_{m,n}, 0 \le m, n < \infty$, be the entries of the matrix of a in the basis B_∞. Then

$$a(b_{j+dn}) = \sum_{k=0}^{\infty} a_{k,j+dn} b_k = \sum_{i=0}^{d-1} \sum_{r=0}^{\infty} a_{i+dr,j+dn} b_i \otimes b_r$$
$$= \sum_{i=0}^{d-1} b_i \otimes \left(\sum_{r=0}^{\infty} a_{i+dr,j+dn} b_r \right),$$

which together with (12) implies that $c_{r,n} = a_{i+dr,j+dn}$ are the entries of $A_{i,j}$ in the basis B_∞, i.e., that $A_{i,j} = \Xi_d(a)_{i,j}$. $\square$

Definition 4.12 Let $\Bbbk$ be a finite field. We say that a linear operator a on $C(X^\omega, \Bbbk)$ is *automatic* if there exists a finite set $\mathfrak{A}$ of operators such that $a \in \mathfrak{A}$, and for every $a' \in \mathfrak{A}$ all entries of the matrix $\Xi_B(a')$ belong to $\mathfrak{A}$.

Proposition 4.13 *Let* B_1, B_2 *be two bases of* $\Bbbk^X$. *A linear operator on* $C(X^\omega, \Bbbk)$ *is automatic with respect to* B_1 *if and only if it is automatic with respect to* B_2.

Proof Let a_1 be an operator which is automatic with respect to B_1. Let $\mathfrak{A}$ be the corresponding finite set of operators, closed with respect to taking entries of the matrix recursion. Let $\mathfrak{A}'$ be the set of all linear combinations of elements of $\mathfrak{A}$, which is finite, since we assume that $\Bbbk$ is finite. If T is the transition matrix from B_1 to B_2, then

$$\Xi_{B_2}(a) = T^{-1} \Xi_{B_1}(a) T$$

for every linear operator a. It follows that $\mathfrak{A}'$ is closed with respect to taking entries of Ξ_{B_2}. The set $\mathfrak{A}'$ is finite, $a_1 \in \mathfrak{A}'$, hence a_1 is also automatic with respect to B_2. $\square$

As a direct corollary of Proposition 4.13 we get the following relation between finite-state automorphisms of the rooted tree X^* and automatic matrices.

Theorem 4.14 *Suppose that* $\Bbbk$ *is finite. Let* B *be a marked basis of* $\Bbbk^X$. *Then the matrix of* $\pi_\infty(g)$ *in the associated basis* B_∞, *where* $g \in \mathrm{Aut}(X^*)$, *is d-automatic if and only if g is finite-state.*

We get, therefore, a subgroup of the group of units of $\mathcal{A}_d(\Bbbk)$ isomorphic to the group $\mathrm{FAut}(X^*)$ of finite-state automorphisms of the tree X^*.

Matrix recursions (i.e., homomorphisms from an algebra A to the algebra of matrices over A) associated with groups acting on rooted trees, and in more general cases were studied in many papers, for instance in [3,4,6,34,37,42,43]. Note that the algebra generated by the natural representation on $C(X^\omega, \Bbbk)$ of a group G acting on the rooted tree X^* is different from the group ring. This algebra (and its analogs) were studied in [4,34,42].

4.4 Creation and annihilation operators

For $h \in \Bbbk^X$, denote by T_h the operator on $C(X^\omega, \Bbbk)$ acting by the rule

$$T_h(f) = h \otimes f.$$

It is easy to see that T_h is linear, and that we have $T_{a_1 h_1 + a_2 h_2} = a_1 T_{h_1} + a_2 T_{h_2}$ for all $h_1, h_2 \in \Bbbk^X$ and $a_1, a_2 \in \Bbbk$.

Consider the dual vector space $(\Bbbk^X)'$ to the space $\Bbbk^X$ of functions. We will denote the value of a functional v on a function $f \in \Bbbk^X$ by $\langle v | f \rangle$. Then for every $v \in (\Bbbk^X)'$ we have an operator T_v on $C(X^\omega, \Bbbk)$ defined by

$$T_v(f)(x_1, x_2, \ldots) = \langle v | f(x, x_1, x_2, \ldots) \rangle,$$

where $f(x, x_1, x_2, \ldots)$ on the right-hand side of the equation is seen for every choice of the variables $x_1, x_2, \ldots$ as a function of one variable x, i.e., an element of $\Bbbk^X$.

Let $\mathsf{B} = \{e_i\}_{i=0}^{d-1}$ be a basis of $\Bbbk^X$. Let $\mathsf{B}' = \{e_i'\}_{i=0}^{d-1}$ be the basis of the dual space defined by $\langle e_i'|e_j\rangle = \delta_{i,j}$. (Here and in the sequel, $\delta_{i,j}$ is the Kronecker's symbol equal to 1 when $i = j$, and to 0 otherwise.) We will denote $T_{e_i'} = T_{e_i}'$.

Then T_{e_i} is an isomorphism between the space $C(X^\omega, \Bbbk)$ and its subspace $e_i \otimes C(X^\omega, \Bbbk)$. It is easy to see that T_{e_i}' restricted onto $e_i \otimes C(X^\omega, \Bbbk)$ is the inverse of this isomorphism, and that T_{e_i}' restricted onto $e_j \otimes C(X^\omega, \Bbbk)$ is equal to zero for $j \neq i$.

The operators T_{e_i} and T_{e_j}' satisfy the relations:

$$T_{e_i}' T_{e_j} = \delta_{e_i, e_j}, \qquad \sum_{e_i \in \mathsf{B}} T_{e_i} T_{e_i}' = 1. \qquad (13)$$

The products $T_{e_i} T_{e_i}'$ are projections onto the summands $e_i \otimes C(X^\omega, \Bbbk)$ of the direct sum decomposition

$$C(X^\omega, \Bbbk) = \bigoplus_{e_i \in \mathsf{B}} e_i \otimes C(X^\omega, \Bbbk).$$

Let B be a marked basis of $\Bbbk^X$, and let $\mathbf{1} = e_0 < e_1 < \cdots < e_{d-1}$ be its elements. Let B_∞ be the associated ordered basis of $C(X^\omega, \Bbbk)$.

If A is (not necessarily square) finite matrix, then we denote by $A^{\oplus\infty}$ the infinite matrix

$$\begin{pmatrix} A & O & O & \cdots \\ O & A & O & \cdots \\ O & O & A & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where O is the zero matrix of the same size as A.

The following is a direct corollary of the definitions.

Proposition 4.15 *The matrices of* $T_{e_0}, T_{e_1}, \ldots, T_{e_{d-1}}$ *in the basis* B_∞ *are equal to*

$$E_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}^{\oplus\infty}, \quad E_1 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}^{\oplus\infty}, \ldots, \quad E_{d-1} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}^{\oplus\infty},$$

respectively. The matrix of T_{e_i}' *is the transpose* $E_i^\top$ *of the matrix* E_i.

The matrices E_i and E_i' have a natural relation to decimation of matrices. The proof of the next proposition is a straightforward computation of matrix products.

Proposition 4.16 *Let $A = (a_{i,j})_{i,j=0}^{\infty}$ be an infinite matrix, and let*

$$\Xi_d(A) = \begin{pmatrix} A_{0,0} & A_{0,1} & \cdots & A_{0,d-1} \\ A_{1,0} & A_{1,1} & \cdots & A_{1,d-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{d-1,0} & A_{d-1,1} & \cdots & A_{d-1,d-1} \end{pmatrix}$$

be the matrix of its d-decimations. Then

$$A_{i,j} = E_i^{\top} A E_j$$

and

$$A = \sum_{i,j=0}^{d-1} E_i A_{i,j} E_j^{\top}.$$

Corollary 4.17 *Let A be an operator on $C(X^{\omega}, \Bbbk)$, and let $\mathsf{B} = \{e_i\}$ be a basis of $\Bbbk^X$. Then the entries of the associated matrix recursion for A are equal to $T'_{e_i} A T_{e_j}$.*

The next proposition is a direct corollary of Proposition 4.15.

Proposition 4.18 *If $h = a_0 e_0 + a_1 e_1 + \cdots + a_{d-1} e_{d-1}$, then the matrix of T_h is equal to*

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{d-1} \end{pmatrix}^{\oplus \infty}.$$

Corollary 4.19 *Let B be a marked basis of $\Bbbk^X$. Order the letters of the alphabet X in a sequence $x_0, x_1, \ldots, x_{d-1}$. Let $T_i = T_{\delta_{x_i}}$ and $T'_i = T'_{\delta_{x_i}}$ be the corresponding operators, defined using the basis $\{\delta_{x_i}\}$.*

Let $S = (a_{ij})_{i,j=0}^{d-1}$ be the transition matrix from the basis $\delta_{x_0} < \delta_{x_1} < \cdots < \delta_{x_{d-1}}$ to B. Let $S^{-1} = (b_{ij})_{i,j=0}^{d-1}$ be the inverse matrix.

Then the matrix of T_i is

$$\begin{pmatrix} b_{0,i} \\ b_{1,i} \\ \cdots \\ b_{d-1,i} \end{pmatrix}^{\oplus \infty},$$

and the matrix of T'_i is

$$\begin{pmatrix} a_{i,0} & a_{i,1} & \cdots & a_{i,d-1} \end{pmatrix}^{\oplus \infty}.$$

Let us consider now the case of the basis $\mathsf{D} = \{\delta_x\}_{x \in X}$. For simplicity, let us denote $T_x = T_{\delta_x}$ and $T'_x = T'_{\delta_x}$. Then the operators T_x and T'_x act on $C(X^\omega, \Bbbk)$ by the rule

$$T_x(f)(x_1, x_2, \ldots) = \begin{cases} f(x_2, x_3, \ldots) & \text{if } x = x_1, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

and

$$T'_x(f)(x_1, x_2, \ldots) = f(x, x_1, x_2, \ldots). \tag{15}$$

In other words, the operator T_x is induced by the natural homeomorphism $X^\omega \longrightarrow xX^\omega : w \mapsto xw$, and T'_x is induced by its inverse map $xX^\omega \longrightarrow X^\omega : xw \mapsto w$.

Proposition 4.20 *Let* B_∞ *be a basis of* $C(X^\omega, \Bbbk)$ *associated with a marked basis* B *(see Definition* 4.10*). Then the matrices of the operators* T_x *and* T'_x*, for* $x \in X$*, in the ordered basis* B_∞ *are* $|X|$*-automatic.*

Note that we do not require in this proposition the field $\Bbbk$ to be finite.

Proof Let $\delta_x = \sum_{i=0}^{d-1} \alpha_{x,i} e_i$ for $x \in X$, $\mathsf{B} = \{e_i\}_{i=0}^{d-1}$ and $\alpha_{x,i} \in \Bbbk$. It follows from Proposition 4.16 that the entries of the matrix $\Xi_d(T_x)$ with respect to the basis B are equal to $\sum_{k=0}^{d-1} \alpha_{x,k} T'_{e_i} T_{e_k} T_{e_j}$. Every product of the form $T'_{e_i} T_{e_k} T_{e_j}$ is equal, by relations (13), either to zero, or to T_{e_j}. It follows that decimations of T_x are either zeros or of the form $\alpha_{x,i} T_{e_j}$. It follows that the set of repeated decimations of the matrix of T_x is contained in $\{\alpha_{x,0}, \alpha_{x,1}, \ldots, \alpha_{x,d-1}\} \cdot \{T_{e_0}, T_{e_1}, \ldots, T_{e_{d-1}}\} \cup \{0\}$. $\square$

Example 4.21 Let us consider the case $X = \{0, 1\}$ and $\mathsf{B} = \{y_0, y_1\}$, where $y_0 = \delta_0 + \delta_1$ and $y_1 = \delta_1$. Then the transition matrix from $\{\delta_0, \delta_1\}$ to $\{y_0, y_1\}$ is $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, whose inverse is $\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$.

It follows then from Corollary 4.19 that the matrices of T_0, T_1, T'_0, and T'_1 in the basis B_∞ are

$$T_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}^{\oplus\infty}, \quad T_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}^{\oplus\infty},$$

and

$$T'_0 = \begin{pmatrix} 1 & 0 \end{pmatrix}^{\oplus\infty}, \quad T'_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}^{\oplus\infty}.$$

Example 4.22 In the case $\Bbbk = \mathbb{C}$, it is natural to consider the operators

$$S_x = \frac{1}{\sqrt{|X|}} T_x.$$

Then S_x are isometries of the Hilbert space $L^2(X^\omega)$, and their conjugates S_x^* are equal to $\sqrt{|X|} T'_x$.

The C^*-algebra of operators on $L^2(X^\omega)$ generated by the operators S_x is called the Cuntz algebra [14], and is usually denoted $\mathcal{O}_{|X|}$. Any isometries satisfying the relations

$$S_x^* S_x = 1, \qquad \sum_{x \in X} S_x S_x^* = 1$$

generate a C^*-algebra isomorphic to $\mathcal{O}_{|X|}$. In particular, the C^*-algebra generated by the matrices E_i is the Cuntz algebra. Representation of the Cuntz algebra by matrices E_i is an example of a *permutational representation* of $\mathcal{O}_d$. More on such and similar representations, see [9].

Recall that, for $X = \{0, 1\}$, the Walsh basis of $L^2(X^\omega)$ is the basis W_∞ constructed starting from the basis $\mathsf{W} = \{y_0, y_1\}$, where $y_0 = \delta_0 + \delta_1$ and $y_1 = \delta_0 - \delta_1$. Then direct computation with of the transition matrices show that the matrices of S_0 and S_1 are

$$S_0 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}^{\oplus \infty}, \qquad S_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}^{\oplus \infty}.$$

4.5 Cuntz algebras and Higman–Thompson groups

If $\psi : X^\omega \longrightarrow X^\omega$ is a homeomorphism, then it induces a linear operator L_ψ on $C(X^\omega, \Bbbk)$ given by

$$L_\psi(f)(w) = f(\psi^{-1}(w))$$

for $f \in C(X^\omega, \Bbbk)$ and $w \in X^\omega$.

Fixing any ordered basis of $C(X^\omega, \Bbbk)$, we get thus a natural faithful representation of the homeomorphism group of X^ω in the group of units of the algebra $\mathsf{M}_\infty(\Bbbk)$ of column-finite matrices over $\Bbbk$.

Proposition 4.23 *Let ψ be a homeomorphism of X^ω. Let $u, v \in X^n$, and denote by $\psi_{u,v}$ the partially defined map given by the formula*

$$\psi_{u,v}(w) = \begin{cases} w' & \text{if } \psi(vw) = uw', \\ \text{not defined} & \text{otherwise.} \end{cases}$$

The following conditions are equivalent.

1. *The homeomorphism ψ is (non-deterministically) synchronously automatic.*
2. *The set of partial maps $\{\psi_{u,v} : u, v \in X^*, |u| = |v|\}$ is finite.*
3. *For every finite field $\Bbbk$, the operator $L_\psi : C(X^\omega, \Bbbk) \longrightarrow C(X^\omega, \Bbbk)$ is automatic.*
4. *For some finite field $\Bbbk$, the operator $L_\psi : C(X^\omega, \Bbbk) \longrightarrow C(X^\omega, \Bbbk)$ is automatic.*

Synchronously automatic homeomorphisms are defined in Definition 3.7.

Proof Equivalence of conditions (2), (3), and (5) follow directly from Proposition 4.16.

Suppose that ψ is synchronously automatic. Let $\mathfrak{A}$ be an initial automaton defining ψ. For every pair $u = a_1 a_2 \cdots a_n$, $v = b_1 b_2 \cdots b_n \in X^*$ of words of equal length, let $Q_{u,v}$ be the set of states q of $\mathfrak{A}$ such that there exists a directed path starting in the initial state q_0 of $\mathfrak{A}$ and labeled by $(a_1, b_1), (a_2, b_2), \ldots, (a_n, b_n)$. Then the set $Q_{u,v}$ defines the map $\psi_{u,v}$ in the following sense. We have $\psi_{u,v}(x_1 x_2 \cdots) = y_1 y_2 \cdots$ if and only if there exists a path starting in an element of $Q_{u,v}$ and labeled by $(x_1, y_1), (x_2, y_2), \ldots$. It follows that the number of possible maps of the form $\psi_{u,v}$ is not larger than the number of subsets of the set of states of $\mathfrak{A}$. This shows that every synchronously automatic homeomorphism satisfies condition (2).

Suppose now that a homeomorphism ψ satisfies condition (2), and let us show that it is synchronously automatic. Construct an automaton $\mathfrak{A}$ with the set of states Q equal to the set of non-empty maps of the form $\psi_{u,v}$. For every $(x, y) \in X^2$ and $\psi_{u,v} \in Q$ we have an arrow from $\psi_{u,v}$ to $\psi_{ux,vy}$, labeled by (x, y), provided the map $\psi_{ux,vy}$ is not empty. The initial state of the automaton is the map $\psi = \psi_{\emptyset,\emptyset}$. Let us show that this automaton defines the homeomorphism ψ. It is clear that if $\psi(x_1 x_2 \cdots) = y_1 y_2 \cdots$, then there exists a path starting at the initial state of $\mathfrak{A}$ and labeled by $(x_1, y_1), (x_2, y_2), \ldots$. On the other hand, if such a path exists for a pair of infinite words $x_1 x_2 \cdots$, $y_1 y_2 \cdots$, then the maps $\psi_{x_1 x_2 \cdots x_n, y_1 y_2 \cdots y_n}$ are non-empty for every n. In other words, for every n the set W_n of infinite sequences $w \in x_1 x_2 \cdots x_n X^\omega$ such that $\psi(w) \in y_1 y_2 \cdots y_n X^\omega$ is non-empty. It is clear that the sets W_n are closed and $W_{n+1} \subset W_n$ for every n. By compactness of X^ω it implies that $\bigcap_{n \geq 1} W_n$ is non-empty. It follows that $\psi(x_1 x_2 \cdots) = y_1 y_2 \cdots$. $\qquad \square$

The next corollary follows directly from condition (2) of Proposition 4.23.

Corollary 4.24 *The set of all automatic homeomorphisms of X^ω is a group.*

We have already seen in Theorem 4.14 that a homeomorphisms g of X^ω defined by an automorphisms of X^* is automatic if and only if it is finite state. Note that in this case $g_{u,v}$ is either empty (if $g(v) \neq u$) or is equal to $g|_v$.

Another example of a group of finitely automatic homeomorphisms of X^ω is the Higman-Thompson group $\mathcal{V}_{|X|}$. It is the set of all homeomorphisms that can be defined in the following way. We say that a subset $A \subset X^*$ is a *cross-section* if the sets uX^ω for $u \in A$ are disjoint and their union is X^ω. Let $A = \{v_1, v_2, \ldots, v_n\}$ and $B = \{u_1, u_2, \ldots, u_n\}$ be cross-sections of equal cardinality together with a bijection $v_i \mapsto u_i$. Define a homeomorphism $\psi : X^\omega \longrightarrow X^\omega$ by the rule

$$\psi(v_i w) = u_i w. \tag{16}$$

The set of all homeomorphisms that can be defined in this way is the Higman–Thompson group $\mathcal{V}_{|X|}$, see [12,47].

Let ψ be the homeomorphism defined by (16). It follows directly from (14) and (15) that the operator L_ψ induced by ψ is equal to

$$L_\psi = \sum_{i=1}^{n} T_{u_i} T'_{v_i},$$

where we use notation

$$T_{x_1 x_2 \cdots x_m} = T_{x_1} T_{x_2} \cdots T_{x_n}, \qquad T'_{x_1 x_2 \cdots x_m} = T'_{x_m} T'_{x_{m-1}} \cdots T'_{x_1}.$$

The next proposition follows then from Proposition 4.20.

Proposition 4.25 *The Higman–Thompson group $\mathcal{V}_{|X|}$ is a subgroup of the group of synchronously automatic homeomorphisms of X^ω.*

The group generated by $\mathcal{V}_2$ and the Grigorchuk group was studied by Roever [40]. He proved that it is a finitely presented simple group isomorphic to the abstract commensurizer of the Grigorchuk group. Generalizations of this group (for arbitrary self-similar group) was studied in [34].

5 Representations by uni-triangular matrices

5.1 Sylow p-subgroup of $\mathrm{Aut}(X^*)$

Let $|X| = p$ be prime. We assume that $X = \{0, 1, \ldots, p-1\}$ is equal to the field $\mathbb{F}_p$ of p elements. From now on, we will write the vertices of the tree X^* as tuples $(x_1, x_2, \ldots, x_n)$ in order not to confuse them with products of elements of $\mathbb{F}_p$.

Denote by $\mathcal{K}_p$ the subgroup of $\mathrm{Aut}(X^*)$ consisting of automorphisms g whose labels $\alpha_{g,v}$ of the vertices of the portrait consist only of powers of the cyclic permutation $\sigma = (0, 1, \ldots, p-1)$. It follows from (2) and (3) that $\mathcal{K}_p$ is a group. The study of the group $\mathcal{K}_p$ (and its finite analogs) were initiated by Kaloujnine [26–28].

Suppose that an element $g \in \mathcal{K}_p$ is represented by a tableau

$$[a_0, a_1(x_1), a_2(x_1, x_2), \ldots],$$

as in Sect. 2.1. Then $a_n(x_1, x_2, \ldots, x_n)$ are maps from X^n to the group generated by the cyclic permutation σ. The elements of this group act on $X = \mathbb{F}_p$ by maps $\sigma^a : x \mapsto x + a$. It follows that we can identify functions a_n with maps $X^n \longrightarrow \mathbb{F}_p$, so that an element $g \in \mathcal{K}_p$ represented by a tableau

$$[a_0, a_1(x_1), a_2(x_1, x_2), \ldots]$$

acts on sequences $v = (x_0, x_1, \ldots) \in X^\omega$ by the rule

$$g(v) = (x_0 + a_0, x_1 + a_1(x_0), x_2 + a_2(x_1, x_2), x_3 + a_3(x_1, x_2, x_3), \ldots).$$

It follows that if $g_1, g_2 \in \mathcal{K}_p$ are represented by the tableaux $[a_n]_{n=0}^\infty$ and $[b_n]_{n=0}^\infty$, then their product $g_1 g_2$ is represented by the tableau

$$\begin{aligned}
[b_0 + a_0, \quad & b_1(x_0) + a_1(x_0 + b_0), \\
& b_2(x_0, x_1) + a_2(x_0 + b_0, x_1 + b_1(x_0)), \\
& b_3(x_0, x_1, x_2) + a_3(x_0 + b_0, x_1 + b_1(x_0), x_2 + b_2(x_0, x_1)), \quad \ldots].
\end{aligned} \tag{17}$$

Denote by $\mathcal{K}_{p,n}$ the quotient of $\mathcal{K}_p$ by the pointwise stabilizer of the nth level of the tree X^*. We can consider $\mathcal{K}_{p,n}$ as a subgroup of the automorphism group of the finite subtree $X^{[n]} = \bigcup_{k=0}^{n} X^k \subset X^*$.

Proposition 5.1 *The group $\mathcal{K}_{p,n}$ is a Sylow subgroup of the symmetric group* Symm(X^n) *and of the automorphism group of the tree $X^{[n]}$.*

Proof The order of Symm(X^n) is $p^n!$, and the maximal power of p dividing it is

$$\frac{p^n}{p} + \frac{p^n}{p^2} + \cdots + \frac{p^n}{p^n} = \frac{p^n - 1}{p - 1}.$$

It follows that the order of the Sylow p-subgroup of Symm(X^n) is $p^{\frac{p^n-1}{p-1}}$. The order of $\mathcal{K}_{p,n}$ is equal to the number of possible tableaux

$$[a_0, a_1(x_1), a_2(x_1, x_2), \ldots a_{n-1}(x_1, \ldots, x_{n-1})],$$

where a_i is an arbitrary map from X^i to the cyclic group $\langle \sigma \rangle$ of order p. The number of possibly maps a_i is hence p^{p^i}. Consequently, the number of possible tableaux is $p^{1+p+p^2+\cdots+p^{n-1}} = p^{\frac{p^n-1}{p-1}}$. Since the group of all automorphisms of the tree $X^{[n]}$ is contained in Symm(X^n) and contains $\mathcal{K}_{p,n}$, the subgroup $\mathcal{K}_{p,n}$ is its Sylow p-subgroup. $\qquad\square$

Proposition 5.2 *Let $g \in \mathcal{K}_p$ be represented by a tableau $[a_0, a_1(x_1), a_2(x_1, x_2), \ldots]$. Consider the map $\alpha : \mathcal{K}_p \longrightarrow \mathbb{F}_p^\omega$, where $\mathbb{F}_p^\omega$ is the infinite Cartesian product of additive groups of $\mathbb{F}_p$, given by*

$$\alpha(g) = \left(a_0, \sum_{x_1 \in \mathbb{F}_p} a_1(x_1), \sum_{(x_1, x_2) \in \mathbb{F}_p^2} a_2(x_1, x_2), \ldots \right). \tag{18}$$

In other words, we just sum up modulo p all the decorations of the portrait of g on each level. Then α is the abelianization epimorphism $\mathcal{K}_p \longrightarrow \mathcal{K}_p / [\mathcal{K}_p, \mathcal{K}_p] \cong \mathbb{F}_p^\omega$.

Proof It is easy to check that α is a homomorphism. It remains to show that its kernel is the derived subgroup of $\mathcal{K}_p$. This is a folklore fact, and we show here how it follows from a more general result of Kaloujnine.

Let $g \in \mathcal{K}_p$ be represented by a tableau

$$[a_0, a_1(x_1), a_2(x_1, x_2), \ldots].$$

Each function $a_n(x_1, x_2, \ldots, x_n)$ can be written as a polynomial

$$\sum_{0 \le k_i \le p-1} c_{k_1, k_2, \ldots, k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

for some coefficients $c_{k_1, k_2, \ldots, k_n} \in \mathbb{F}_p$.

It is proved in [27, Theorem 6] (see also Equation (5.4) in [28]) that the derived subgroup $[\mathcal{K}_p, \mathcal{K}_p]$ of $\mathcal{K}_p$ is the set of elements defined by tableaux in which a_0 and the coefficient $c_{p-1, p-1, \ldots, p-1}$ at the eldest term $x_1^{p-1} x_2^{p-1} \cdots x_n^{p-1}$ are equal to zero for every n.

Note that $\sum_{x \in \mathbb{F}_p} x^k$ is equal to zero for $k = 0, 1, \ldots, p-2$ and is equal to -1 for $k = p - 1$. Therefore,

$$\sum_{(x_1, x_2, \ldots, x_n) \in \mathbb{F}_p^n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n} = \prod_{i=1}^{n} \sum_{x \in \mathbb{F}_p} x^{k_i}$$

is equal to zero for all n-tuples $(k_1, k_2, \ldots, k_n) \in \{0, 1, \ldots, p-1\}^n$ except for $(p-1, p-1, \ldots, p-1)$, when it is equal to $(-1)^n$. It follows that the coefficient at the eldest term of $a_n(x_1, x_2, \ldots, x_n)$ is equal to zero if and only if $\sum_{(x_1, x_2, \ldots, x_n) \in \mathbb{F}_p^n} a_n(x_1, x_2, \ldots, x_n) = 0$. $\qquad\square$

5.2 Polynomial bases of $C(X^\omega, \mathbb{F}_p)$

Proposition 5.3 *Suppose that an ordered basis* B *of* $\Bbbk^X$ *is such that the matrices of* $\pi_1(g)$ *for* $g \in G \leq \mathrm{Aut}(X^*)$ *are all upper uni-triangular, and the minimal element of* B *is the constant one function. Then the matrices of* $\pi_n(g)$ *in the basis* $\mathsf{B}^{\otimes n}$ *and of* $\pi_\infty(g)$ *in the associated basis* B_∞ *are upper uni-triangular for all* $g \in G$.

See Sect. 2.4 for the definition of the representations π_n. We say that a matrix is *upper uni-triangular* if all its elements below the main diagonal are equal to zero, and all the elements on the diagonal are equal to one. From now on, unless the contrary is specifically mentioned, "uni-triangular" will mean "upper uni-triangular". Note that if G is a subgroup of the group of finite automata, then the matrices $\pi_\infty(g)$ are automatic, by Theorem 4.14.

Proof Let $b_0 < b_1 < \cdots < b_{d-1}$ be the ordered basis B. Let $\mathsf{Y} = \{y_0, y_1, \ldots, y_{d-1}\}$ be the corresponding basis of the right module Φ_G. Namely, we take for every $\sum_{x \in X} a_x \delta_x \in \mathsf{B}$ the corresponding element $\sum_{x \in X} a_x x \in \Phi = \Bbbk[X \cdot G]$, see Sect. 2.3. Then $\mathsf{B} = \mathsf{Y} \otimes \varepsilon$, where $[\varepsilon]$ is the left G-module of the trivial representation of G, see 2.4.

If the matrices of $\pi_1(g)$ are uni-triangular, then

$$\pi_1(g)(b_i) = b_i + a_{i-1,i} b_{i-1} + a_{i-2,i} b_{i-2} + \cdots + a_{0,i} b_0$$

for some $a_{k,i} \in \Bbbk$ and all i. It follows that, in the bimodule Φ, we have relations

$$g \cdot y_i = y_i \cdot g_{i,i} + a_{i-1,i} y_{i-1} \cdot g_{i-1,i} + a_{i-2,i} y_{i-2} \cdot g_{i-2,i} + \cdots + a_{0,i} y_0 \cdot g_{0,i} \quad (19)$$

for some $g_{j,i} \in \Bbbk[G]$ such that $g_{j,i} \cdot \epsilon = \epsilon$ (Recall that the last equality just means that the sum of coefficients of $g_{j,i} \in \Bbbk[G]$, i.e., the value of the augmentation map,

is equal to one). Consequently, relation (19) together with the condition $g_{j,i} \cdot \varepsilon = \varepsilon$ hold for all $g \in \Bbbk[G]$ such that $g \cdot \varepsilon = \varepsilon$.

It follows that every element $g \cdot y_{i_1} \otimes y_{i_2} \otimes \cdots \otimes y_{i_n} \in \Phi^{\otimes n}$ is equal to $y_{i_1} \otimes y_{i_2} \otimes \cdots \otimes y_{i_n} \cdot h$ plus a sum of elements of the form $y_{j_1} \otimes y_{j_2} \otimes \cdots y_{j_n} \cdot a_{j_1, j_2,..., j_n}$, where $h \in \Bbbk[G]$ is such that $h \cdot \varepsilon = \varepsilon$, $a_{j_1, j_2,..., j_n} \in \Bbbk[G]$, and $j_k \leq i_k$ for all $k = 1, 2, \ldots, n$, and $(j_1, j_2, \ldots, j_n) \neq (i_1, i_2, \ldots, i_n)$. Taking tensor product with ε and applying Proposition 2.10, we conclude that for every function $b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_n} \in \Bbbk^{X^n}$ the function $\pi_\infty(g)(b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_n})$ is equal to $b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_n}$ plus a linear combination of functions $b_{j_1} \otimes b_{j_2} \otimes \cdots \otimes b_{j_n}$ such that $j_k \leq i_k$ for all $k = 1, 2, \ldots, n$, and $(j_1, j_2, \ldots, j_n) \neq (i_1, i_2, \ldots, i_n)$. But any such function $b_{j_1} \otimes b_{j_2} \otimes \cdots \otimes b_{j_n}$ is an element of B_∞, which is smaller than $b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_n}$ in the inverse lexicographic order. This proves that the matrix of $\pi_\infty(g)$ in the basis B_∞ is uni-triangular. $\square$

Throughout the rest of our paper we assume that $|X| = p$ is prime, $\Bbbk$ is the field $\mathbb{F}_p$ of p elements, G is a subgroup of $\mathcal{K}_p$, and we identify X with $\mathbb{F}_p$. We will be able then to use Proposition 5.3 to construct bases of $C(X^\omega, \mathbb{F}_p)$ in which the representation π_∞ of $\mathcal{K}_p$ (and hence of G) are uni-triangular.

Every function $f \in \mathbb{F}_p^X$ can be represented as a polynomial $f(x) \in \mathbb{F}_p[x]$, using the formula

$$\delta_t(x) = \frac{x(x-1)(x-2)\cdots(x-p+1)}{(x-t)},$$

where $(x - t)$ in the numerator and the denominator cancel each other (Recall that $(p-1)! = -1 \pmod p$, by Wilson's theorem).

Since $x^p = x$ as a function on $\mathbb{F}_p$ (by Fermat's little theorem), polynomials that differ by an element of the ideal generated by $x^p - x$ represent the same function. Note that the ring $\mathbb{F}_p[x]/(x^p - x)$ has cardinality p^p, hence we get a natural bijection between $\mathbb{F}_p[x]/(x^p - x)$ and $\mathbb{F}_p^X$, mapping a polynomial to the function it defines on $\mathbb{F}_p$. From now on, we will thus identify the space of functions $\mathbb{F}_p^X$ with the $\mathbb{F}_p$-algebra $\mathbb{F}_p[x]/(x^p - x)$.

Following Kaloujnine, we will call the elements of $\mathbb{F}_p[x]/(x^p - x)$ *reduced polynomials*. We write them as usual polynomials $a_0 + a_1 x + \cdots + a_{p-1} x^{p-1}$ (but keeping in mind reduction, when performing multiplication).

Suppose that $g \in G$ is such that $g(x) = x + 1$ for all $x \in X$. Then $\pi_1(g)$ acts on the functions $f \in V_1 = \mathbb{F}_p^X$ by the rule

$$\pi_1(g)(f)(x) = f(x - 1).$$

In particular, if we represent f as a polynomial, then $\pi_1(g)$ does not change its degree and the coefficient of the leading term. It follows that the matrix of the operator $\pi_1(g)$ in the basis $e_0(x) = \mathbf{1}$, $e_1(x) = x$, $e_2(x) = x^2, \ldots, e_{p-1}(x) = x^{p-1}$ is uni-triangular. Let us denote this marked basis by E.

Definition 5.4 The basis E_∞ of $C(X^\omega, \mathbb{F}_p)$ corresponding to E and consisting of all monomial functions $x_1^{k_1} x_2^{k_2} \cdots$ on X^ω ordered inverse lexicographically, so that

$$e_0 = 1, \quad e_1 = x_1, \quad e_2 = x_1^2, \ldots, e_{p-1} = x_1^{p-1}, \quad e_p = x_2, \quad e_{p+1} = x_1 x_2, \ldots$$

is called the *Kaloujnine basis* of monomials.

It is easy to see that $e_n \in \mathsf{E}_\infty$ is equal to the monomial function

$$e_n(x_1 x_2 \cdots) = x_1^{k_1} x_2^{k_2} \ldots, \tag{20}$$

where $k_1, k_2, \ldots$ are the digits of the base p expansion of n, i.e., such that $k_i \in \{0, 1, \ldots, p-1\}$ and

$$n = k_1 + k_2 \cdot p + k_3 \cdot p^2 + \cdots.$$

Coordinates of a function $f \in C(X^\omega, \mathbb{F}_p)$ in the basis E_∞ are the coefficients of the representation of the function f as a polynomial in the variables $x_1, x_2, \ldots$. Since we are dealing with functions, we assume that these polynomials are reduced, i.e., are elements of the ring $\mathbb{F}_p[x_1, x_2, \ldots]/(x_1^p - x_1, x_2^p - x_2, \ldots)$.

As an immediate corollary of Proposition 5.3 we get the following.

Theorem 5.5 *The representation π_∞ of $\mathcal{K}_p$ in the Kaloujnine basis E_∞ is uni-triangular. In particular, the representations π_∞ of all its subgroups $G \le \mathcal{K}_p$ are uni-triangular in E_∞.*

We can change the ordered basis $\mathsf{E} = \{e_0 = 1, e_1 = x, \ldots, e_{p-1}(x) = x^{p-1}\}$ to any ordered basis $\mathsf{F} = (f_0, f_1, \ldots, f_{p-1})$ consisting of polynomials of degrees $0, 1, 2, \ldots, p-1$, respectively, since then the transition matrix from E to F will be triangular, hence the representation of G in the basis F will be also uni-triangular.

For example, a natural choice is the basis B in which the matrix of the cyclic permutation $x \mapsto x + 1 : X \longrightarrow X$ is the Jordan cell

$$\begin{pmatrix} 1 & 1 & 0 & \cdots \\ 0 & 1 & 1 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

To get such a basis, define the functions $b_0, b_1, \ldots, b_{p-1} \in V_1$ by the formula

$$b_k = (\pi_1(g) - 1)^{p-1-k}(\delta_0)$$

for $k = 0, 1, \ldots, p-2$ and $b_{p-1} = \delta_0$. Then $(\pi_1(g) - 1)(b_k) = (\pi_1(g) - 1)^{p-k}(\delta_0) = b_{k-1}$, for all $k = 1, 2, \ldots, p-1$, i.e.,

$$b_{k-1}(x) = b_k(x - 1) - b_k(x).$$

Note that

$$b_0 = (\pi_1(g) - 1)^{p-1}(\delta_0) = (\pi_1(g)^{p-1} + \pi_1(g)^{p-2} + \cdots + 1)(\delta_0) = \sum_{k=0}^{p-1} \delta_k = \mathbf{1},$$

i.e., the basis $b_0 < b_1 < \cdots < b_{p-1}$ is marked.

Proposition 5.6 *For every $k \in \{1, \ldots, p-1\}$ and $x \in X = \mathbb{F}_p$ we have*

$$b_k(x) = (-1)^k \binom{x+k}{k} = (-1)^k \frac{(x+1)(x+2)\cdots(x+k)}{k!}.$$

Note that $k! \neq 0$ in $\mathbb{F}_p$ for every $k = 1, 2, \ldots, p-1$.

Proof We have $(p-1)! = 1$ and $(-1)^{p-1} = 1$ in $\mathbb{F}_p$. We also have $(x+1)(x+2)\cdots(x+p-1) = 0$ for all $x \in \mathbb{F}_p \setminus \{0\}$. It follows that $(-1)^{p-1}\binom{x+p-1}{p-1} = \delta_0 = f_{p-1}$.

It is enough now to check that the functions $(-1)^k \binom{x+k}{k}$ satisfy the recurrent relation $b_{k-1}(x) = b_k(x-1) - b_k(x)$. But we have

$$(-1)^k \binom{x-1+k}{k} - (-1)^k \binom{x+k}{k} = (-1)^{k-1}\left(\binom{x+k}{k} - \binom{x+k-1}{k}\right)$$
$$= (-1)^{k-1}\binom{x+k-1}{k-1}$$

by the well known identity

$$\binom{a}{b} = \binom{a-1}{b} + \binom{a-1}{b-1}.$$

$\square$

Proposition 5.7 *The transition matrix from the basis $(\delta_0, \delta_1, \ldots, \delta_{p-1})$ to the basis $\mathsf{B} = (b_0, b_1, \ldots, b_{p-1})$ is*

$$T = \begin{pmatrix} \binom{p-1}{0} & \binom{p-1}{1} & \binom{p-1}{2} & \binom{p-1}{3} & \cdots & \binom{p-1}{p-1} \\ \binom{p-2}{0} & \binom{p-2}{1} & \binom{p-2}{2} & \binom{p-2}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 3 & 3 & 1 & \cdots & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Its inverse is obtained by transposing T with respect to the secondary diagonal:

$$
T^{-1} = \begin{pmatrix}
0 & 0 & 0 & 0 & \cdots & \binom{p-1}{p-1} \\
0 & 0 & 0 & 0 & \cdots & \binom{p-1}{p-2} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 1 & \cdots & \binom{p-1}{3} \\
0 & 0 & 1 & 3 & \cdots & \binom{p-1}{2} \\
0 & 1 & 2 & 3 & \cdots & \binom{p-1}{1} \\
1 & 1 & 1 & 1 & \cdots & \binom{p-1}{0}
\end{pmatrix}.
$$

Proof It follows from Proposition 5.6 that the entry t_{ij} of the transition matrix, where $i = 0, 1, \ldots, p-1$ and $j = 0, 1, \ldots, p-1$ is equal to

$$
\begin{aligned}
t_{ij} &= (-1)^j \binom{i+j}{j} = (-1)^j \frac{(i+j)(i+j-1)\cdots(i+j-j+1)}{j!} \\
&= \frac{(-i-j)(-i-j+1)\cdots(-i-1)}{j!} \\
&= \frac{(p-i-j)(p-i-j+1)\cdots(p-i-1)}{j!} = \binom{p-1-i}{j},
\end{aligned}
$$

which proves the first claim of the proposition.

In order to prove the second claim, we have to show that the product

$$
\left(\binom{p-1-i}{j} \right)_{i,j=0}^{p-1} \cdot \left(\binom{j}{p-1-i} \right)_{i,j=0}^{p-1}
$$

is equal to the identity matrix. The general entry of the product is equal to

$$
a_{ij} = \sum_{k=0}^{p-1} \binom{p-1-i}{k} \cdot \binom{j}{p-1-k} = \binom{p-1-i+j}{p-1}.
$$

But $\binom{x+p-1}{p-1} = \frac{(x+p-1)(x+p-2)\cdots(x+1)}{(p-1)!}$ is equal to one for $x = 0$ and is equal to zero for $x \neq 0$, which shows that the product is equal to the identity matrix. □

Example 5.8 In the case $p = 2$, the transition matrix is $T = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, and its inverse is $T^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Let us use this to find the matrix recursions for some self-similar groups acting on the binary tree in the new basis $\mathsf{B} = \{b_0, b_1\}$.

For the adding machine (see Examples 2.3, 2.5, 2.6), we have

$$
\Xi_2(a) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1+a & 1 \end{pmatrix}.
$$

Proposition 5.9 *The matrix of the binary adding machine $\pi_\infty(a)$ in the basis B_∞ is the infinite Jordan cell*

$$
\begin{pmatrix}
1 & 1 & 0 & 0 & \cdots \\
0 & 1 & 1 & 0 & \cdots \\
0 & 0 & 1 & 1 & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}.
$$

Proof Let us prove the statement by induction, using the matrix recursion from Example 5.8. The matrix of a on V_0 is (1). The four 2-decimations $\Xi_2(J)_{ij}$ of the Jordan cell are

$$
\Xi_2(J)_{00} :
\begin{pmatrix}
① & 1 & ⓪ & 0 & \cdots \\
0 & 1 & 1 & 0 & \cdots \\
⓪ & 0 & ① & 1 & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\mapsto
\begin{pmatrix}
1 & 0 & 0 & \cdots \\
0 & 1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
$$

$$
\Xi_2(J)_{01} :
\begin{pmatrix}
1 & ① & 0 & ⓪ & \cdots \\
0 & 1 & 1 & 0 & \cdots \\
0 & ⓪ & 1 & ① & \cdots \\
0 & 0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\mapsto
\begin{pmatrix}
1 & 0 & 0 & \cdots \\
0 & 1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
$$

$$
\Xi_2(J)_{10} :
\begin{pmatrix}
1 & 1 & 0 & 0 & \cdots \\
⓪ & 1 & ① & 0 & \cdots \\
0 & 0 & 1 & 1 & \cdots \\
⓪ & 0 & ⓪ & 1 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\mapsto
\begin{pmatrix}
0 & 1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
$$

$$
\Xi_2(J)_{11} :
\begin{pmatrix}
1 & 1 & 0 & 0 & \cdots \\
0 & ① & 1 & ⓪ & \cdots \\
0 & 0 & 1 & 1 & \cdots \\
0 & ⓪ & 0 & ① & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\mapsto
\begin{pmatrix}
1 & 0 & 0 & \cdots \\
0 & 1 & 0 & \cdots \\
0 & 0 & 1 & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{pmatrix},
$$

which agrees with the recursion $\Xi_2(a) = \begin{pmatrix} 1 & 1 \\ 1+a & 1 \end{pmatrix}$. $\qquad\square$

Lemma 5.10 *If $g \in \mathrm{Aut}(X^*)$ satisfies the wreath recursion $g = (g_0, g_1)$, then its matrix recursion in the basis B_∞ (over the field $\mathbb{F}_p$ for $p = 2$) is*

$$
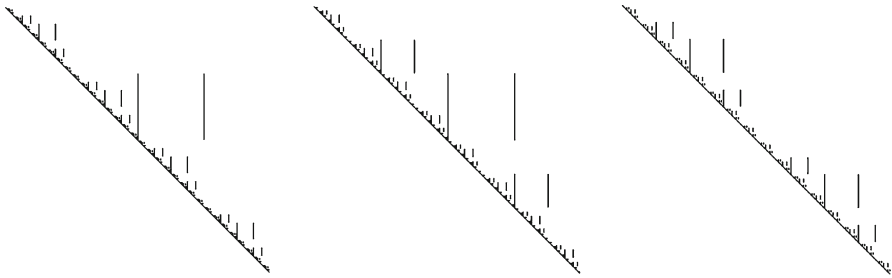\Xi_2(g) = \begin{pmatrix} g_1 & 0 \\ g_0 + g_1 & g_0 \end{pmatrix}.
$$

Fig. 7 The matrices of the generators b, c, d of the Grigorchuk group

If it satisfies $g = \sigma(g_0, g_1)$, then

$$\Xi_2(g) = \begin{pmatrix} g_0 & g_0 \\ g_0 + g_1 & g_0 \end{pmatrix}.$$

Proof We have, in the first case,

$$\Xi_2(g) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} g_0 & 0 \\ 0 & g_1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} g_1 & 0 \\ g_0 + g_1 & g_0 \end{pmatrix}.$$

In the second case:

$$\Xi_2(g) = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & g_1 \\ g_0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} g_0 & g_0 \\ g_0 + g_1 & g_0 \end{pmatrix}.$$

$\square$

Example 5.11 It follows from Lemma 5.10 that the matrix recursion for the generators of the Grigorchuk group (see Example 2.4) in the basis B is

$$\Xi_2(a) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad \Xi_2(b) = \begin{pmatrix} c & 0 \\ a+c & a \end{pmatrix},$$
$$\Xi_2(c) = \begin{pmatrix} d & 0 \\ a+d & a \end{pmatrix}, \quad \Xi_2(d) = \begin{pmatrix} b & 0 \\ 1+b & 1 \end{pmatrix}.$$

See a visualization of the matrices b, c, d on Fig. 7, where black pixels correspond to ones, and white pixels to zeros.

Denote $U_n = \langle b_0, b_1, \ldots, b_n \rangle = \langle e_0, e_1, \ldots, e_n \rangle < C(X^\omega, \mathbb{F}_p)$.

Proposition 5.12 *Each space U_n is $\mathcal{K}_p$-invariant, and the kernel in $\mathcal{K}_p$ of the restriction of π_∞ onto $U_{p^{n-1}} = \langle b_0, b_1, \ldots, b_{p^{n-1}} \rangle$ coincides with the kernel of π_n. In other words, restriction of π_∞ onto $U_{p^{n-1}}$ defines a faithful representation of $\mathcal{K}_{p,n}$.*

Proof The subspace $\langle b_0, b_1, \ldots, b_{p^{n-1}} \rangle < C(X^\omega, \mathbb{F}_p)$ is equal to the span of the product $V_{n-1} \cdot e_{p^{n-1}}$, where $e_{p^{n-1}}$ is the function on X^ω given by

$$e_{p^{n-1}}(x_1 x_2 \cdots) = x_n,$$

according to (20). In other words, it is the tensor product $V_{n-1} \otimes \langle e_1 \rangle$, where $e_1(x) = x$.

Suppose that $g \in G$ belongs to the kernel of the restriction of π_∞ onto $V_{n-1} \otimes \langle e_1 \rangle$. Then for every $v \in X^{n-1}$ we have $\pi_1(g)(\delta_v) = \delta_v$, since $\delta_v \in V_{n-1}$. Then

$$\pi_\infty(g)(\delta_v \otimes e_1) = \delta_v \otimes (e_1 \circ \pi_1(g|_v)^{-1}),$$

hence $\pi_1(g|_v)$ is identical for every $v \in X^{n-1}$. It follows that g acts trivially on X^n, i.e., that $\pi_n(g)$ is trivial. $\qquad\square$

Thus, we get a faithful representation of $\mathcal{K}_{p,n} = \wr_{k=1}^n C_p$ by uni-triangular matrices of dimension $p^{n-1} + 1$. Note that this is the smallest possible dimension for a faithful representation, since the nilpotency class of $\mathcal{K}_{p,n}$ is equal to p^{n-1}, while the nilpotency class of the group of uni-triangular matrices of dimension d is equal to $d - 1$.

5.3 The first diagonal

Let $\alpha : \mathcal{K}_p \longrightarrow \mathbb{F}_p^\omega$ be the abelianization homomorphism given by (18). We write

$$\alpha(g) = (\alpha_0(g), \alpha_1(g), \ldots).$$

If $A = (a_{ij})_{i,j=0}^\infty$ is an infinite matrix, then its *first diagonal* is the sequence $(a_{01}, a_{12}, a_{23}, \ldots)$, i.e., the first diagonal above the main diagonal of A.

Theorem 5.13 *Let $g \in G$, and let $A_g = (a_{ij})_{i,j=0}^\infty$ be the matrix of $\pi_\infty(g)$ in the basis B_∞, constructed in the previous section. Let $(s_1, s_2, \ldots) = (a_{01}, a_{12}, \ldots)$ be the first diagonal of A_g. Then*

$$s_n = \alpha_k(g),$$

where p^k is the maximal power of p dividing n.

For example, if $p = 2$, and $\alpha(g) = (a_0, a_1, a_2, \ldots)$, then the first diagonal of A_g is

$$a_0, a_1, a_0, a_2, a_0, a_1, a_0, a_3, a_0, a_1, a_0, a_2, \ldots$$

Proof The first diagonal of a product of two upper uni-triangular matrices A and B is equal to the sum of the first diagonals of the matrices A and B. It follows that it is enough to prove the theorem for rooted automorphisms of $\mathrm{Aut}(X^*)$ (i.e., automorphisms g such that $g|_v$ is trivial for all non-empty words $v \in X^*$) and for automorphisms acting trivially on the first level X.

If the automorphism is rooted, then it is a power of the automorphism

$$a : x_1 x_2 \cdots x_n \mapsto (x_1 + 1) x_2 \cdots x_n.$$

It follows from the definition of the basis B that the matrix of $\pi_\infty(a)$ is the block-diagonal matrix consisting of the Jordan cells of size p. Consequently, its first diagonal is the periodic sequence of period $(1, 1, \ldots, 1, 0)$ of length p. Hence, the first diagonal of a^s is $(s, s, \ldots, s, 0)$ repeated periodically. This proves the statement of the theorem for the automorphisms of the form a^s.

Suppose that g acts trivially on the first level of the tree. Then its matrix recursion in the basis $\{\delta_x\}_{x \in X}$ is the diagonal matrix with the entries $g|_x$ on the diagonal. It follows that the matrix recursion for g in the basis $\{b_i\}_{i=0}^{p-1}$ is equal to the product of the matrices

$$
\begin{pmatrix}
0 & 0 & 0 & \cdots & \binom{p-1}{p-1} \\
0 & 0 & 0 & \cdots & \binom{p-1}{p-2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 1 & \cdots & \binom{p-1}{2} \\
0 & 1 & 2 & \cdots & \binom{p-1}{1} \\
1 & 1 & 1 & \cdots & 1
\end{pmatrix}
\cdot
\begin{pmatrix}
g|_0 & 0 & 0 & \cdots & 0 \\
0 & g|_1 & 0 & \cdots & 0 \\
0 & 0 & g|_2 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & g|_{p-1}
\end{pmatrix}.
$$

$$
\begin{pmatrix}
1 & \binom{p-1}{1} & \binom{p-1}{2} & \cdots & \binom{p-1}{p-1} \\
1 & \binom{p-2}{1} & \binom{p-2}{2} & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 2 & 1 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0
\end{pmatrix}.
$$

It is easy to see that the entries on the first diagonal above the main diagonal of the product are equal to zero, and that the entry in the left bottom corner is equal to $g|_0 + g|_1 + \cdots + g|_{p-1}$.

If we apply the stencil map

$$
\begin{aligned}
\Xi_p(s_1, s_2, \ldots) &= ((s_1, s_{1+p}, \ldots), (s_2, s_{2+p}, \ldots), \ldots, (s_p, s_{2p}, \ldots)) \\
&= (w_1, w_2, \ldots, w_p)
\end{aligned}
$$

to the first diagonal of A_g, then $w_1, w_2, \ldots, w_{p-1}$ are the main diagonals of the p-decimations $B_1, B_2, \ldots, B_{p-1}$ of A_g, where $(B_1, B_2, \ldots, B_{p-1})$ is the first diagonal above the main diagonal of $\Xi_p(A_g)$. The sequence w_p is the first diagonal of the entry in the lower left corner of $\Xi_p(A_g)$. It follows that $(s_1, s_2, \ldots)$ is of the form $(0, 0, \ldots, s_1', 0, 0, \ldots, s_2', \ldots)$, where there are $p - 1$ zeros at the beginning and between the entries s_i', and $(s_1', s_2', \ldots)$ is the first diagonal of $\pi_\infty(g|_0 + g|_1 + \cdots + g|_{p-1})$ in the basis $(b_0, b_1, \ldots, b_{p-1})$. This provides us with an inductive proof of the statement of the theorem. □

Example 5.14 Consider the matrices a, b, c, d generating the Grigorchuk group, as it is described in Example 5.11. It follows directly from the description of the action of the elements a, b, c, d on the rooted tree (see Example 2.4) that

$$\alpha(a) = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0, \ldots),$$
$$\alpha(b) = (0, 1, 1, 0, 1, 1, 0, 1, 1, 0, \ldots),$$
$$\alpha(c) = (0, 1, 0, 1, 1, 0, 1, 1, 0, 1, \ldots),$$
$$\alpha(d) = (0, 0, 1, 1, 0, 1, 1, 0, 1, 1, \ldots).$$

It follows from Theorem 5.13 that the first diagonal of a is $(1, 0, 1, 0, \ldots)$. The first diagonal of b is $(s_1, s_2, \ldots)$ where

$$s_{2^{3k}(2m+1)} = 0, \quad s_{2^{3k+1}(2m+1)} = 1, \quad s_{2^{3k+2}(2m+1)} = 1,$$

where $k, m \geq 0$ are integers.

Sequence s_n from the previous example is a *Toeplitz sequence*, see [39,49] and [2, Exercise 10.11.42]. If G is a finitely generated finite state self-similar group, then the sequence $\alpha_k(g)$ in Theorem 5.13 is eventually periodic for every $g \in G$, and then the sequence s_n is Toeplitz.

5.4 Generating series

Let $w = (a_0, a_1, \ldots) \in \Bbbk^\omega$, and consider the corresponding formal power series $a_0 + a_1 x + a_2 x^2 + \cdots = G_w(x) \in \Bbbk[[x]]$. It is easy to see that if

$$\Xi_d(w) = (w_0, w_1, \ldots, w_{d-1}),$$

then

$$G_w(x) = G_{w_0}(x^d) + x G_{w_1}(x^d) + \cdots + x^{d-1} G_{w_{d-1}}(x^d).$$

Note that if $\Bbbk = \mathbb{F}_p$ and $d = p$, then we get

$$G_w(x) = (G_{w_0}(x))^p + x(G_{w_1}(x))^p + \cdots + x^{p-1}(G_{w_{p-1}}(x))^p.$$

We have the following characterization of automatic sequences, due to Christol, see [2, Theorem 12.2.5].

Theorem 5.15 *Let $\Bbbk$ be a finite field of characteristic p. Then a sequence $w \in \Bbbk^\omega$ is p-automatic if and only if the generating series $G_w(x)$ is algebraic over $\Bbbk(x)$.*

Similarly, if $A = (a_{ij})_{i,j=0}^\infty$ is a matrix over a field $\Bbbk$, then we can consider the formal series $G_A(x, y) = \sum_{i,j=0}^\infty a_{ij} x^i y^j \in \Bbbk[[x, y]]$. If

$$\Xi_d(A) = \left(A_{ij}\right)_{i,j=0}^{d-1},$$

then

$$G_A = \sum_{i,j=0}^{d-1} x^i y^j G_{A_{ij}}(x^d, y^d).$$

We also have a complete analog of Christol's theorem for matrices, see [2, Theorem 14.4.2].

Theorem 5.16 *Let* $\Bbbk$ *be a finite field of characteristic* p. *Then a matrix* $A = (a_{ij})_{i,j=0}^{\infty}$ *is* p-*automatic if and only if the series* G_A *is algebraic over* $\Bbbk(x, y)$.

In the case when A is triangular, it may be natural to use the generating function

$$T_A(t, s) = \sum a_{ij} t^{j-i} s^i,$$

so that $T_A(t, s) = H_0(s) + H_1(s)t + H_2(s)t^2 + \cdots$, where $H_i(y)$ are generating functions of the diagonals of A. Note that T_A and G_A are related by the formula:

$$T_A(t, s) = G_A(s/t, t)$$

Example 5.17 Consider the generators of the Grigorchuk group a, b, c, d given by the matrices from Example 5.11. Let $A(x, y)$, $B(x, y)$, $C(x, y)$, and $D(x, y)$ be the corresponding generating series. Note that the generating series of the unit matrix is $I(x, y) = \frac{1}{1+xy} = \frac{1}{1+s}$.

It follows from the recursions in Example 5.11 that

$$A = I(x^2, y^2) + yI(x^2, y^2) + xyI(x^2, y^2) = \frac{1+y+xy}{1+x^2y^2} = \frac{1}{1+s} + \frac{t}{1+s^2}.$$
$$B = C^2 + x(A^2 + C^2) + xyA^2 = (1+x)C^2 + (x+xy)A^2$$
$$C = (1+x)D^2 + (x+xy)A^2$$
$$D = (1+x)B^2 + (x+xy)I^2.$$

Let us make a substitution $A = y\tilde{A} + \frac{1}{1+xy}$, $B = y\tilde{B} + \frac{1}{1+xy}$, $C = y\tilde{C} + \frac{1}{1+xy}$, and $D = y\tilde{D} + \frac{1}{1+xy}$. Note that $I = \frac{1}{1+xy}$, so we set $\tilde{I} = 0$. The series $\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D}$ are the generating series of the matrices obtained from the matrices a, b, c, d by removing the main diagonal, and shifting all columns to the left by one position.

We have then

$$\tilde{A} = \frac{1}{1+x^2y^2} = \frac{1}{1+s^2}$$

and

$$y\tilde{B} + \frac{1}{1+xy} = (1+x)\left(y^2\tilde{C}^2 + \frac{1}{1+x^2y^2}\right) + (x+xy)\left(y^2\tilde{A}^2 + \frac{1}{1+x^2y^2}\right),$$

hence

$$y\tilde{B} = (1 + x)y^2\tilde{C}^2 + (x + xy)y^2\tilde{A}^2,$$

hence

$$\tilde{B} = (y + xy)\tilde{C}^2 + \frac{xy + xy^2}{1 + x^4 y^4} = (s + t)\tilde{C}^2 + \frac{s + ts}{1 + s^4}.$$

Similarly,

$$\tilde{C} = (y + xy)\tilde{D}^2 + \frac{xy + xy^2}{1 + x^4 y^4} = (s + t)\tilde{D}^2 + \frac{s + ts}{1 + s^4},$$

and

$$\tilde{D} = (y + xy)\tilde{B}^2 = (s + t)\tilde{B}^2.$$

Let us denote $F = \frac{s+ts}{1+s^4}$. Then $\tilde{B}$, $\tilde{C}$, and $\tilde{D}$ are solutions of the equations

$$(s + t)^7 \tilde{B}^8 + \tilde{B} + (s + t)F^2 + F = 0$$
$$(s + t)^7 \tilde{C}^8 + \tilde{C} + (s + t)^3 F^4 + F = 0,$$
$$(s + t)^7 \tilde{D}^8 + \tilde{D} + (s + t)^3 F^4 + (s + t)F^2 = 0.$$

Substituting $t = 0$, we get equations for the generating functions $B_1(s)$, $C_1(s)$, $D_1(s)$ of the first diagonals above the main in the matrices b, c, d:

$$s^7 B_1^8 + B_1 + \frac{s^3}{1 + s^8} + \frac{s}{1 + s^4} = 0, \tag{21}$$

$$s^7 C_1^8 + C_0 + \frac{s^7}{1 + s^{16}} + \frac{s}{1 + s^4} = 0, \tag{22}$$

$$s^7 D_1^8 + D_1 + \frac{s^7}{1 + s^{16}} + \frac{s^3}{1 + s^8} = 0. \tag{23}$$

Denote

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Then every upper uni-triangular matrix M can be written as

$$I + D_1(M)J + D_2(M)J^2 + D_3(M)J^3 + \cdots, \tag{24}$$

where $D_i(M)$ are diagonal matrices whose main diagonals are equal to the ith diagonals of M.

The generating series $T_M(t, s)$ is equal to

$$\frac{1}{1 + s} + M_1(s)t + M_2(s)t^2 + M_3(s)t^3 + \cdots, \tag{25}$$

where $M_i(s)$ is the usual generating series of the main diagonal of $D_i(M)$. Addition and multiplication of diagonal matrices $\mathrm{diag}(a_0, a_1, \ldots)$ corresponds to the usual addition and the *Hadamard* (coefficient-wise) multiplication of the power series $a_0 + a_1 s + a_2 s^2 + \cdots$. Note that we have

$$J \, \mathrm{diag}(a_0, a_1, \ldots) = \mathrm{diag}(a_1, a_2, \ldots) J,$$

which gives an algebraic rule for multiplication of the power series (25) corresponding to multiplication of matrices.

Namely, we can replace the matrix M by the formal power series (25), where the series in the variable s are added and multiplied coordinate-wise, while the series in the variable t are multiplied in the usual (though non-commutative) way subject to the relation

$$t(a_0 s^0 + a_1 s + a_2 s^2 + \cdots) = (a_1 s^0 + a_2 s + a_3 s^2 + \cdots)t.$$

Let $A = (a_{ij})_{i,j=0}^{\infty}$ be a matrix, and denote by $\Delta_i(M) = (a_{0,i}, a_{1,i+1}, a_{2,i+2}, \ldots)$ the sequence equal to the ith diagonal of A.

Let $d > 2$, $k \in \{0, 1, \ldots, d-1\}$, $n \geq 1$ be integers, and let $k + n = dq + r$ for $q \in \mathbb{Z}$ and $r \in \{0, 1, \ldots, d-1\}$. Then the sequence $\Xi_d(\Delta_n(M))_k = (a_{k,dq+r}, a_{k+d,d(q+1)+r}, a_{k+2d,d(q+2)+r}, \ldots)$ is equal to the qth diagonal of the matrix $(\Xi_d M)_{k,r}$.

Example 5.18 Consider, as in the previous examples, the matrices a, b, c, d generating the Grigorchuk group. Denote $A_n = \Delta_n(a)$, $B_n = \Delta_n(b)$, $C_n = \Delta_n(c)$, $D_n = \Delta_n(d)$. Note that $A_0 = (1, 1, \ldots)$, $A_1 = (1, 0, 1, 0, \ldots)$, and $A_n = 0$ for all $n \geq 2$. We have

$$\Xi_2(B_{2n}) = (C_n, A_n)$$
$$\Xi_2(B_{2n+1}) = (0, C_{n+1} + A_{n+1}).$$

Similarly,

$$\Xi_2(C_{2n}) = (D_n, A_n)$$
$$\Xi_2(C_{2n+1}) = (0, A_{n+1} + D_{n+1}),$$

and

$$\Xi_2(D_{2n}) = (B_n, I_n),$$
$$\Xi_2(D_{2n+1}) = (0, I_{n+1} + B_{n+1}).$$

These stencil recursions give us recursive formulas for the corresponding generating functions, which we will denote $B_n(s), C_n(s), D_n(s)$. Recall that $A_n(s) = 0$ for $n \geq 2$, $I_n(s) = 0$ for $n \geq 1$, $A_1(s) = \frac{1}{1+s^2}$, and $A_0(s) = I_0(s) = \frac{1}{1+s}$.

$$B_{2n}(s) = C_n^2 + s A_n^2$$
$$B_{2n+1}(s) = s(C_{n+1}^2 + A_{n+1}^2)$$
$$C_{2n}(s) = D_n^2 + s A_n^2$$
$$C_{2n+1}(s) = s(D_{n+1}^2 + A_{n+1}^2)$$
$$D_{2n}(s) = B_n^2 + s I_n^2$$
$$D_{2n+1}(s) = s(I_{n+1}^2 + B_{n+1}^2)$$

Note that iterations of the map $2n \mapsto n, 2n+1 \mapsto n+1$ on the set of non-negative integers are attracted to two fixed points $0 \mapsto 0$ and $1 \mapsto 1$. Consequently, we get the following

Proposition 5.19 *For every $n \geq 1$ the generating functions $B_n(s), C_n(s), D_n(s)$ are of the form*

$$\frac{p_0(s)}{1 + s^{2^k}} + p_1(s) B_1(s)^{2^l},$$

where $k, l \geq 0$ are integers, and $p_0(s), p_1(s)$ are polynomials over $\mathbb{F}_2$.

5.5 Principal columns

Let $g \in \mathcal{K}_p$, let A_g be the matrix of $\pi_\infty(g)$ in the basis E_∞ of monomials. Recall that we number the columns and rows of the matrix A_g starting from zero.

Proposition 5.20 *Every entry $a_{i,j}$ of the matrix A_g is a polynomial function (not depending on g) of the entries of the columns number $p, p^2, \ldots, p^{\lfloor \log_p j \rfloor}$.*
The same statement is true for the matrices of $\pi_\infty(g)$ in the basis B_∞.

Proof Let

$$[a_0, a_1(x_1), a_2(x_1, x_2), \ldots]$$

be the tableau of g, as in Sect. 5.1.

Recall that e_{p^n} is the monomial x_n. Consequently,

$$\pi_\infty(g)(e_{p^n}) = x_n + u_n(x_1, x_2, \ldots, x_{n-1}).$$

It follows that the entries of column number p^n of the matrix A_g above the main diagonal are the coefficients of the representation of u_n as a linear combination of monomials e_i for $i < p^n$, i.e., are the coefficients of the polynomial u_n. (The entries below the diagonal are zeros, and the entry on the diagonal is equal to one, of course.)

If k is a natural number such that $k < p^{n+1}$, then $e_k = x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$, where $r_n, r_{n-1}, \ldots, r_1$ are the digits of the representation of k in the base p numeration system, and

$$\pi_\infty(g)(e_k) = (x_1 + u_1)^{r_1} (x_2 + u_2(x_1))^{r_2} \cdots (x_n + u_n(x_1, x_2, \ldots, x_{n-1}))^{r_n},$$

which implies the statement of the proposition.

For the basis B_∞ we have $b_{p^n} = -\binom{x_n+1}{1} = -x_n - 1$, and a similar proof works.

$\square$

Definition 5.21 Columns number p^n, $n = 0, 1, 2, \ldots$, of the matrix A_g are called the *principal columns* of A_g.

Example 5.22 Let, for $p = 2$, the first four principal columns of the matrix A_g, $g \in \mathcal{K}_p$ be $(a_{01}, 1, 0, \ldots)^\top$, $(a_{02}, a_{12}, 1, 0, \ldots)^\top$, and $(a_{04}, a_{14}, a_{24}, a_{34}, 1, 0, \ldots)^\top$. Then the columns number 3, 5, 6, and 7 (when numeration of the columns starts from zero) are

$$
\begin{pmatrix}
a_{01}a_{02} \\
a_{01}a_{12} + a_{12} + a_{02} \\
a_{01} \\
1
\end{pmatrix},
\quad
\begin{pmatrix}
a_{01}a_{04} \\
a_{01}a_{14} + a_{14} + a_{04} \\
a_{01}a_{24} \\
a_{24} + a_{34}a_{01} + a_{34} \\
a_{01} \\
1
\end{pmatrix},
$$

$$
\begin{pmatrix}
a_{02}a_{04} \\
a_{02}a_{14} + a_{12}a_{04} + a_{12}a_{14} \\
a_{04} + a_{24}a_{02} + a_{24} \\
a_{14} + a_{34} + a_{02}a_{34} + a_{12}a_{34} + a_{12}a_{24} \\
a_{02} \\
a_{12} \\
1
\end{pmatrix},
$$

and

$$
\begin{pmatrix}
a_{01}a_{02}a_{04} \\
(a_{04} + a_{14}a_{01} + a_{14})(a_{02} + a_{12}) + a_{01}a_{12}a_{04} \\
a_{01}a_{04} + a_{01}a_{24}a_{02} + a_{01}a_{24} \\
(a_{02} + a_{12} + 1)(a_{24} + a_{34} + a_{01}a_{34}) + a_{01}(a_{14} + a_{12}a_{24}) + a_{04} + a_{14} \\
a_{01}a_{02} \\
a_{02} + a_{12} + a_{01}a_{12} \\
a_{01} \\
1
\end{pmatrix},
$$

respectively.

5.6 Uniseriality

Recall that the basis E_∞ of $C(X^\omega, \mathbb{F}_p)$ consists of monomial functions $e_0, e_1, \ldots$, where

$$e_{k_0+k_1p+k_2p^2+\ldots}(x_1, x_2, \ldots) = x_1^{k_0} x_2^{k_1} x_3^{k_2} \cdots,$$

where $k_i \in \{0, 1, \ldots, p-1\}$ are almost all equal to zero. Let us call, following [27] (though we use a slightly different definition), $n = k_0 + k_1 p + k_2 p^2 + \cdots$ the *height* of the monomial $e_n = x_1^{k_0} x_2^{k_1} x_3^{k_2} \cdots$.

Height of a reduced polynomial f is defined as the maximal height of its monomials, and is denoted $\gamma(f)$. We define $\gamma(0) = -1$ (note that our definition is different from the definition of Kaloujnine, which uses $\gamma(f) + 1$, so that height of 0 is zero).

Let us describe, following [13], an algorithm for computing height of a function $f \in \mathbb{F}_p^{X^n}$. Let B be the basis $b_0 < b_1 < \cdots < b_{p-1}$ of $\mathbb{F}_p^X$, constructed in 5.2.

Let $\{\delta_0', \delta_1', \ldots, \delta_{p-1}'\}$ and $\{b_0', b_1', \ldots, b_{p-1}'\}$ be the bases of the space of functionals $\left(\mathbb{F}_p^X\right)'$ dual to the bases $\{\delta_0, \delta_1, \ldots, \delta_{p-1}\}$ and B, respectively, i.e., δ_i' and b_i' are defined by the condition

$$\langle \delta_i' | \delta_j \rangle = \langle b_i' | b_i \rangle = \delta_{i,j}$$

for all $0 \le i, j \le p - 1$.

Then we have

$$\langle \delta_x' | f \rangle = f(x),$$

for all $f \in \mathbb{F}_p^X$ and $x \in X = \mathbb{F}_p$.

It follows from Proposition 5.7 and elementary linear algebra that the transition matrix from the basis $\{\delta_0', \delta_1', \ldots, \delta_{p-1}'\}$ to $\{b_0', b_1', \ldots, b_{p-1}'\}$ is the matrix transposed to the matrix T^{-1} of Proposition 5.7, i.e., the matrix

$$
\begin{pmatrix}
0 & \cdots & 0 & 0 & 0 & 1 \\
0 & \cdots & 0 & 0 & 1 & 1 \\
0 & \cdots & 0 & 1 & 2 & 1 \\
0 & \cdots & 1 & 3 & 3 & 1 \\
\vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\
\binom{p-1}{p-1} & \cdots & \binom{p-1}{3} & \binom{p-1}{2} & \binom{p-1}{1} & \binom{p-1}{0}
\end{pmatrix}.
$$

In other words,

$$b_k' = \sum_{l=0}^{p-1} \binom{l}{p-1-k} \delta_l', \tag{26}$$

so that

$$\langle b'_k | f \rangle = \sum_{l=0}^{p-1} \binom{l}{p-1-k} f(l).$$

For instance, $\langle b'_{p-1} | f \rangle = \sum_{x \in X} f(x)$.

Define linear maps $R_k : \mathbb{F}_p^{X^{n+1}} \longrightarrow \mathbb{F}_p^{X^n}$, $k = 0, 1, \ldots, p-1$, as the linear extension of the map

$$b_{i_1} \otimes \cdots b_{i_n} \otimes b_{i_{n+1}} \mapsto \langle b'_k | b_{i_{n+1}} \rangle \cdot b_{i_1} \otimes \cdots b_{i_n}.$$

In other terms, the map R_k is given by

$$R_k(f)(x_1, x_2, \ldots, x_n) = \langle b'_k | f(x_1, x_2, \ldots, x_n, x) \rangle,$$

where $f(x_1, x_2, \ldots, x_n, x) \in \mathbb{F}_p^{X^{n+1}}$ on the right-hand side of the equality is treated as a function of x for every choice of $(x_1, x_2, \ldots, x_n) \in \mathbb{F}_p^{X^n}$.

Using (26), we see that R_k can be computed using the formula

$$R_k(f)(x_1, x_2, \ldots, x_n) = \sum_{x=0}^{p-1} \binom{x}{p-1-k} f(x_1, x_2, \ldots, x_n, x).$$

Proposition 5.23 *Let $f \in \mathbb{F}_p^{X^n}$. Define j_n as the maximal value of $j = 0, 1, \ldots, p-1$ such that $R_j(f) \neq 0$, and then define inductively j_k for $1 \leq k < n$ as the maximal value of j such that $R_j \circ R_{j_{k+1}} \circ \cdots \circ R_{j_n}(f) \neq 0$. Then*

$$\gamma(f) = j_1 + j_2 p + j_3 p^2 + \cdots + j_n p^{n-1}.$$

Proof For any monomial $b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_m}$ and any $0 \leq j \leq p-1$, we have

$$\langle b'_j | b_{i_m} \rangle \cdot b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_{m-1}} = \begin{cases} 0 & \text{if } j \neq i_m, \\ b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_{m-1}} & \text{otherwise.} \end{cases}$$

The proof of the proposition is now straightforward. $\qquad\qquad\square$

In [13] a different basis of the dual space was used (formal dot products with b_k), but the transition matrix from their basis to $\{b'_i\}$ is triangular, so a statement similar to Proposition 5.23 holds.

One can find the height of a function $f \in \mathbb{F}_p^{X^n}$ "from the other end" by applying the maps $T_{b'_k}$, defined in Sect. 4.4. Recall, that these maps act by the rule

$$T_{b'_k}(f)(x_1, x_2, \ldots, x_{n-1}) = \langle b'_k | f(x, x_1, x_2, \ldots, x_{n-1}) \rangle,$$

i.e., they are linear extensions of the maps

$$T_{b_k'}(b_{i_1} \otimes b_{i_2} \otimes \cdots \otimes b_{i_n}) = \langle b_k' | b_{i_1} \rangle b_{i_2} \otimes \cdots \otimes b_{i_n}, \tag{27}$$

and are computed by the rule

$$T_{b_k'}(f)(x_1, x_2, \ldots, x_{n-1}) = \sum_{x=0}^{p-1} \binom{x}{p-1-k} f(x, x_1, x_2, \ldots, x_{n-1}).$$

Equation (27) imply then the following algorithm for computing height of a function.

Proposition 5.24 *Let* $f \in \mathbb{F}_p^{X^n}$, *and let* $h_k = T_{b_k'}(f)$ *for* $0 \le k \le p - 1$. *Let* $h_{k_1}, h_{k_2}, \ldots, h_{k_s}$ *be the functions of the maximal height among the functions* h_k, *and let* $j_1 = \max(k_i)$. *Then*

$$\gamma(f) = j_1 + d \cdot \gamma(h_{j_1}).$$

Proposition 5.24 seems to be less efficient than Proposition 5.23 in general, but it is convenient in the case $p = 2$. Let $f \in \mathbb{F}_2^{X^n}$. Denote $f_0 = T_{\delta_0'}(f)$, $f_1 = T_{\delta_1'}(f)$, i.e.,

$$f_0(x_1, x_2, \ldots, x_{n-1}) = f(0, x_1, x_2, \ldots, x_{n-1}),$$
$$f_1(x_1, x_2, \ldots, x_{n-1}) = f(1, x_1, x_2, \ldots, x_{n-1}).$$

Proposition 5.25 *For every* $f \in \mathbb{F}_2^{X^n}$ *we have*

$$\gamma(f) = \begin{cases} 2\max(\gamma(f_0), \gamma(f_1)) + 1 & \text{if } \gamma(f_0) \ne \gamma(f_1), \\ 2\gamma(f_0) & \text{if } \gamma(f_0) = \gamma(f_1). \end{cases}$$

Proof We have $h_0 = f_1$, $h_1 = f_0 + f_1$.

If $\gamma(f_0) < \gamma(f_1)$, then $\gamma(h_1) = \gamma(f_0 + f_1) = \gamma(f_1)$, hence $\gamma(f) = 2\gamma(h_1) + 1 = 2\gamma(f_1) + 1$.

If $\gamma(f_0) > \gamma(f_1)$, then $\gamma(h_1) = \gamma(f_0 + f_1) = \gamma(f_0) > \gamma(f_1) = \gamma(h_0)$, hence $\gamma(f) = 2\gamma(h_1) + 1 = 2\gamma(f_0) + 1$.

If $\gamma(f_0) = \gamma(f_1)$, then $\gamma(h_1) = \gamma(f_0 + f_1) < \gamma(f_1) = \gamma(f_0)$, hence $\gamma(f) = 2\gamma(h_0)$. $\qquad\square$

For more on height of functions on trees, and its generalizations, see [13].

Denote, as before, $U_n = \langle e_0, e_1, \ldots, e_n \rangle$. Then U_n consists of reduced polynomials of height not bigger than n.

Since the representation of G is uni-triangular in the basis E_∞, the spaces U_n are G-invariant, i.e., are sub-modules of the G-module $C(X^\omega, \mathbb{F}_p)$. Note also that U_{n-1} has co-dimension 1 in U_n.

Proposition 5.26 *Let* $g \in \mathcal{K}_p$ *be the adding machine. Then* $U_n = (g - 1)U_{n+1}$ *for every* n.

Proof We know that the matrix of g in the basis B_∞ is the infinite Jordan cell. Consequently, $(g - 1)(b_0) = 0$, and $(g - 1)(b_{n+1}) = b_n$ for all $n \geq 0$. It follows that

$$(g - 1)(U_{n+1}) = (g - 1)(\langle b_0, b_1, \ldots, b_{n+1} \rangle) = \langle b_0, b_1, \ldots, b_n \rangle = U_n.$$

$\square$

Theorem 5.27 *If V is a sub-module of the $\mathcal{K}_p$-module $C(X^\omega, \mathbb{F}_p)$, then either $V = \{0\}$, or $V = C(X^\omega, \mathbb{F}_p)$, or $V = U_n$ for some n.*

Proof Let $v \in V$ and $n \geq 0$ be such that $v \in U_n \backslash U_{n-1}$. Let $g \in \mathcal{K}_p$ be the adding machine defined as the automorphism of the tree X^* acting by the rule

$$g(x_1, x_2, \ldots, x_n) = \begin{cases} (x_1 + 1, x_2, \ldots, x_n) & 0 \leq x_1 \leq p - 2, \\ (0, g(x_2, \ldots, x_n)) & x_1 = p - 1. \end{cases}$$

Then $(g - 1)^k(v) \in U_{n-k} \backslash U_{n-k-1}$ for all $1 \leq k \leq n$. (We assume that $U_{-1} = \{0\}$.) It follows that $\langle v, (g-1)(v), (g-1)^2(v), \ldots, (g-1)^n(v) \rangle_{\mathbb{F}_p} = U_n \subset V$.

Let n be the maximal height of an element of V. If n is finite, then by the proven above, $V = U_n$. If n is infinite, then, by the proven above, V contains $\bigcup_{n=0}^\infty U_n = C(X^\omega, \mathbb{F}_p)$.

$\square$

We adopt therefore, the following definition.

Definition 5.28 Let $G \leq \mathcal{K}_p$. We say that the action of G on $C(X^\omega, \mathbb{F}_p)$ is *uniserial* if for every $n \geq 0$ the set $\bigcup_{g \in G}(g - 1)U_{n+1}$ generates U_n.

A module M is said to be uniserial if its lattice of sub-modules is a chain. It is easy to see that the same arguments as in the proof of Theorem 5.27 show that if the action of G on $C(X^\omega, \mathbb{F}_p)$ is uniserial, then U_n are the only proper sub-modules of the G-module $C(X^\omega, \mathbb{F}_p)$. Consequently, the G-module $C(X^\omega, \mathbb{F}_p)$ is uniserial.

In group theory (see [15,30]) an action of a group G on a finite p-group U is said to be uni-serial, if $|N : [N, G]| = p$ for every non-trivial G-invariant subgroup $N \leq U$. Here $[N, G]$ is the subgroup of N generated by the elements $h^g h^{-1}$ for $h \in H$ and $g \in G$, where h^g denotes the action of $g \in G$ on $h \in H$.

Let $g \in \mathcal{K}_p$, and let

$$[f_0, f_1(x_1), f_2(x_1, x_2), \ldots]$$

be the tableau of g. We have seen in Sect. 5.5 (see the proof of Proposition 5.20) that the entries of the principal columns $(a_{0,p^n}, a_{1,p^n}, \ldots, a_{p^n-1,p^n})^\top$ of the matrix $(a_{i,j})_{i,j=0}^\infty$ of $\pi_\infty(g)$ in the basis E_∞ are precisely the coefficients of the polynomials f_n:

$$f_n(x_1, x_2, \ldots, x_{n-1}) = \sum_{k=0}^{p^n-1} a_{k,p^n} e_k,$$

where e_k is the monomial of height k.

It follows that the height of f_n is equal to the largest index of a non-zero non-diagonal entry of the column number p^n of the matrix of $\pi_\infty(g)$ in the basis E_∞. Note that the same is true for the matrix of $\pi_\infty(g)$ in the basis B_∞.

Proposition 5.29 *Let $G \le \mathcal{K}_p$, and let $\alpha : \mathcal{K}_p \longrightarrow \mathbb{F}_p^\omega : g \mapsto (\alpha_0(g), \alpha_1(g), \ldots)$ be the abelianization homomorphism given by* (18). *The action of the group G on $C(X^\omega, \mathbb{F}_p)$ is uniserial if and only if every homomorphism $\alpha_k : G \longrightarrow \mathbb{F}_p$ is non-zero.*

Proof It follows from Theorem 5.13 that all homomorphisms α_k are non-zero if and only if for every $k = 1, 2, \ldots$ there exists $g_k \in G$ such that the entry number k on the first diagonal of $\pi_\infty(g_k)$ is non-zero.

Then for every monomial e_k the height of $(1 - g_k)(e_k)$ is equal to $k - 1$, which shows that $\bigcup_{i=1}^k (1 - g_k)(U_k)$ generates U_{k-1}, hence the action of G is uniserial. □

Corollary 5.30 *Let S be a generating set of $G \le \mathcal{K}_p$. Then the action of G on $C(X^\omega, \mathbb{F}_p)$ is uniserial if and only if for every $k = 0, 1, \ldots$ there exists $g_k \in S$ such that $\alpha_k(g_k) \ne 0$.*

Note that it also follows from Theorem 5.13 and from the fact that the entries in the principal columns are the coefficients of the polynomials in the tableau, that $\alpha_n(g) \ne 0$ if and only if height of the polynomial f_n of the tableau $[f_0, f_1(x_1), f_2(x_1, x_2), \ldots]$ representing g is equal to $p^n - 1$, i.e., has the maximal possible value.

Example 5.31 The cyclic group generated by an element $g \in \mathcal{K}_p$ is transitive on the levels X^n if and only if $\alpha_n(g) \ne 0$ for all n. It follows that if G contains a level-transitive element, then its action is uniserial. But there exist torsion groups with uniserial action on $C(X^\omega, \mathbb{F}_p)$, as the following example shows.

Example 5.32 It is easy to check that for the generators a, b, c, d of the Grigorchuk group, we have $\alpha(a) = (1, 0, 0, 0, \ldots)$, and

$$\alpha(b) = (0, 1, 1, 0, 1, 1, 0, 1, 1, 0, \ldots)$$
$$\alpha(c) = (0, 1, 0, 1, 1, 0, 1, 1, 0, 1, \ldots)$$
$$\alpha(d) = (0, 0, 1, 1, 0, 1, 1, 0, 1, 1, \ldots).$$

(In the last three equalities, each sequence have a pre-period of length 1 and a period of length 3). It follows that the action of the Grigorchuk group is uniserial.

Example 5.33 Gupta–Sidki group [24] is generated by two elements a, b acting on $\{0, 1, 2\}^*$, where a is the cyclic permutation $\sigma = (012)$ on the first level of the tree (i.e., changing only the first letter of a word), and b is defined by the wreath recursion

$$b = (a, a^{-1}, b).$$

Then $\alpha(a) = (1, 0, 0, \ldots)$, and $\alpha(b) = (0, 0, 0, \ldots)$, hence the group $\langle a, b \rangle$ does not act uniserially on $\{0, 1, 2\}^\omega$.

References

1. Aleshin, S.V.: Finite automata and the Burnside problem for periodic groups. Mat. Zametki **11**, 319–328 (1972). (in Russian)
2. Allouche, J.-P., Shallit, J.: Automatic Sequences. Theory, Applications, Generalizations. Cambridge University Press, Cambridge (2003)
3. Bacher, R.: Determinants related to Dirichlet characters modulo 2, 4 and 8 of binomial coefficients and the algebra of recurrence matrices. Int. J. Algebra Comput. **18**(3), 535–566 (2008)
4. Bartholdi, L.: Branch rings, thinned rings, tree enveloping rings. Isr. J. Math. **154**, 93–139 (2006)
5. Bartholdi, L., Grigorchuk, R.I.: Lie methods in growth of groups and groups of finite width. In: Atkinson, M., et al. (eds.) Computational and Geometric Aspects of Modern Algebra. London Math. Soc. Lect. Note Ser., vol. 275, pp. 1–27. Cambridge Univ. Press, Cambridge (2000)
6. Bartholdi, L., Grigorchuk, R.I.: On the spectrum of Hecke type operators related to some fractal groups. Proc. Steklov Inst. Math. **231**, 5–45 (2000)
7. Bartholdi, L., Grigorchuk, R., Nekrashevych, V.: From Fractal Groups to Fractal Sets. Fractals in Graz 2001. In: Grabner, P., Woess, W. (eds.) Analysis–Dynamics–Geometry–Stochastics, pp. 25–118. Birkhäuser Verlag, Basel, Boston, Berlin (2003)
8. Bartholdi, L., Grigorchuk, R.I., Šunik, Z.: Branch Groups. Handbook of Algebra, vol. 3, pp. 989–1112. North-Holland, Amsterdam (2003)
9. Bratteli, O., Jorgensen, P.E.T.: Iterated function systems and permutation representations of the Cuntz algebra. Mem. Am. Math. Soc. **139**(663), x+89 (1999)
10. Bartholdi, L., Nekrashevych, V.V.: Thurston equivalence of topological polynomials. Acta Math. **197**(1), 1–51 (2006)
11. Bartholdi, L., Virág, B.: Amenability via random walks. Duke Math. J. **130**(1), 39–56 (2005)
12. Cannon, J.W., Floyd, W.I., Parry, W.R.: Introductory notes on Richard Thompson groups. L'Enseignement Mathematique **42**(2), 215–256 (1996)
13. Ceccherini-Silberstein, T.G., Leonov, Y.G., Scarabotti, F., Tolli, F.: Generalized Kaloujnine groups, uniseriality and height of automorphisms. Int. J. Algebra Comput. **15**(3), 503–527 (2005)
14. Cuntz, J.: Simple C^*-algebras generated by isometries. Commun. Math. Phys. **57**, 173–185 (1977)
15. Dixon, J.D., du Sautoy, M.P.F., Mann, A., Segal, D.: Analytic pro-p groups, 2nd edn. Cambridge Studies in Advanced Mathematics, vol. 61. Cambridge University Press, Cambridge (1999). MR 1720368 (2000m:20039)
16. Grigorchuk, R.I., Linnell, P., Schick, T., Żuk, A.: On a question of Atiyah. C. R. Acad. Sci. Paris Sér. I Math. **331**(9), 663–668 (2000)
17. Glushkov, V.M.: Abstract theory of automata. Uspehi Mat. Nauk **16**(5), 3–62 (1961)
18. Grigorchuk, R.I., Nekrashevich, V.V.: The group of asynchronous automata and rational homeomorphisms of the Cantor set. Math. Notes **67**(5), 577–581 (2000)
19. Grigorchuk, R.I., Nekrashevich, V.V., Sushchanskii, V.I.: Automata, dynamical systems and groups. Proc. Steklov Inst. Math. **231**, 128–203 (2000)
20. Golod, E.S.: On nil-algebras and finitely approximable p-groups. Izv. Akad. Nauk SSSR Ser. Mat. **28**, 273–276 (1964)
21. Grigorchuk, R.I.: On Burnside's problem on periodic groups. Funct. Anal. Appl. **14**(1), 41–43 (1980)
22. Grigorchu, R.I.: An example of a finitely presented amenable group that does not belong to the class EG. Mat. Sb. **189**(1), 79–100 (1998)
23. Grigorchuk, R.I.: Just infinite branch groups. New Horizons in pro-p Groups. In: Shalev, A., du Sautoy, M.P.F., Segal, D. (eds.) Progress in Mathematics, vol. 184, pp. 121–179. Birkhäuser Verlag, Basel (2000)

24. Gupta, N.D., Sidki, S.N.: On the Burnside problem for periodic groups. Math. Z. **182**, 385–388 (1983)
25. Hořejš, J.: Transformations defined by finite automata. Problemy kibernetiki **9**, 23–26 (1963). (in Russian)
26. Kaloujnine, L.: Sur le groupe $\mathcal{P}_\infty$ des tableaux infinis. C. R. Acad. Sci. Paris **224**, 1097–1099 (1947)
27. Kaloujnine, L.: La structure des p-groupes de Sylow des groupes symètriques finis. Ann. Sci. Ecole Norm. Sup. (3) **65**, 239–276 (1948)
28. Kaloujnine, L.: Über eine Verallgemeinerung der p-Sylow-gruppen symmetrischer Gruppen. Acta Math. Acad. Sci. Hungar. **2**, 197–221 (1951)
29. Leonov, Yu.: On the representation of groups approximated by finite p-groups. Ukrainian Math. J. **63**(11), 1706–1718 (2012)
30. Leedham-Green, C.R., McKay, S.: The Structure of Groups of Prime Power Order. London Mathematical Society Monographs. New Series, vol. 27. Oxford University Press, Oxford, Oxford Science Publications (2002)
31. Leonov, Y., Nekrashevych, V., Sushchanky, V.: Representations of wreath products by unitriangular matrices. Dopov. Nats. Akad. Nauk Ukr., Mat. Pryr. Tekh. Nauky (4), 29–33 (2005) (in Ukrainian)
32. Lysionok, I.G.: A system of defining relations for the Grigorchuk group. Mat. Zametki **38**, 503–511 (1985)
33. Nekrashevych, V.: Iterated monodromy groups. Dopov. Nats. Akad. Nauk Ukr., Mat. Pryr. Tekh. Nauky (4), 18–20 (2003) (in Ukrainian)
34. Nekrashevych, V.: Cuntz–Pimsner algebras of group actions. J. Oper. Theory **52**(2), 223–249 (2004)
35. Nekrashevych, V.: Self-Similar Groups. Mathematical Surveys and Monographs, vol. 117. Amer. Math. Soc, Providence (2005)
36. Nekrashevych, V.: Symbolic dynamics and self-similar groups. Holomorphic dynamics and renormalization. A volume in honour of John Milnor's 75th birthday In: Lyubich, M., Yampolsky, M. (eds.) Fields Institute Communications, vol. 53, A.M.S., pp. 25–73 (2008)
37. Nekrashevych, V.: C^*-algebras and self-similar groups. Journal für die reine und angewandte Mathematik **630**, 59–123 (2009)
38. Oliĭnyk, A.S., Sushchanskiĭ, V.I.: A free group of infinite unitriangular matrices. Mat. Zametki **67**(3), 382–386 (2000)
39. Prodinger, H., Urbanek, F.J.: Infinite 0–1 sequences without long adjacent identical blocks. Discrete Math. **28**(3), 277–289 (1979) MR 548627 (81a:05032)
40. Röver, C.E.: Constructing finitely presented simple groups that contain Grigorchuk groups. J. Algebra **220**, 284–313 (1999)
41. Serre, J.-P.: Trees. Springer-Verlag, New York (1980)
42. Sidki, S.N.: A primitive ring associated to a Burnside 3-group. J. Lond. Math. Soc. (2) **55**, 55–64 (1997)
43. Said, N.: Functionally recursive rings of matrices—two examples. J. Algebra **322**(12), 4408–4429 (2009)
44. Sushchansky, V.I.: Periodic permutation p-groups and the unrestricted Burnside problem. DAN SSSR **247**(3), 557–562 (1979) (in Russian)
45. Sushchansky, V.I.: Groups of finitely automatic permutations. Dop. NAN Ukrainy (2), 29–32 (1999) (in Ukrainian)
46. Sylvester, J.J.: Thoughts on inverse orthogonal matrices, simultaneous sign successions, and tessellated pavements in two or more colours, with applications to Newton's rule, ornamental tile-work, and the theory of numbers. Philos. Mag. **34**, 461–475 (1867)
47. Thompson, R.J.: Embeddings into finitely generated simple groups which preserve the word problem. In: Adian, S.I., Boone, W.W., Higman, G. (eds.) Word Problems II. Studies in Logic and Foundations of Math., vol. 95, pp. 401–441. North-Holland Publishing Company, Amsterdam (1980)
48. Tits, J.: Free subgroups in linear groups. J. Algebra **20**, 250–270 (1972)
49. Toeplitz, O.: Ein Beispiel zur Theorie der fastperiodischen Funktionen. Math. Ann. **98**(1), 281–295 (1928) MR 1512405
50. von Haeseler, F.: Automatic Sequences. de Gruyter Expositions in Mathematics, vol. 36. Walter de Gruyter & Co., Berlin (2003) MR 1958215 (2004b:11027)
51. Walsh, J.L.: A closed set of normal orthogonal functions. Am. J. Math. **45**(1), 5–24 (1923) MR 1506485