

FINITELY GENERATED SUBGROUPS OF FREE GROUPS AS FORMAL LANGUAGES AND THEIR COGROWTH

ARMAN DARBINYAN, ROSTISLAV GRIGORCHUK, AND ASIF SHAIKH

Department of Mathematics, Texas A&M University, College Station, TX, USA
e-mail address: adarbina@math.tamu.edu

Department of Mathematics, Texas A&M University, College Station, TX, USA
e-mail address: grigorch@math.tamu.edu

Department of Mathematics, Texas A&M University, College Station, TX, USA & Department of
Mathematics & Statistics, R. A. Podar College of Commerce and Economics, Mumbai, India
e-mail address: asif.shaikh@rapodar.ac.in

ABSTRACT. Let $H \leq F_m$ be a finitely generated subgroup of the free group F_m of finite rank m . It is well-known that the language L_H of reduced words from F_m representing elements of H is regular. In the current article, using the (extended) core of the Schreier graph of H , we construct the minimal deterministic finite automaton that recognizes L_H . Then we characterize the finitely generated subgroups $H \leq F_m$ for which L_H is irreducible, and for each such H we explicitly construct an ergodic automaton that recognizes L_H . This construction gives us an efficient way to compute the cogrowth series $L_H(z)$ as well as the entropy of L_H . Several examples are provided to illustrate the method. Also, a comparison is made with the method of calculation of $L_H(z)$ based on the use of Nielsen system of generators of H .

1. INTRODUCTION

In [9, 10, 11], the notion of *cogrowth* of a subgroup H of a free group F_m was introduced and a cogrowth criterion for amenability of the factor group F_m/H (in case H is normal) was proved. The concept of cogrowth was used to construct counterexamples concerning various versions of the von Neumann conjecture about the existence of an invariant mean on groups and homogeneous spaces [1, 10, 20].

In [9, 10, 11], it was observed that when $H \leq F_m$ is a finitely generated (f.g.) subgroup, then the cogrowth series

$$H(z) = \sum_{w \in H} z^{|w|} = \sum_{n=0}^{\infty} |H_n| z^n \quad (1.1)$$

is a rational function, where $|w|$ denotes the length of $w \in H$ and H_n denotes the set of the reduced elements of length n in H with respect to a fixed basis of F_m .

Key words and phrases: Free group, Subgroups, Schreier graphs, Cogrowth, Stallings foldings, Deterministic finite automata, Regular languages.

Also, by [9], for $H \leq F_m$, the following formula holds:

$$u(z) = R\left(z, \sqrt{1 - \frac{2m-1}{m^2}z^2}\right) H\left(\frac{m\left(1 - \sqrt{1 - \frac{2m-1}{m^2}z^2}\right)}{(2m-1)z}\right),$$

where $R(x, y)$ is a rational function explicitly determined in [9], and

$$u(z) = \sum_{n=1}^{\infty} P_{1,1}^{(n)} z^n,$$

where $P_{1,1}^{(n)}$ is the probability of return to the identity $1 \in G = F_m/H$ in the simple random walk on G that starts at the identity. This formula shows a close relation between analytic properties of functions $H(z)$ and $u(z)$. In particular, the algebraicity of $u(z)$ is equivalent to the algebraicity of $H(z)$. If H is finitely generated, then the fact that $H(z)$ is rational was proven in [11] using Nielsen system of generators for H . The topics related to growth and cogrowth have gotten a lot of attention and popularity, and are widely presented in the literature. See, for example, [8].

Among various open questions related to cogrowth, the authors suggest a conjecture that $H(z)$ is rational if and only if H is finitely generated. Note that when $H \triangleleft F_m$ is a normal subgroup, the conjecture follows from the result of D. Kouksov [16].

The alternative approach to proving the rationality of $H(z)$ is via the theory of formal languages. Recall that the classical Chomsky hierarchy of languages begins with the class of regular (also called rational) languages, that is, languages recognizable by finite-automata acceptors [13]. Already in the 1950-60's, Chomsky and Schützenberger were aware that the rationality of a language $L \subset \Sigma^*$ (where Σ is finite alphabet and Σ^* denotes the set of finite words over Σ) implies rationality of the growth series

$$L(z) = \sum_{n=0}^{\infty} |B_n(L)| z^n,$$

where $B_n(L)$ is the set of words in L of length n . In fact, back then it was more popular to consider the following noncommutative version of the growth series,

$$\hat{L} = \sum_{w \in L} w,$$

the rationality of which is equivalent to the rationality of L [3, 21]. A concept closely related to cogrowth is the *entropy* of languages, which is defined as

$$h(L) = \limsup_{n \rightarrow \infty} \frac{\log |B_n(L)|}{n}.$$

For a fixed free basis $A = \{a_1, \dots, a_m\}$ of F_m , we denote by L_H the set of all reduced finite words in $(A \cup A^{-1})^*$ that represent an element of $H \leq F_m$. The fact that regularity of the language L_H is equivalent to the finite generation of H was observed by Anissimov and Seifert [2]. Since the intersection of two regular languages is regular (see [13]), a direct consequence of the theorem of Anissimov and Seifert is that the intersection of two finitely generated subgroups of F_m is again finitely generated. The last observation is in fact a well-known theorem of Howson from 1954, [14]. A proof of the theorem of Anissimov and Seifer, based on the ideas of geometric group theory, is presented in [7]. In fact, there are several ways to prove that if $H \leq F_m$ is finitely generated, then L_H is a regular language. An

elegant proof of this statement is presented by I. Kapovich and A. Myasnikov in [15]. Their proof is based on the idea of J. Stallings from [22], which is now known as *Stallings foldings*, and on the notion of the core of a Schreier graph associated with the triple (F_m, H, A) . One of the goals of the current article is to explore this approach from the finite automata theory point of view and make it more detailed and accessible.

We conclude the introduction by outlining the content of the rest of the paper. In Section 2, we recall some of the basic definitions and terminology from the theory of finite automata, formal languages and theory of graphs that will be needed later. In Section 3, we give two versions of the definition of Schreier graph, and hence two versions of the core, one of which we call *the extended core*. We also discuss a combinatorial procedure to obtain the (extended) core of $H \leq F_m$. In Section 4, we recall the definition and properties of the Nielsen system of generators of $H \leq F_m$. We also recall how to obtain a Nielsen system geometrically. In Section 5, using the DFA \mathcal{A}_{F_m} that recognizes the language of freely reduced words of F_m and the DFA $\mathcal{A}_{\widehat{\Delta}_H}$ whose Moore diagram is the extended core $\widehat{\Delta}_H$, we build another automaton, \mathcal{A}_H , that recognizes the language L_H of freely reduced words from F_m that represent the elements in H . Then we extract from \mathcal{A}_H a minimal sub-automaton \mathcal{D}_H such that $L(\mathcal{D}_H) = L_H$. In Subsection 5.4, we characterize the f.g. subgroups $H \leq F_m$ for which L_H is irreducible (see Proposition 5.17) and for such H we explicitly construct ergodic automaton $\widehat{\mathcal{D}}_H$ that recognizes L_H (see Theorem 5.14). The ergodicity of $\widehat{\mathcal{D}}_H$ allows us, by applying Perron-Frobenius theory, to obtain in Subsection 5.5 a matrix formula for the entropy of L_H , allowing its efficient computation (see Theorem 5.18). One more ingredient for showing this formula is based on relating to each finite state automaton so called *base automaton*, and by showing that under certain natural restrictions on the automaton, its entropy coincides with the entropy of its base automaton (see, Proposition 5.21).

The standard *transfer matrix* method, that goes back to Kolmogorov's theory of finite Markov chains, leads to the system of linear equations that allows us to compute $H(z)$. All of the above, in principle, is applicable to arbitrary subgroup $H \leq F_m$, but the related automata and the system of equations are finite only when H is finitely generated. In Section 6, we also include reproduction of the proof from [11] of rationality of $H(z)$ via Nielsen system of generators. The produced algorithms are of polynomial complexity. An interesting question for further investigation is to check which of the two approaches on a given finite set $\{w_1, \dots, w_k\}$ of generators of H is more efficient for computing $H(z)$.

Finally, in the last section, several concrete examples illustrating the theoretical part from the previous sections are included.

2. PRELIMINARIES

Throughout this paper, by $A = \{a_1, \dots, a_m\}$ we denote a fixed basis of the free group F_m , elements of which, along with their inverses, we regard as formal letters, whenever they are considered in the context of formal languages. Correspondingly, the set of generators $\Sigma = A \cup A^{-1}$ of F_m will be regarded as an alphabet whenever it is in the context of formal languages. The set of all finite words over the alphabet Σ is denoted by Σ^* . Algebraically, Σ^* is the free monoid generated by the finite set Σ . The *length* of a word $w \in \Sigma$, denoted by $|w|$, means the number of letters in w when each letter is counted as many times as it occurs. By $red(w)$ we denote the word that is obtained from $w \in \Sigma$ by free reduction. The subsets of Σ^* are referred to as (formal) languages over the alphabet Σ . A language L is called *regular* if it is recognized by some finite automaton. A finite automaton \mathcal{A} is a

quintuple, $\mathcal{A} = (Q, \Sigma, \delta, \mathcal{I}, \mathcal{F})$, consisting of finite set of states Q , alphabet Σ , transition function $\delta : Q \times \Sigma \rightarrow 2^Q$, the set of initial states $\mathcal{I} \subseteq Q$ and the set of final states $\mathcal{F} \subseteq Q$.

Let $G_{\mathcal{A}}$ be the *Moore* or transition diagram of \mathcal{A} , that is $G_{\mathcal{A}}$ is a labelled directed graph with vertex set Q and the directed labelled edges are described by the transition function δ with labels from Σ . Namely, vertex q is connected with vertex q' with an edge labeled by $a \in \Sigma$, if $q' \in \delta(q, a)$. (For example, Figure 4 is a depiction of a Moore diagram for an automaton that we define later.)

Let e be an edge in $G_{\mathcal{A}}$. Then by $o(e), t(e)$ and $l(e)$ we denote origin, terminus and the label of the edge e , respectively. A directed path $p = e_1 \cdots e_n$ in $G_{\mathcal{A}}$ is called *admissible* if $o(e_1) \in \mathcal{I}, t(e_i) = o(e_{i+1}),$ for $i = 1, \dots, (n-1), t(e_n) \in \mathcal{F}$. Let $w = x_1 \cdots x_n$ be a word over Σ . The automaton \mathcal{A} accepts the word w if there is an admissible path p in $G_{\mathcal{A}}$ such that $l(p) = l(e_1) \cdots l(e_n) = w$. The set of words that \mathcal{A} accepts is the language recognized by \mathcal{A} and this language is denoted by $L(\mathcal{A})$.

An automaton \mathcal{A} is *ergodic* if its Moore diagram $G_{\mathcal{A}}$ is strongly connected, that is, for any two states q and $q' \in Q$ there exists a path connecting q to q' . A language $L \subseteq \Sigma^*$ is *irreducible* if, given two words $w_1, w_2 \in L$, there exists a word $w \in \Sigma^*$ such that the concatenation $w_1 w w_2 \in L$. A regular language L is irreducible if and only if it is generated by some ergodic automaton, see Theorem 3.3.11 of [17]. An automaton \mathcal{A} is *unambiguous* if for every $w \in L(\mathcal{A})$, there is a unique admissible path $p \in G_{\mathcal{A}}$ such that $l(p) = w$. An automaton \mathcal{A} is *deterministic*, if for each state of Q , all outgoing edges carry distinct labels. It is obvious that a deterministic automaton with one initial state is unambiguous. Note that \mathcal{A} is deterministic if the codomain of δ is $\{\emptyset\} \cup Q$, that is $\delta : Q \times \Sigma \rightarrow \{\emptyset\} \cup Q$. The automaton $\mathcal{A}' = (Q', \Sigma', \delta', \mathcal{I}', \mathcal{F}')$ is said to be a *subautomaton* of $\mathcal{A} = (Q, \Sigma, \delta, \mathcal{I}, \mathcal{F})$, if $Q' \subseteq Q, \Sigma' \subseteq \Sigma, \mathcal{I}' \subseteq \mathcal{I}, \mathcal{F}' \subseteq \mathcal{F}$, and for each $q' \in Q', a' \in \Sigma', \delta'(q', a') = \delta(q, a) \cap Q'$. In the language of Moore diagrams, an equivalent definition would be: \mathcal{A}' is a subautomaton of \mathcal{A} if its Moore diagram $G_{\mathcal{A}'}$ is a subdiagram of $G_{\mathcal{A}}$.

Let \mathcal{A} be a finite automaton. We say that \mathcal{A} is *essential* if in its Moore diagram $G_{\mathcal{A}}$ every vertex (hence, also every edge) belongs to some path connecting an initial state to a final state, i.e. to an admissible path. If \mathcal{A}' is an essential subautomaton of \mathcal{A} such that $L(\mathcal{A}') = L(\mathcal{A})$, then we say that \mathcal{A}' is an *essential part* of \mathcal{A} .

Let $k \geq 1$. An automaton \mathcal{A} has *homogeneous ambiguity* k if, for any nonempty word $w \in L(\mathcal{A})$, there are exactly k admissible paths p_1, \dots, p_k in $G_{\mathcal{A}}$ with label w . In case the number of such paths for each $w \in L(\mathcal{A})$ is bounded from above by k , we say that \mathcal{A} has *bounded ambiguity*. We shall use the terminology *DFA* \mathcal{A} for the deterministic (unambiguous) finite automaton with exactly one initial state i.e. $\mathcal{A} = (Q, \Sigma, \delta, \{q_0\}, \mathcal{F})$. The DFA \mathcal{A} is said to be *minimal* if there is no DFA with smaller number of states that recognizes the same language $L(\mathcal{A})$. It is known that for any regular language L , up to isomorphism, there is a unique minimal DFA recognizing L , called *the minimal deterministic finite automaton* of L (see Theorem 3.10 on page 67 of [13]).

By *inaccessible state* we mean a non-initial state that does not have any incoming edge. Let \mathcal{A} be a DFA and let $w, w' \in \Sigma^*$. For every state q and word w , the value of $\delta(q, w)$ is the end state of the path in \mathcal{A} that starts at q and reads the input word w . In case such path does not exist, we define $\delta(q, w) = \emptyset$. Let $L = L(\mathcal{A})$ be the language recognized by DFA \mathcal{A} and let $w, w' \in \Sigma^*$. Then there is a natural equivalence relation R_L on words associated with L , given by

$$wR_L w' \iff \forall z \in \Sigma^* \text{ either both or neither of } wz \text{ and } w'z \text{ is in } L \quad (2.1)$$

Observe that the number of equivalence classes of R_L is at most the number of states of \mathcal{A} , which is finite. Now we recall a version of Myhill-Nerode Theorem.

Theorem 2.1. (Theorem 3.9 and Theorem 3.10 of [13]) *Let $L \subseteq \Sigma^*$ be a regular language. Then, the relation R_L defines a DFA $\mathcal{A}' = (Q', \Sigma, \delta', \{q'_0\}, \mathcal{F}')$ for L whose states correspond to the equivalence classes of R_L . Moreover, this is the unique minimal DFA for L (up to isomorphism), where*

$$\begin{aligned} Q' &= \{[w] \mid w \in \Sigma^*\} \\ \{q'_0\} &= \text{the equivalence class of the empty word.} \\ \mathcal{F}' &= \{[w] \mid w \in L\} \text{ and} \\ \delta'([w], a) &= [wa]. \end{aligned}$$

A graph $G = (V, E)$ is called locally finite if the degree of each vertex of V is finite. The *in* degree of a vertex v of the directed graph is the number of edges in the graph that have v as the terminus. Similarly the *out* degree of a vertex v of the graph is the number of edges in the graph that have v as the origin. We denote *in* and *out* degree of the vertex v by $\deg^-(v)$ and $\deg^+(v)$. If for every vertex v of the graph $\deg^+(v) = \deg^-(v)$, then we will ignore the signs and by $\deg(v)$ we denote the degree of the vertex v . For further details on the theory of finite automata, we refer the reader to [13] and [21].

3. THE SCHREIER GRAPH AND THE CORE ASSOCIATED WITH A SUBGROUP H OF F_m

This section is devoted to the Schreier graph and the core of a f.g. subgroup H of F_m . We shall also discuss the procedure to obtain the core of H using Stallings foldings and some of the important properties of the core.

3.1. The Schreier graph of subgroup H of F_m . We define two versions of the Schreier graph associated with $H \leq F_m$, which we denote by Γ and $\widehat{\Gamma}$, respectively. The set of vertices of Γ and $\widehat{\Gamma}$ is the same and is the set $V = \{H_g \mid g \in F_m\}$ of right cosets. The set of edges E of Γ is the set $E = \{(H_g, H_{ga}) \mid g \in F_m, a \in A\}$ consisting of pairs $e = (H_g, H_{ga})$ of cosets. The edges are oriented and H_g is the origin $o(e)$ of e while H_{ga} is the terminus $t(e)$ of e . Moreover, such an edge has the label $\mu(e) = a$. Each vertex in Γ has m outgoing edges whose labels constitute the set A . The graph $\widehat{\Gamma}$ is obtained from Γ by adding edges from the set $\overline{E} = \{\overline{e} \mid e \in E\}$ where $\overline{e} = (H_{ga}, H_g)$ if $e = (H_g, H_{ga})$ and the label $\mu(\overline{e}) = \mu(e)^{-1} = a^{-1} \in A^{-1}$. Thus $\Gamma = (V, E, \mu)$ and $\widehat{\Gamma} = (V, E \cup \overline{E}, \widehat{\mu})$, where $\widehat{\mu}(e) = \mu(e)$ if $e \in E$ and $\widehat{\mu}(\overline{e}) = \mu(e)^{-1}$ if $\overline{e} \in \overline{E}$. Each vertex of $\widehat{\Gamma}$ has $2m$ outgoing edges and $2m$ incoming edges, whose labels constitute the set $\Sigma = A \cup A^{-1}$. We call Γ the *Schreier graph* and $\widehat{\Gamma}$ the *extended Schreier graph* of H . The vertex $v_0 = H_1 = H$ is the distinguished vertex, so in fact Γ and $\widehat{\Gamma}$ are rooted graphs with root v_0 . Observe that in fact, according to the standard terminology in graph theory, Γ and $\widehat{\Gamma}$ are multigraphs as they may have loops and multiple edges. We will use the obvious notion of path p in directed graph Γ or $\widehat{\Gamma}$ and its label $\mu(p) \in A^*$ or $\widehat{\mu}(p) \in \Sigma^*$, respectively.

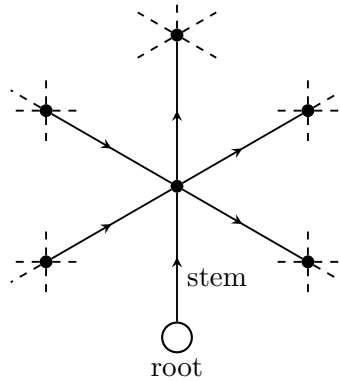


FIGURE 1. A branch

3.2. The core graph of subgroup H of F_m . The *core* $\Delta_H = (\widehat{V}, E_{\Delta_H})$ is the subgraph of the Schreier graph Γ that is defined as the union of closed paths containing the root vertex v_0 .

A branch of a k -regular tree is a subtree which has one degree 1 vertex, which we call the *root* of the branch and all the other vertices have degree k . Such a branch is uniquely determined by its *stem*, which is the oriented edge going from the root to the interior of the branch, see Figure (1). A subgraph of the Schreier graph Γ isomorphic to a branch in the Cayley graph X_m of F_m (with its labeling) is called a *hanging branch*. The Cayley graph $X_m = \text{Cay}(F_m, \Sigma)$ is a homogeneous tree of degree $|\Sigma|$. The core of the Schreier graph Γ can be obtained also by removing the hanging branches. Moreover, if the core Δ_H is known, then the graph Γ can be obtained from the core Δ_H by filling the deficient valencies of the vertices of Δ_H with maximal hanging branches (so that all the degrees of the resulting graph have the degree $2m$). Thus, since the Schreier graph Γ is connected, its core Δ_H is also connected. We refer the reader to [12] for the descriptions of the Hopf decomposition of the boundary in terms of Γ , Δ_H and the collection of hanging branches.

Let $\overline{E}_{\Delta_H} = \{\overline{e} \mid e \in E_{\Delta_H}\}$. We now define the *extended core* graph $\widehat{\Delta}_H = (\widehat{V}, \widehat{E}, \widehat{\mu})$ from the core Δ_H , where $\widehat{E} = E_{\Delta_H} \cup \overline{E}_{\Delta_H}$. Observe that the extended core $\widehat{\Delta}_H$ is a subgraph of $\widehat{\Gamma}$. It is easy to see that if $e \in \widehat{E}$ then $\overline{e} \in \widehat{E}$ (i.e. if e belongs to the path p , then \overline{e} belongs to the path \overline{p} obtained from p by obvious inversion of the direction). We say that a labeled path is *reduced* if it does not contain adjacent edges with labels of the form aa^{-1} , otherwise, we say that the path is not reduced or we say that it *backtracks*. Note that paths in the graph $\widehat{\Delta}_H$ are not necessarily reduced and may backtrack. For example, a path $p = e\overline{e}$ in $\widehat{\Delta}_H$ from v to v , where $e, \overline{e} \in \widehat{E}$ and $o(e) = v = t(\overline{e}), t(e) = v' = o(\overline{e})$ is not a reduced path. In further applications we regard $\widehat{\Delta}_H$ as a DFA $\mathcal{A}_{\widehat{\Delta}_H}$ having the root vertex v_0 as the initial and the final states. More precisely, the Moore diagram of the DFA

$$\mathcal{A}_{\widehat{\Delta}_H} = (\widehat{V}, \Sigma, \delta_{\widehat{\Delta}_H}, \{v_0\}, \{v_0\})$$

is the extended core graph

$$\widehat{\Delta}_H = (\widehat{V}, \widehat{E}, \widehat{\mu}),$$

where for each $e \in \widehat{E}$ connecting vertex v to va_j^ϵ , we have

$$\delta_{\widehat{\Delta}_H}(v, a_j^\epsilon) = va_j^\epsilon, \tag{3.1}$$

where $\widehat{\mu}(e) = a_j^\epsilon \in \Sigma$. The language $L(\mathcal{A}_{\widehat{\Delta}_H})$ of the DFA $\mathcal{A}_{\widehat{\Delta}_H}$ contains words $w = \widehat{\mu}(p)$, where p is a admissible path of $\mathcal{A}_{\widehat{\Delta}_H}$. Notice that the admissible paths p in $\mathcal{A}_{\widehat{\Delta}_H}$ may or may not be reduced. Hence not all words in the language $L(\mathcal{A}_{\widehat{\Delta}_H})$ are reduced. We denote by L_H the language of reduced elements of a f.g. subgroup H of F_m . Theorem 5.1 from [15] can be read as

Theorem 3.1. $L_H \subseteq L(\mathcal{A}_{\widehat{\Delta}_H})$. Moreover, the words in $L(\mathcal{A}_{\widehat{\Delta}_H}) \setminus L_H$ are not reduced.

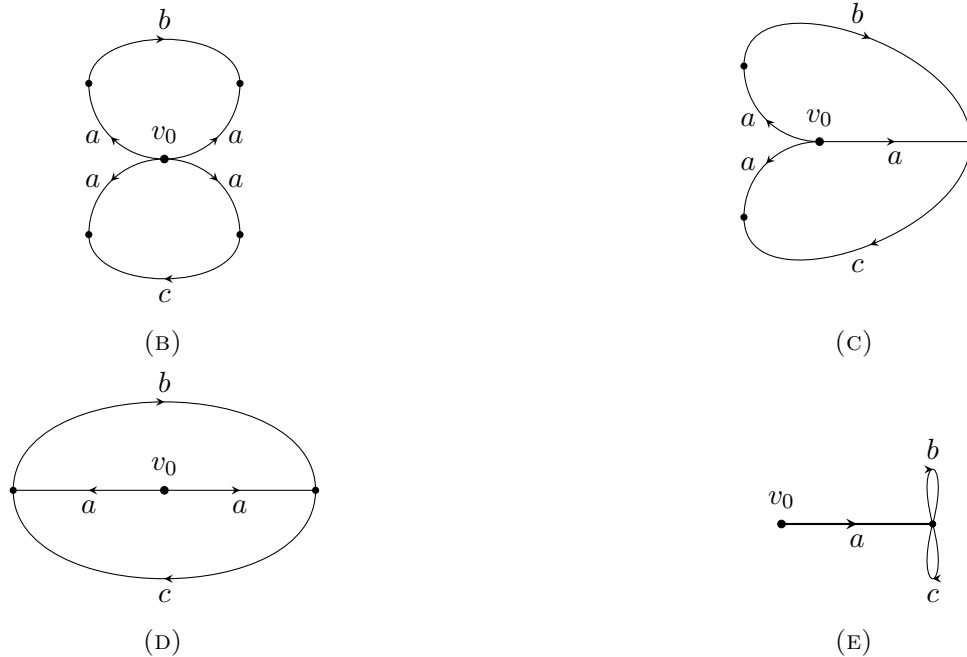
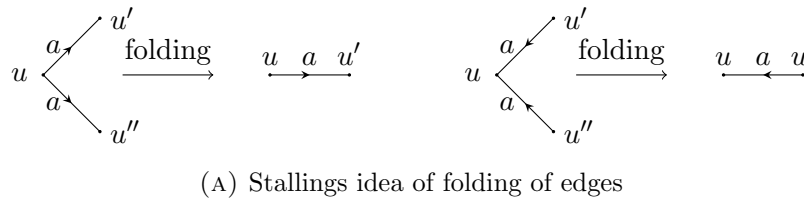


FIGURE 2. Construction of the core Δ_H of $H = \langle aba^{-1}, aca^{-1} \rangle$. In (2b), we start with a bouquet of two loops attached to a base vertex v_0 . Each loop we split into 3 oriented edges with positive labels a, b, a and a, c, a , respectively. Observe that the first two edges of each loop have clockwise orientation whereas the third edge in both the loops has the reverse orientation. In Figures (2c) and (2d), we use Stallings foldings (see 2a) to obtain the required Δ_H (see 2e).

3.3. Stallings foldings. Let H be generated by elements w_1, \dots, w_k . We identify w_i with freely reduced words in the alphabet Σ . Then there is a simple combinatorial procedure to obtain the core Δ_H . This procedure is based on the topological idea of *folding* developed by J. Stallings in [22]. Roughly the procedure can be described as follows.

Start with a bouquet of k circles glued together along a vertex v_0 . Split the i -th circle, $1 \leq i \leq k$, into $|w_i|$ edges which are oriented and labeled by the letters from the set $A \cup A^{-1}$ so that the label of the i -th circle (as read from v_0 to v_0) is precisely the word w_i . Reverse the edges with (negative) label x^{-1} from the set A^{-1} and assign the (positive) label x from the set A (see Figure 2).

Suppose e_1, e_2 are edges of this graph with a common origin and the same label $x \in A$. Then, informally, folding the graph at e_1, e_2 means identifying e_1 and e_2 in a single new edge labeled by x .

At the first step, fold the graph at the edges that are originated at the root vertex v_0 . After performing these foldings, fold the graph at the edges that are originated at the other vertices and continue the process. As we assumed that the subgroup H is finitely generated, the process will stop after applications of finitely many steps (i.e. it stops when no more folding is possible). The resulting graph, up to isomorphism of labeled graphs, does not depend on the performed sequence of foldings. And the resulting graph is isomorphic to the core Δ_H (see [15, 22]). Moreover, the algorithm based on above procedure has polynomial time complexity (see [15]).

In fact, the above procedure works also for the situation of an infinitely generated group $H = \langle w_1, \dots, w_n, \dots \rangle$. One just has to begin with folding of the bouquet of the two loops labeled by w_1 and w_2 , after the process stops add to the obtained graph a new loop labeled by w_3 , apply folding, then add w_4 etc. The process will converge to the infinite folded (i.e. no more folding is possible) graph Δ_H with rooted vertex v_0 .

The following lemma lists some of the well-known properties of the graph $\widehat{\Delta}_H$ of $H \leq F_m$.

Lemma 3.2. *Let $\widehat{\Delta}_H = (\widehat{V}, \widehat{E}, \widehat{\mu})$ be the extended core of H . Then the following holds.*

- (1) *The graph $\widehat{\Delta}_H$ has no degree one vertices, except possibly for the root vertex v_0 .*
- (2) *The degree of each vertex in $\widehat{\Delta}_H$ is at most $2m$.*
- (3) *$[F_m : H] < \infty$ if and only if $\widehat{\Delta}_H$ is a finite $2m$ -regular graph. In this case $[F_m : H] = |\widehat{V}|$.*
- (4) *H is normal in F_m if and only if $\widehat{\Delta}_H$ is a $2m$ -regular graph and any vertex v of \widehat{V} can be considered as the root vertex of $\widehat{\Delta}_H$.*
- (5) *For any edge $e \in \widehat{E}$ there exists a reduced path p in $\widehat{\Delta}_H$ such that it travels e only once and $o(p) = t(p) = v_0$.*
- (6) *The DFA $\mathcal{A}_{\widehat{\Delta}_H}$ is essential.*

Proof. For the proofs of above properties (1)-(4) we refer the reader to Propositions 3.8 and 8.3, and Theorem 8.14 of [15]. Property (5) follows immediately from the definition of the extended core graph $\widehat{\Delta}_H$. Property (6) follows from Lemma 3.9 of [15]. \square

Convention 3.3. From now on, we will assume that H is a non-trivial finitely generated subgroup of a free group $F_m = F(A)$, where $A = \{a_1, a_2, \dots, a_m\}$. $\Sigma = A \cup A^{-1}$. By v_0 we always mean the root vertex of the (extended) core of H . By $\deg(v_0)$ we denote the in (=out) degree of v_0 in the extended core graph $\widehat{\Delta}_H$.

4. THE NIELSEN SYSTEM OF GENERATORS

Let H be a subgroup of F_m generated by the set $S = \{w_i\}_{i=1}^k, 1 \leq k \leq \infty$ of freely reduced words over Σ . We further assume that S is a Nielsen basis. Recall that a set S of freely reduced words from Σ^* has the Nielsen property if the following two conditions hold:

(1) If $u, v \in S \cup S^{-1}$ and $u \neq v^{-1}$ then

$$|u \cdot v|_{\Sigma} \geq |u|_{\Sigma}, |u \cdot v|_{\Sigma} \geq |v|_{\Sigma}.$$

(2) If $u, v, w \in S \cup S^{-1}$ and $u \neq w^{-1}, v \neq w^{-1}$ then

$$|u \cdot v \cdot w|_{\Sigma} > |u|_{\Sigma} + |w|_{\Sigma} - |v|_{\Sigma},$$

where by $|w|_{\Sigma}$ we mean the length of the reduced word w over Σ .

Condition (1) means that not more than a half of u and not more than a half of v freely cancels in the product $u \cdot v$. Condition (2) means that assuming (1) after free cancellation in the product $u \cdot v \cdot w$ at least one letter of v will remain un-cancelled. See [19]. From (1) and (2) it follows that S is a free basis of the subgroup $\langle S \rangle$ of F_m .

Nielsen was the first who proved that every non trivial subgroup of F_m has a set of generators with properties (1) and (2). His argument was quite involved. A simpler proof is given in [19]. In Theorem 3.4 of [19], it is shown that any minimal Schreier system of generators of subgroup H of F_m has the Nielsen property. See also Proposition 6.7 in [15] or [18].

Let us recall briefly how to get a Nielsen system geometrically. Let H be a subgroup of F_m and Γ be the corresponding Schreier graph. Recall that labels of edges of Γ belong to the set $A = \{a_1, \dots, a_m\}$. Let T be the spanning tree in Γ . The set of vertices of T is same as the set of vertices V of Γ . The tree T is obtained from Γ by deletion of some edges. Let E' be the set of deleted edges. With each edge $e \in E'$ we associate an element w_e of H which is the word $\hat{\mu}(p)$, where p is the unique path in Γ as described in Figure (3). That

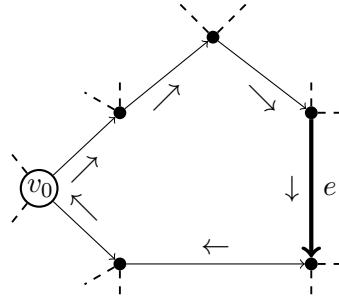


FIGURE 3. A Schreier generator

means, we connect the root v_0 with $o(e)$ by a reduced path that goes through the edges in T , then the path p goes along the edge e and after that goes from $t(e)$ to v_0 moving through the edges of T in opposite direction (thus the return path $t(e) \rightarrow v_0$ goes in fact through the edges in the extended Schreier graph $\hat{\Gamma}$).

It is obvious that w_e belongs to H . It is not obvious but the result of Schreier is that, the set $S = \{w_e\}_{e \in E'}$ is a free basis of H . Such a basis is called Schreier basis. The choice of a spanning tree is not canonical and usually there are a plenty of such choices (hence plenty of choices for Schreier system of generators). Some of the choices of T are better

than others. A spanning tree T is geodesic (or minimal) with respect to the root v_0 if for any vertex $v \in V$ the combinatorial distance from v to v_0 in Γ is the same as the distance from v to v_0 in T . Here we assume that we convert both graphs Γ and T into non-oriented graphs (by forgetting the direction of each edge). In this case the combinatorial distance (i.e. the number of edges in the closest path connecting two vertices) is the metric.

It is known that a spanning tree in the connected locally finite graph always exists and there is an effective procedure to find such a tree if the graph itself is defined in an effective way. It is well-known that for any spanning tree the corresponding Schreier system $S = \{w_e\}_{e \in E'}$ satisfies the Nielsen properties (1) and (2). It is straightforward that H is a finite index subgroup in F_m if and only if the core Δ_H coincides with the Schreier graph Γ of H (or the extended core $\widehat{\Delta}_H$ coincides with the extended Schreier graph $\widehat{\Gamma}_H$ of H). So if $[F_m : H] < \infty$, then we can have geodesic spanning trees T_{Δ_H}, T_Γ of Δ_H and Γ , respectively that coincide (i.e. $T_{\Delta_H} = T_\Gamma$). If $[F_m : H] = \infty$, then we can have $T_{\Delta_H} \subset T_\Gamma$. In this case, observe that the subtree $T_\Gamma \setminus T_{\Delta_H}$ is disconnected and the corresponding connected components are the hanging branches of Γ . Therefore, similarly to the finite index case, if $[F_m : H] = \infty$, then in order to find the Nielsen generating set it is sufficient to consider the T_{Δ_H} . Let $E'_{\Delta_H} = E_{\Delta_H} \setminus E(T_{\Delta_H})$ be the set of deleted edges of Δ_H . Then $S = S'$, where $S' = \{w_e\}_{e \in E'_{\Delta_H}}$. The highlighted geodesic spanning tree T_{Δ_H} in Δ_H of Figure (2e) can be used to find a Nielsen generating set $\{aba^{-1}, aca^{-1}\}$ of H .

5. THE CONSTRUCTION OF \mathcal{D}_H AND $\widehat{\mathcal{D}}_H$, THEIR PROPERTIES AND CONSEQUENCES

The first goal of this section is to construct the minimal DFA \mathcal{D}_H that recognizes the language L_H of reduced words of a finitely generated subgroup $H \leq F_m$. We shall approach this construction by defining an unambiguous automaton \mathcal{A}_H as the (Cartesian) product of two automata. Then we obtain the minimal DFA \mathcal{D}_H as the essential part of \mathcal{A}_H . At the end we obtain the multi-initial state automaton $\widehat{\mathcal{D}}_H$ by replacing the initial state of \mathcal{D}_H and provide the complete description when $\widehat{\mathcal{D}}_H$ is ergodic. In Theorem 5.18, we utilize the ergodicity property of $\widehat{\mathcal{D}}_H$ to obtain an entropy formula for L_H .

Recall that a finite automaton \mathcal{A} is essential if in its Moore diagram $G_{\mathcal{A}}$ every vertex (hence, also every edge) belongs to some path connecting an initial state to a final state, i.e. to an admissible path. The following lemma will be used later.

Lemma 5.1. *Every finite automaton \mathcal{A} has an essential part. If \mathcal{A} is an unambiguous finite automaton, then it has only one essential part, called the essential part of \mathcal{A} .*

Proof. Define \mathcal{A}' to be the subautomaton of \mathcal{A} such that $G_{\mathcal{A}'}$ is the union of all admissible paths of $G_{\mathcal{A}}$. Then, clearly, $L(\mathcal{A}') = L(\mathcal{A})$ and \mathcal{A}' is essential, hence \mathcal{A}' is an essential part of \mathcal{A} .

If \mathcal{A} is unambiguous, then no proper subautomaton of \mathcal{A}' generates the language $L(\mathcal{A})$. Also, if a vertex in $G_{\mathcal{A}}$ does not belong to some admissible path in $G_{\mathcal{A}}$, then, by definition, it will not belong to any essential subautomaton of \mathcal{A} . Thus, if \mathcal{A} is unambiguous, then \mathcal{A}' is the only essential part of \mathcal{A} . \square

5.1. **The automaton \mathcal{A}_H .** One of the important closure properties of regular languages is that the intersection of two regular languages is regular. See Theorem 3.3 on page 59 of [13]. A DFA recognizing the intersection of two regular languages can be constructed as follows. Let $L_1, L_2 \subseteq \Sigma^*$ be regular languages. Also, let $\mathcal{A}_1 = (Q_1, \Sigma, \delta_1, \{q_0\}, \mathcal{F}_1)$ and $\mathcal{A}_2 = (Q_2, \Sigma, \delta_2, \{q'_0\}, \mathcal{F}_2)$ be two DFA such that $L_1 = L(\mathcal{A}_1)$ and $L_2 = L(\mathcal{A}_2)$. Define the product $\mathcal{A}_1 \times \mathcal{A}_2$ of \mathcal{A}_1 and \mathcal{A}_2 as follows.

$$\mathcal{A}_1 \times \mathcal{A}_2 = (Q_1 \times Q_2, \Sigma, \delta, \{(q_0, q'_0)\}, \mathcal{F}_1 \times \mathcal{F}_2)$$

such that for all $q \in Q_1, q' \in Q_2$ and $x \in \Sigma$, we define

$$\begin{aligned} \delta((q, q'), x) &= (\delta_1(q, x), \delta_2(q', x)) \text{ if } \delta_1(q, x) \neq \emptyset, \delta_2(q', x) \neq \emptyset, \text{ and} \\ &\delta((q, q'), x) = \emptyset, \text{ otherwise.} \end{aligned}$$

Then $\mathcal{A}_1 \times \mathcal{A}_2$ is a DFA such that $L(\mathcal{A}_1 \times \mathcal{A}_2) = L_1 \cap L_2$.

To construct a DFA \mathcal{A}_H that recognizes the language L_H of reduced words of H , we shall consider first the automaton $\mathcal{A}_1 = \mathcal{A}_{\widehat{\Delta}_H} = (\widehat{V}, \Sigma, \delta_{\widehat{\Delta}_H}, \{v_0\}, \{v_0\})$ discussed in Section 3. Recall that the Moore diagram of $\mathcal{A}_{\widehat{\Delta}_H}$ is the extended core $\widehat{\Delta}_H$.

We take a second automaton $\mathcal{A}_2 = \mathcal{A}_{F_m} = (Q_{F_m}, \Sigma, \delta_{F_m}, \{q_0\}, \mathcal{F}_{F_m})$, where the set of states and of final states are both equal to

$$Q_{F_m} = \mathcal{F}_{F_m} = \{q_0\} \cup \{q_i^\epsilon \mid a_i^\epsilon \in \Sigma\}$$

and the transition function $\delta_{F_m} : Q_{F_m} \times \Sigma \rightarrow Q_{F_m}$ is

$$\delta_{F_m}(q_0, a_j^{\epsilon'}) = q_j^{\epsilon'}, \quad \text{for all } a_j^{\epsilon'} \in \Sigma, \quad (5.1)$$

$$\delta_{F_m}(q_i^\epsilon, a_j^{\epsilon'}) = q_j^{\epsilon'}, \quad \text{if } a_i^\epsilon \neq (a_j^{\epsilon'})^{-1}. \quad (5.2)$$

See Figure 4 for the Moore diagram of \mathcal{A}_{F_m} . We define $\mathcal{A}_H = \mathcal{A}_{\widehat{\Delta}_H} \times \mathcal{A}_{F_m}$. Namely,

$$\mathcal{A}_H = \mathcal{A}_{\widehat{\Delta}_H} \times \mathcal{A}_{F_m} = (Q_H, \Sigma, \delta_H, (v_0, q_0), \mathcal{F}_H),$$

where

$$\begin{aligned} Q_H &= \widehat{V} \times Q_{F_m} = \{(v, q) \mid v \in \widehat{V} \text{ and } q \in Q_{F_m}\}, \\ \mathcal{F}_H &= \{v_0\} \times \mathcal{F}_{F_m}, \end{aligned}$$

and

$$\delta_H((v_0, q_0), a_j^{\epsilon'}) = \left(\delta_{\widehat{\Delta}_H}(v_0, a_j^{\epsilon'}), \delta_{F_m}(q_0, a_j^{\epsilon'}) \right),$$

in case $\delta_{\widehat{\Delta}_H}(v_0, a_j^{\epsilon'}) \neq \emptyset, \delta_{F_m}(q_0, a_j^{\epsilon'}) \neq \emptyset$. Otherwise, $\delta_H((v_0, q_0), a_j^{\epsilon'}) = \emptyset$.

$$\delta_H((v, q_i^\epsilon), a_j^{\epsilon'}) = \left(\delta_{\widehat{\Delta}_H}(v, a_j^{\epsilon'}), \delta_{F_m}(q_i^\epsilon, a_j^{\epsilon'}) \right),$$

in case $\delta_{\widehat{\Delta}_H}(v, a_j^{\epsilon'}) \neq \emptyset, \delta_{F_m}(q_i^\epsilon, a_j^{\epsilon'}) \neq \emptyset$. Otherwise, $\delta_H((v, q_i^\epsilon), a_j^{\epsilon'}) = \emptyset$.

Remark 5.2. \mathcal{A}_H inherits from \mathcal{A}_{F_m} and $\mathcal{A}_{\widehat{\Delta}_H}$ the property of being deterministic and having only one initial state. In particular, \mathcal{A}_H is an unambiguous automaton.

Recall that $L_H \subseteq \Sigma^*$ denotes the set of reduced words that represent elements of $H \leq F_m$.

Proposition 5.3. $L(\mathcal{A}_H) = L_H$.

Proof. Recall that $L(\mathcal{A}_H) = L(\mathcal{A}_{\widehat{\Delta}_H}) \cap L(\mathcal{A}_{F_m})$. Also $L_H \subseteq L(\mathcal{A}_{\widehat{\Delta}_H}) \cap L(\mathcal{A}_{F_m})$, and $L(\mathcal{A}_{\widehat{\Delta}_H}) \setminus L_H$ consists of only non-reduced words, whereas $L(\mathcal{A}_{F_m})$ consists of only reduced words. Therefore, $L(\mathcal{A}_H) = L_H$. \square

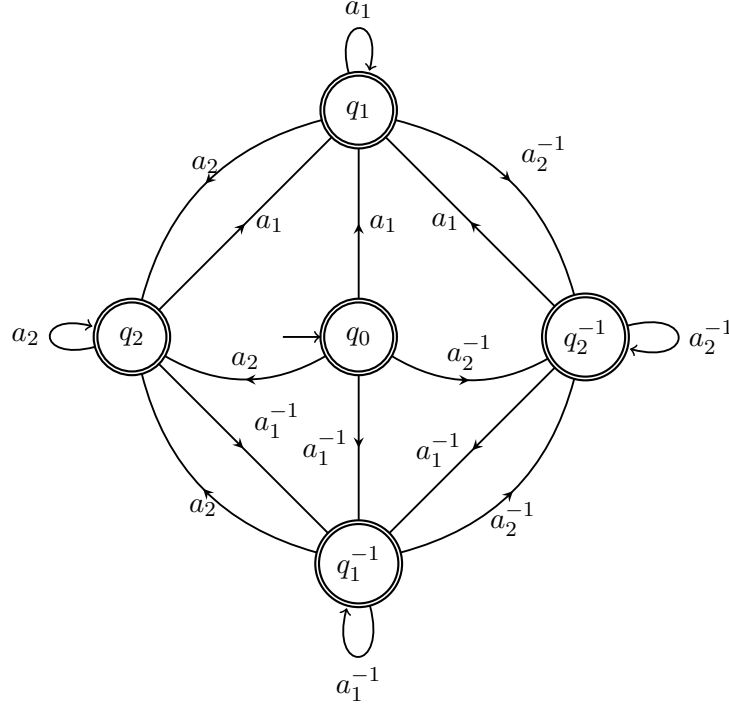


FIGURE 4. The Moore diagram of \mathcal{A}_{F_2}

5.2. Definition and main properties of \mathcal{D}_H . Define $Q_{\mathcal{D}_H} \subseteq \widehat{V} \times Q_{F_m}$ as

$$Q_{\mathcal{D}_H} = \{(v_0, q_0)\} \cup \{(v, q_i^\epsilon) \mid v \in \widehat{V} \text{ and } \exists e \in \widehat{E} \text{ s.t. } a_i^\epsilon = \widehat{\mu}(e), t(e) = v\}.$$

In other words, $Q_{\mathcal{D}_H}$ is the set of accessible states of \mathcal{A}_H .

Lemma 5.4.

- (1) If there is a path p in $\mathcal{A}_{\widehat{\Delta}_H}$ with a reduced label w such that $o(p) = v_0$ and $t(p) = v \in \widehat{V}$, and the suffix of w is a_i^ϵ , then $(v, q_i^\epsilon) \in Q_{\mathcal{D}_H}$.
- (2) If $(v, q_i^\epsilon) \in Q_{\mathcal{D}_H}$, then there exists a reduced word $w_1 a_i^\epsilon w_2 \in L_H$ such that $\delta_H(v_0, w_1 a_i^\epsilon) = v$.

Proof.

- (1) Let $w = a_{j_1}^{\epsilon_1} a_{j_2}^{\epsilon_2} \dots a_{j_l}^{\epsilon_l} \in L_H$ such that $a_{j_l}^{\epsilon_l} = a_i^\epsilon$. Suppose that the path $v_0 \xrightarrow{a_{j_1}^{\epsilon_1}} v_1 \xrightarrow{a_{j_2}^{\epsilon_2}} \dots \xrightarrow{a_{j_l}^{\epsilon_l}} v_l = v$ in $\mathcal{A}_{\widehat{\Delta}_H}$ with label w and starting at v_0 terminates at $v \in \widehat{V}$. Then, by definition of \mathcal{A}_H , the path

$$(v_0, q_0) \xrightarrow{a_{j_1}^{\epsilon_1}} (v_1, q_{j_1}^{\epsilon_1}) \xrightarrow{a_{j_2}^{\epsilon_2}} \dots \xrightarrow{a_{j_l}^{\epsilon_l}} (v_l, q_{j_l}^{\epsilon_l}) = (v, q_i^\epsilon)$$

is well defined in \mathcal{A}_H and has the label w . Also, by definition of $Q_{\mathcal{D}_H}$, $(v_l, q_{j_l}^{\epsilon_l}) \in Q_{\mathcal{D}_H}$.

- (2) Now assume that $(v, q_i^\epsilon) \in Q_{\mathcal{D}_H}$. Then, by definition, there exists $e \in \widehat{E}$ such that $t(e) = v$ and $\widehat{\mu}(e) = a_i^\epsilon$. Assume that $v' = o(e) \neq v_0$. Then, since by Lemma 3.2, the number of incoming edges for v' is at least two, there exists $e' \in \widehat{E}$ such that $t(e') = v'$ and its label $a_{i'}^{\epsilon'}$ is different from a_i^ϵ . Therefore, $(v', q_{i'}^{\epsilon'}) \in Q_{\mathcal{D}_H}$.

Note that, by Lemma 5, $\widehat{\Delta}_H$ has an admissible path $p = p'ep''$ with reduced label such that $t(p') = v'$ and $o(p'') = v$. As the discussion in the proof of part (1) shows, p corresponds to a path \bar{p} in \mathcal{A}_H with the same label as p . Now, if the label of p' is w_1 and the label of p'' is w_2 , then the label of \bar{p} is $w_1a_i^\epsilon w_2$. By part (1) of the lemma, the sub-path of \bar{p} with label $w_1a_i^\epsilon$ will terminate at v . Therefore, by part (1) of the lemma, $\delta_H(v_0, w_1a_i^\epsilon) = v$. \square

Corollary 5.5. $(v, q_i^\epsilon) \in Q_{\mathcal{D}_H}$ if and only if there is an admissible path in \mathcal{A}_H that contains (v, q_i^ϵ) and whose vertices belong to $Q_{\mathcal{D}_H}$.

Proof. The ‘only if’ part follows immediately from part (2) of Lemma 5.4. For the ‘if’ part, note that, by definition, any state that is on some admissible path in \mathcal{A}_H is accessible, hence belongs to $Q_{\mathcal{D}_H}$. \square

Define \mathcal{D}_H to be the subautomaton of \mathcal{A}_H induced by the states $Q_{\mathcal{D}_H}$.

Proposition 5.6. \mathcal{D}_H is the essential part of $\mathcal{A}_H = \mathcal{A}_{\widehat{\Delta}_H} \times \mathcal{A}_{F_m}$. In particular, $L_H = L(\mathcal{D}_H)$.

Proof. The fact that \mathcal{D}_H is essential follows directly from Corollary 5.5. Also, from Lemma 5.4 and the fact that the set of reduced words in $L(\mathcal{A}_{\widehat{\Delta}_H})$ coincides with L_H it follows that $L_H = L(\mathcal{D}_H)$. Combining this with Lemma 5.1 and the fact that \mathcal{A}_H is unambiguous, we get that \mathcal{D}_H is the essential part of \mathcal{A}_H . \square

Remark 5.7. \mathcal{D}_H inherits from \mathcal{A}_H the properties of being deterministic and unambiguous. **Automaton presentation of \mathcal{D}_H .** For expository reasons, in the sequel we replace the notation q_i^ϵ for a state of \mathcal{D}_H by a_i^ϵ , and denote the initial state (v_0, q_0) simply by q^* . Thus \mathcal{D}_H will have the following presentation:

$$\mathcal{D}_H = (Q_{\mathcal{D}_H}, \Sigma, \delta_{\mathcal{D}_H}, \{q^*\}, \mathcal{F}_{\mathcal{D}_H}),$$

where

$$Q_{\mathcal{D}_H} = \{q^*\} \cup \left\{ (v, a_i^\epsilon) \mid v \in \widehat{V}, a_i^\epsilon \in \Sigma, \text{ and } \exists e \in \widehat{E} \text{ s.t. } a_i^\epsilon = \widehat{\mu}(e), t(e) = v \right\},$$

$$\mathcal{F}_{\mathcal{D}_H} = \{q^*\} \cup \left\{ (v_0, a_i^\epsilon) \mid \exists e \in \widehat{E} \text{ s.t. } a_i^\epsilon = \widehat{\mu}(e), t(e) = v_0 \right\},$$

$$\delta_{\mathcal{D}_H}(q^*, a_i^\epsilon) = (v_0 a_i^\epsilon, a_i^\epsilon), \text{ for all } a_i^\epsilon \in \Sigma,$$

$$\delta_{\mathcal{D}_H}\left((v, a_i^\epsilon), a_j^{\epsilon'}\right) = \left(v a_j^{\epsilon'}, a_j^{\epsilon'}\right), \text{ if } a_i^\epsilon \neq (a_j^{\epsilon'})^{-1}.$$

5.2.1. *Minimality of \mathcal{D}_H .* The Myhill-Nerode Theorem (see Theorem 2.1) suggests the existence of a minimal DFA (which is unique up to isomorphism). For any given DFA \mathcal{A} , in order to construct the minimal DFA \mathcal{A}' such that $L(\mathcal{A}) = L(\mathcal{A}')$, we need to define an equivalence relation \equiv among states of \mathcal{A} .

We write $u \equiv v$ if the following holds: for all $w \in \Sigma^*$, $\delta(u, w)$ is a final state if and only if $\delta(v, w)$ is a final state. Notice that the relations R_L (see (2.1)) and \equiv express exactly the same idea, i.e.

$$wR_Lw' \iff \delta(q_0, w) \equiv \delta(q_0, w'). \quad (5.3)$$

The states u and v of \mathcal{A} are *equivalent* if $u \equiv v$. When the states u and v are not equivalent, then we say that they are *distinguishable*. That is, there exists at least one state w such that one of $\delta(u, w)$ and $\delta(v, w)$ is an accepting state and the other is not.

Theorem 5.8. *\mathcal{D}_H is a minimal DFA.*

Proof. In order to show that the DFA \mathcal{D}_H is minimal, according to the Myhill-Nerode Theorem, it is enough to show that any two distinct states (u, a_i^ϵ) and (v, a_j^ν) of \mathcal{D}_H are distinguishable. We prove this by contradiction.

Assume that $(u, a_i^\epsilon) \neq (v, a_j^\nu)$ are R_L -equivalent states. Then $\forall w \in \Sigma^*$, $\delta_{\mathcal{D}_H}((u, a_i^\epsilon), w)$ is a final state if and only if $\delta_{\mathcal{D}_H}((v, a_j^\nu), w)$ is a final state. Let us fix one such w (the existence of w follows from \mathcal{D}_H being essential.) Since \mathcal{D}_H is an essential automaton, there are w_1 and w_2 in Σ^* such that $\delta_{\mathcal{D}_H}(q^*, w_1) = (u, a_i^\epsilon)$ and $\delta_{\mathcal{D}_H}(q^*, w_2) = (v, a_j^\nu)$. Therefore, $w_1w, w_2w \in L(\mathcal{D}_H) = L_H$, the suffix of w_1 is a_i^ϵ , the suffix of w_2 is a_j^ν . Then $w_1(w_2)^{-1} = (w_1w)(w_2w)^{-1} \in H$, which implies that $Hw_1 = Hw_2$, and hence $u = v$. Therefore, it must be that $a_i^\epsilon \neq a_j^\nu$, which, in particular, implies that $w_1(w_2)^{-1} \in L_H$. As the states (u, a_i^ϵ) and (v, a_j^ν) are equivalent, under the assumption that (u, a_i^ϵ) and (v, a_j^ν) are R_L -equivalent, $w_1(w_2)^{-1} \in L_H$ implies $w_2(w_2)^{-1} \in L_H$, which is not true – a contradiction. \square

5.3. **Definition of $\widehat{\mathcal{D}}_H$.** In this subsection we introduce the automaton $\widehat{\mathcal{D}}_H$, which is obtained from \mathcal{D}_H by simply removing its initial state q^* and proclaiming the final states $\mathcal{F}_{\mathcal{D}_H}$ of \mathcal{D}_H at the same time the initial and final states of $\widehat{\mathcal{D}}_H$. Thus the automaton $\widehat{\mathcal{D}}_H$ would have the following presentation:

$$\widehat{\mathcal{D}}_H = \left(Q_{\widehat{\mathcal{D}}_H}, \Sigma, \delta_{\widehat{\mathcal{D}}_H}, \mathcal{I}_{\widehat{\mathcal{D}}_H}, \mathcal{F}_{\widehat{\mathcal{D}}_H} \right),$$

where

$$Q_{\widehat{\mathcal{D}}_H} = \left\{ (v, a_i^\epsilon) \mid v \in \widehat{V}, a_i^\epsilon \in \Sigma, \text{ and } \exists e \in \widehat{E} \text{ s.t. } a_i^\epsilon = \widehat{\mu}(e), t(e) = v \right\},$$

$$\mathcal{I}_{\widehat{\mathcal{D}}_H} = \mathcal{F}_{\widehat{\mathcal{D}}_H} = \left\{ (v_0, a_i^\epsilon) \mid a_i^\epsilon = \widehat{\mu}(e), t(e) = v_0 \right\}, \quad (5.4)$$

$$\delta_{\widehat{\mathcal{D}}_H} \left((v, a_i^\epsilon), a_j^{\epsilon'} \right) = (va_j^{\epsilon'}, a_j^{\epsilon'}), \text{ if } a_i^\epsilon \neq (a_j^{\epsilon'})^{-1}. \quad (5.5)$$

The main motivation behind consideration of the automaton $\widehat{\mathcal{D}}_H$ is that it will define the language L_H and be ergodic whenever L_H is irreducible, and at the same time will provide a convenient way of computing the entropy of L_H . All these will be revealed in detail in the next section. See, for example, Theorems 5.14, 5.18 and Proposition 5.17.

To describe the main properties of $\widehat{\mathcal{D}}_H$, first we need the following definition.

Definition 5.9. $H \leq F_m$ is said to be *conjugacy reduced* if there does not exist $a \in \Sigma$ such that every word from L_H is of the form $a^{-1}wa$. In other words, H is conjugacy reduced if there exist two words from L_H with different last letters.

Proposition 5.10. For $H \neq 1$, $\widehat{\mathcal{D}}_H$ has exactly one initial state if and only if H is not conjugacy reduced. Moreover, in such case this state is isolated in the sense that there is no outgoing edge from it in the Moore diagram.

Proof. By definition of $\widehat{\mathcal{D}}_H$, it has only one initial state if and only if there is only one incoming edge for the root vertex of the extended core $\widehat{\Delta}_H$. The last property for $\widehat{\Delta}_H$ is equivalent to saying that all words from L_H end with the same letter.

Finally, if the words in L_H start with a^{-1} and end with a for some $a \in \Sigma$, then the initial state in $\widehat{\mathcal{D}}_H$ is (v_0, a) , which does not have an outgoing edge, because edges from (v_0, a) with label a^{-1} are not permitted by the definition of $\widehat{\mathcal{D}}_H$. Thus the proposition is proved. \square

Lemma 5.11. If $H \leq F_m$ is conjugacy reduced, then $\widehat{\mathcal{D}}_H$ is an essential automaton and $L_{\widehat{\mathcal{D}}_H} = L_H$.

Proof. By Proposition 5.10, H is conjugacy reduced means that $\mathcal{I}_{\widehat{\mathcal{D}}_H}$ contains at least two elements (v_0, a) and (v_0, b) for $a \neq b \in \Sigma$. This means that whichever state one can move from q^* in \mathcal{D}_H is possible to move from at least one state from $\mathcal{I}_{\widehat{\mathcal{D}}_H}$ (in fact, from precisely $|\mathcal{I}_{\widehat{\mathcal{D}}_H}| - 1$ states.) Therefore, the fact that $\widehat{\mathcal{D}}_H$ is essential follows from \mathcal{D}_H being essential, see Proposition 5.6.

On the other hand, every state that is possible to attain from the states $\mathcal{I}_{\widehat{\mathcal{D}}_H}$ in $\widehat{\mathcal{D}}_H$ is possible to attain from q^* in \mathcal{D}_H . Therefore, $L_{\widehat{\mathcal{D}}_H} = L_{\mathcal{D}_H} = L_H$. \square

Proposition 5.12. If $H \leq F_m$ is conjugacy reduced, then $\widehat{\mathcal{D}}_H$ is deterministic and has homogeneous ambiguity $\deg(v_0) - 1$.

Proof. The fact that $\widehat{\mathcal{D}}_H$ is deterministic follows straightforwardly from the definition of $\widehat{\mathcal{D}}_H$. The fact that $\widehat{\mathcal{D}}_H$ has homogeneous ambiguity $\deg(v_0) - 1$ follows from the observations that \mathcal{D}_H is a DFA (hence, is unambiguous) and the states that can be attained from q^* in \mathcal{D}_H are precisely the states that can be attained from precisely $|\mathcal{I}_{\widehat{\mathcal{D}}_H}| - 1$ initial states of $\widehat{\mathcal{D}}_H$ (note that except the initial states, the states of \mathcal{D}_H and $\widehat{\mathcal{D}}_H$ coincide). Therefore, $\widehat{\mathcal{D}}_H$ has homogeneous ambiguity $|\mathcal{I}_{\widehat{\mathcal{D}}_H}| - 1 = \deg(v_0) - 1$. \square

5.4. Ergodicity of $\widehat{\mathcal{D}}_H$. The main goal of this subsection is to give a complete description of when $\widehat{\mathcal{D}}_H$ is ergodic for finitely generated subgroups H of F_m . In Subsection 5.5 we discuss some of its consequences that are connected with the entropy of L_H . Everywhere in this section we assume that H is a non-trivial finitely generated subgroup of the free group $F_m = F(A)$, $\Sigma = A \cup A^{-1}$.

A finite state automaton is said to be *ergodic* if its Moore diagram, regarded as a directed graph, is strongly connected. That is for each ordered pair of vertices of the Moore diagram there is a path from the first to the second vertex. An alternative characterization of ergodicity of a finite state automaton is that the formal language $L \subseteq \Sigma^*$ that it defines is *irreducible*, which means that for each $w_1, w_2 \in L$ there exists $w \in \Sigma^*$ such that $w_1ww_2 \in L$.

See [4]. Yet another characterization of ergodicity is that the adjacency matrix M of the Moore diagram of the automaton is *irreducible*: if M is of size $m \times m$, then for each $1 \leq i, j \leq m$, there exists $n \in \mathbb{N}$ such that the (i, j) -th coefficient of M^n is strictly positive. The latest property of ergodic automata of bounded ambiguity can be used to compute the entropy of the corresponding language, which we discuss later in Subsection 5.5.

Proposition 5.13. *$1 \neq H \leq F_m$ is conjugacy reduced and cyclic if and only if there exist $a \neq b \in \Sigma$ such that all words from L_H either start with a^{-1} and end with b or start with b^{-1} and end with a . Moreover, in such case $\widehat{\mathcal{D}}_H$ has exactly two initial states, (v_0, a) and (v_0, b) , which belong to two different connected components in the Moore diagram of $\widehat{\mathcal{D}}_H$.*

Proof. First, assume that L_H satisfies to the property from the statement. Then it follows from Proposition 5.10 that H is conjugacy reduced. Now, by contradiction assume that H is not cyclic. Then, let $a^{-1}w_1b$ be the shortest word in L_H ending with b and let $a^{-1}w_2b \in L_H$ be the shortest word ending with b that is not contained in the maximal cyclic subgroup of H containing $a^{-1}w_1b \in H$. Since H is cyclically reduced, $a \neq b^{-1}$. Then $1 \neq \text{red}(b^{-1}w_2^{-1}w_1b) \in L_H$. However, since by the assumption no word in L_H starts with b^{-1} and ends with b , it must be that after the free cancellation of $b^{-1}w_2^{-1}w_1b$ either the prefix or the suffix of $b^{-1}w_2^{-1}w_1b$ cancels out. In the first case we end up with a word that is shorter than $a^{-1}w_1b$, in the second case we end up with a word that is shorter than $a^{-1}w_2b$. Therefore, we end up with a contradiction because of the minimality condition on the lengths of $a^{-1}w_1b$ and $a^{-1}w_2b$. Thus H is cyclic.

Now assume that H is cyclic and conjugacy reduced. Then there exists $w \in H$ that generates H . In other words, $L_H = \{w^n \mid n \in \mathbb{Z}\}$. Since H is conjugacy reduced, the prefix of w is not the inverse of the suffix. Therefore, the prefix corresponds to a^{-1} , the suffix corresponds to b , and the “only if” part of the first statement of the proposition follows as well.

Now let us show the second statement. The condition that all words of L_H are of the form $(a^{-1}wb)^{\pm 1}$ implies that the root vertex v_0 of $\widehat{\Delta}_H$ has exactly two incoming edges with corresponding labels a and b , and it has exactly two outgoing edges with the same labels. This means that $\mathcal{I}_{\widehat{\mathcal{D}}_H} = \mathcal{F}_{\widehat{\mathcal{D}}_H} = \{(v_0, a), (v_0, b)\}$. Moreover, since by definition of $\widehat{\mathcal{D}}_H$ there is no outgoing edge of (v_0, a) with label a^{-1} in the Moore diagram, we have that the label of any accepted path that starts at (v_0, a) starts with b^{-1} and hence ends with a^{-1} , hence is a loop. The same statement in regard with (v_0, b) is true as well. Therefore, the second assertion of the proposition follows as well. \square

Theorem 5.14. *Let H be a finitely generated subgroup of F_m . Then $\widehat{\mathcal{D}}_H$ is an ergodic automaton if and only if H is conjugacy reduced and non cyclic.*

Proof. First, assume that H is conjugacy reduced and non cyclic.

By Lemma 5.11, $\widehat{\mathcal{D}}_H$ is essential, which means that any state of $\widehat{\mathcal{D}}_H$ is on some admissible path. Therefore, since $\mathcal{I}_{\widehat{\mathcal{D}}_H} = \mathcal{F}_{\widehat{\mathcal{D}}_H}$, in order to show that $\widehat{\mathcal{D}}_H$ is strongly connected, it is enough to show that for any ordered pair of different initial (=final) stats of $\widehat{\mathcal{D}}_H$, there is a directed path connecting the first state to the second one. (Note that, by Proposition 5.10, $\widehat{\mathcal{D}}_H$ has at least two different initial states.)

Let $(v_0, a), (v_0, b), a \neq b \in \Sigma$, be an arbitrary pair of initial states. We want to show that there exists a path from (v_0, a) to (v_0, b) , which is equivalent to the existence of a word $w \in L_H$ which does not start with a^{-1} and ends with b , as in such case $\delta_{\widehat{\mathcal{D}}_H}((v_0, a), w) = (v_0, b)$.

The fact that $(v_0, b) \in \mathcal{S}_{\widehat{\mathcal{D}}_H}$ means that there exists $u \in L_H$ such that it ends with letter b . Since by our assumption H is not cyclic, there exists $1 \neq v \in L_H$ such that $\langle u \rangle \cap \langle v \rangle = \{1\}$. Then, for large enough $n \in \mathbb{N}$, the word $\bar{v} := \text{red}(u^{-n}vu^n) \in L_H$ starts with b^{-1} and ends with b . Therefore, $\delta_{\widehat{\mathcal{D}}_H}((v_0, a), \bar{v}) = (v_0, b)$. Thus we showed that if H is conjugacy reduced and non cyclic, then $\widehat{\mathcal{D}}_H$ is ergodic.

Now assume that H is not conjugacy reduced. Then, by Proposition 5.10, $\widehat{\mathcal{D}}_H$ has exactly one initial state, which is isolated. Therefore, in such case $\widehat{\mathcal{D}}_H$ is not ergodic.

Finally, assume that H is conjugacy reduced but cyclic. Then, by Proposition 5.13, $\widehat{\mathcal{D}}_H$ has exactly two initial states which are not connected one to the other. Therefore, in such case $\widehat{\mathcal{D}}_H$ is not ergodic either. Thus the theorem is proved. \square

Combining Theorem 5.14 with Propositions 5.10 and 5.13, we get the following corollary.

Corollary 5.15. *If $|\mathcal{S}_{\widehat{\mathcal{D}}_H}| > 2$, then $\widehat{\mathcal{D}}_H$ is ergodic.*

Remark 5.16. Besides the case when H is cyclic and conjugacy reduced, the only other case when $|\mathcal{S}_{\widehat{\mathcal{D}}_H}| = 2$ is when, for some $a \in \Sigma$, L_H consists of words only, and inclusively, of the forms $awa, awa^{-1}, a^{-1}wa^{-1}$ or $a^{-1}wa$. The latest corresponds, for example, to the case when $H = \langle aba^{-1}, a^2 \rangle < F(a, b) = F_2$.

Proposition 5.17. *For $H \leq F_m$, L_H is irreducible if and only if H is conjugacy reduced and non cyclic.*

Proof. The ‘if’ part of the statement follows from Theorem 5.14.

Now assume that H is not conjugacy reduced, then for some $a \in \Sigma$, elements of L_H are (reduced) words of the form $a^{-1}wa$. Since concatenation of any such words is not reduced, it means that L_H is not irreducible in such case.

Finally, assume that $H = \langle w \rangle$ is cyclic. Then, $w, w^{-1} \in L_H$. However, there is no $v \in L_H$ such that wvw^{-1} is also in L_H . Therefore, L_H is not irreducible in this case either. Thus the proposition is proved. \square

5.5. Entropy of L_H . Recall that for a formal language L its *entropy*, $h(L)$, which is a fundamental numerical invariant of a language, is defined as

$$h(L) = \limsup_{n \rightarrow \infty} \frac{\log |B_n(L)|}{n}, \tag{5.6}$$

where $B_n(L)$ is the subset of L of words of length n .

If L is a language generated by an automaton \mathcal{A} , then by $h(\mathcal{A})$ we denote $h(L(\mathcal{A}))$. If for $H \leq F_m$, then the entropy of H is $h(H) = h(L_H)$.

One important feature of ergodicity of $\widehat{\mathcal{D}}_H$ is that the adjacency matrix $M_{\widehat{\mathcal{D}}_H}$ of $\widehat{\mathcal{D}}_H$ is irreducible and is of bounded ambiguity (as follows from Theorem 5.14 and Proposition 5.12). Therefore, one can apply the Perron-Frobenius theory to obtain the following theorem on entropy of L_H . (For discussion on Perron-Frobenius theory see Chapter 4 in [17].)

Theorem 5.18. *If $H \leq F_m$ is a conjugacy reduced and non cyclic finitely generated group, then the entropy of L_H is equal to $\log \lambda$, where λ is the maximal eigenvalue of the adjacency matrix $M_{\widehat{\mathcal{D}}_H}$ of $\widehat{\mathcal{D}}_H$.*

Remark 5.19. If H is cyclic, then its entropy is equal to 0. If H is not conjugacy reduced, then there is $g \in F_m$ such that $gHg^{-1} \leq F_m$ is conjugacy reduced, and $h(H) = h(gHg^{-1})$.

Definition 5.20 (Base automaton). Let \mathcal{A} be a finite automaton and let (V, E) be its underlying graph. Then *the base automaton* $\check{\mathcal{A}}$ of \mathcal{A} is the automaton with the same underlying graph (V, E) and such that

- (1) all states of $\check{\mathcal{A}}$ are at the same time initial and final states,
- (2) all edges of its Moore diagram are labeled with different labels. To be more specific, we assume that each $e \in E$ is labeled by the letter x_e .

To prove Theorem 5.18, we need the following fact.

Proposition 5.21. *If \mathcal{A} is an essential automaton with bounded ambiguity, then $h(\mathcal{A}) = h(\check{\mathcal{A}})$.*

Proof of Theorem 5.18. Since for conjugacy reduced and non cyclic finitely generated $H \leq F_m$, by Lemma 5.11 and Proposition 5.12, $\widehat{\mathcal{D}}_H$ is essential and of bounded ambiguity, then, by Proposition 5.21, $h(\widehat{\mathcal{D}}_H) = h(\check{\mathcal{D}}_H)$, where by $\check{\mathcal{D}}_H$ we denote the base automaton of $\widehat{\mathcal{D}}_H$. Additionally, by Theorem 5.14, $\widehat{\mathcal{D}}_H$ is an ergodic automaton, which implies that $\check{\mathcal{D}}_H$ is ergodic as well. Moreover, the adjacency matrices of $\widehat{\mathcal{D}}_H$ and $\check{\mathcal{D}}_H$ coincide and are irreducible. Therefore, the proof of Theorem 5.18 follows from Perron-Frobenius theory. See, for example, Theorem 4.4.4. in [17]. \square

Proof of Proposition 5.21. Assuming that \mathcal{A} is an essential automaton, for every vertex $v \in V$ in its underlying graph (V, E) , there exist paths that connect v to an initial state and to a final state, respectively. For each v let us fix a pair (\check{v}, \hat{v}) of such paths, where \check{v} is a path that connects an initial state to v and \hat{v} is a path that connects v to a final state. Also, for each path p in (V, E) let us denote by p_- its origin and by p_+ is terminus, and by $\phi(p)$ let us denote its label in the Moore diagram of \mathcal{A} and by $\psi(p)$ let us denote its label in the Moore diagram of $\check{\mathcal{A}}$. Note that for each path p , $\psi(p) \in L(\check{\mathcal{A}})$. Then we have the following map $\Lambda : L(\check{\mathcal{A}}) \rightarrow L(\mathcal{A})$: for each path p in (V, E) , define

$$\Lambda(\psi(p)) = \phi(\check{p}_- p \hat{p}_+).$$

Lemma 5.22. *There exists $C \in \mathbb{N}$ such that for each path p from (V, E) , $|p| \leq |\Lambda(p)| \leq |p| + C$.*

Proof. The left inequality is obvious. For the right inequality, one can take $C = \max\{|\check{v}| \mid v \in V\} + \max\{|\hat{v}| \mid v \in V\}$. \square

From Lemma 5.22, for all $n \in \mathbb{N}$, we immediately get

$$\Lambda(B_n(L(\check{\mathcal{A}}))) \subseteq \bigcup_{i=0}^C B_{n+i}(L(\mathcal{A})). \quad (5.7)$$

Lemma 5.23. *If \mathcal{A} has bounded ambiguity, then there exists $D > 0$ such that for all $w \in L(\mathcal{A})$, $|\Lambda^{-1}(w)| \leq D$.*

Proof. Let $K > 0$ be an upper bound of ambiguity of \mathcal{A} . Let $w \in L(\mathcal{A})$. Also, let $(u, v) \in V \times V$. Then the number of paths p in (V, E) such that $p_- = u$, $p_+ = v$, and $\Lambda(\psi(p)) = w$ is bounded from above by K . The number of pairs $(u, v) \in V \times V$ is equal to $|V|^2$. Therefore, one can take $D = K|V|^2$. \square

Lemma 5.24. *For all $n \in \mathbb{N}$, there exists $0 \leq c_n \leq C$ such that*

$$|B_n(L(\mathcal{A}))| \leq |B_n(L(\check{\mathcal{A}}))| \leq (C + 1)D|B_{n+c_n}(L(\mathcal{A}))|. \tag{5.8}$$

Proof. The left inequality of (5.8) is trivial. For the right one, note that for all $n \in \mathbb{N}$, by Lemma 5.23 and by (5.7), we get

$$D \sum_{i=0}^C |B_{n+i}(L(\mathcal{A}))| \geq |B_n(L(\check{\mathcal{A}}))|.$$

Therefore, one can take c_n in (5.8) to be such that

$$|B_{n+c_n}(L(\mathcal{A}))| = \max\{|B_{n+i}(L(\mathcal{A}))| \mid 0 \leq i \leq C\}.$$

□

From Lemma 5.24 we immediately get that there exists a constant $C' > 0$ and a sequence $\{c_n\}_{n=1}^\infty$, $0 \leq c_n \leq C'$, such that

$$\frac{\log |B_n(L(\mathcal{A}))|}{n} \leq \frac{\log |B_n(L(\check{\mathcal{A}}))|}{n} \leq \frac{\log C' |B_{n+c_n}(L(\mathcal{A}))|}{n}.$$

Proposition 5.21 follows immediately from the last inequality.

6. COMPUTING COGROWTH SERIES OF H

With an arbitrary subgroup H of F_m one associates the *growth function*

$$\gamma_H(n) = |H_n|,$$

where H_n is the set of elements in H of length n with respect to the basis A of F_m i.e. length of the element is the length of the reduced word from Σ^* , where $\Sigma = A \cup A^{-1}$ representing the element. Also, following [11] we introduce the *cogrowth series*

$$H(z) = \sum_{n=0}^\infty |H_n|z^n \tag{6.1}$$

The upper limit

$$\alpha_H = \limsup_{n \rightarrow \infty} |H_n|^{1/n} \tag{6.2}$$

is called the *growth rate* of H with respect to the basis A of F_m . The radius of convergence of the series (6.1) is

$$R = \frac{1}{\alpha_H}$$

and

$$R \geq \frac{1}{2m-1} \text{ as, for } H = F_m, |H_n| = 2m(2m-1)^{n-1} \text{ if } n \geq 1.$$

The cogrowth series (6.1) represents a function of complex variable $z \in \mathbb{C}$ analytic at disc around $z = 0$ of radius $R \geq \frac{1}{2m-1}$. The major questions are: Under what conditions $H(z)$ is rational, algebraic and belongs to distinguished class of analytic functions, like for instance the class of D -finite functions studied in [6].

Let us recall the second author's argument from [11] on rationality of $H(z)$ (when H is finitely generated) using a Nielsen system of generators of H .

6.1. The Nielsen basis approach. Let $H = \langle w_1, \dots, w_k \rangle$ and let $\{w_i\}_{i=1}^k$ be a Nielsen system of generators. Given two reduced words $u, v \in \Sigma^*$ denote by $\beta(u, v)$ the number of a -symbols (by a -symbols we mean elements of Σ) which will be cancelled when reducing the product $u \cdot v$. For each $i, 1 \leq i \leq k$ and $\epsilon \in \{-1, 1\}$ let $H_n^{i, \epsilon}$ be the set of words of length n in H that can be presented by S -reduced product of generators from $S \cup S^{-1}, S = \{w_1, \dots, w_k\}$ ending with w_i^ϵ . That is

$$u \in H_n^{i, \epsilon} \iff u = \text{red}(w_{i_1}^{\epsilon_1} \cdots w_{i_l}^{\epsilon_l} w_i^\epsilon)$$

for some $i_1, \dots, i_l, \epsilon_1, \dots, \epsilon_l \in \{-1, 1\}$ and in the last product none of the factors $w_{i_j}^{\epsilon_j}$ is inverse to the previous or next factor. Let us recall the corollary of the Statement 3.6 of [11].

Theorem 6.1. *If H is a finitely generated subgroup of F_m , then $H(z)$ is rational.*

Proof. Let $H_i^\epsilon(z) = \sum_{n=0}^{\infty} |H_n^{i, \epsilon}| z^n$. Then

$$H(z) = 1 + \sum_{i, \epsilon} H_i^\epsilon(z).$$

The equation (3.5) from [11] shows that the functions $H_i^\epsilon(z)$ satisfy the system of linear equations.

$$H_i^\epsilon(z) = z^{|w_i^\epsilon|} + \sum_{j, \epsilon'} z^{|w_i^\epsilon| - \beta(w_j^{\epsilon'}, w_i^\epsilon)} H_j^{\epsilon'}(z) \quad (6.3)$$

Taking the summation term to the left, we can rewrite the above system as follows

$$BY = Z \quad (6.4)$$

where B is a $2k \times 2k$ matrix with $(i, \epsilon), (j, \epsilon')$ -th entry

$$B((i, \epsilon), (j, \epsilon')) = \chi_i(j) - z^{|w_i^\epsilon| - \beta(w_j^{\epsilon'}, w_i^\epsilon)},$$

where

$$\chi_i(j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Y and Z are $2k \times 1$ column vectors whose i -th entries are

$$Y_i = H_i^\epsilon(z), \quad Z_i = z^{|w_i^\epsilon|},$$

respectively. Observe that, when $z \in \mathbb{R}$ and $|z| < 1$, the determinant of the matrix B is non-zero. Hence the system (6.3) has a unique solution. Solving this system by standard methods (for instance using Cramer's rule) we get a solution.

$$H_i^\epsilon(z) = \frac{P_i^\epsilon(z)}{Q_i^\epsilon(z)}, \quad i = 1, \dots, k; \epsilon \in \{-1, 1\},$$

where $P_i^\epsilon, Q_i^\epsilon$ are polynomials and hence we get the rational expression for $H(z)$

$$H(z) = 1 + \sum_{i, \epsilon} \frac{P_i^\epsilon(z)}{Q_i^\epsilon(z)}.$$

□

6.2. The finite automata approach. In the previous section we have seen the construction of the DFA \mathcal{D}_H . It is a very old observation, going back to Chomsky and Schützenberger and even to A. Kolmogorov in view of his theory of finite Markov chains, that the growth series $L_H(z)$ of the language $L(\mathcal{D}_H) = L_H$ which by Proposition 5.6 from the previous sections coincides with $H(z)$ and are rational. Moreover, it can be computed using a standard method which often is called the transfer matrix method. See page 573 of [23] or Section V.5 and in particular Proposition V.6 of [5]. In this section, using the DFA \mathcal{D}_H we shall compute the cogrowth series $H(z)$. At the end, we shall present the computations of growth using Nielsen set of generators.

Let M be the adjacency matrix of the labelled directed graph G with t vertices. Assume that, for every vertex of the graph G , all outgoing edges carry distinct labels. Let M^n be the n th power of M . It is well known that the (i, j) th entry of M^n which we denote by $M^n(i, j)$ is just the number of paths of length n from the i th vertex to the j th vertex of G . To evaluate the $M^n(i, j)$, we shall use the transfer matrix method. This method uses linear algebra to analyze the behavior of the $M^n(i, j)$. Following [23], we define the growth series (also known as generating function) of paths in G from i to j .

$$\gamma_{ij}(G, z) = \sum_{n \geq 0} M^n(i, j)z^n.$$

Observe that $\gamma_{ij}(G, z)$ is the (i, j) th entry of the matrix

$$\sum_{n \geq 0} M^n z^n = (I - zM)^{-1}$$

where I is the identity matrix of dimension $t \times t$. In order to compute the (i, j) th entry $\gamma_{ij}(G, z)$, we recall Theorem 4.7.2 of [23].

Theorem 6.2. *The growth series $\gamma_{ij}(G, z)$ is given by*

$$\gamma_{ij}(G, z) = \frac{(-1)^{i+j} \det(I - zM : j, i)}{\det(I - zM)} \tag{6.5}$$

where $(B : j, i)$ denotes the minor obtained by removing the j th row and i th column of B . Thus in particular $\gamma_{ij}(G, z)$ is a rational function of z whose degree is strictly less than the multiplicity n_0 of 0 as an eigenvalue of M .

Let $L_H = L(\mathcal{D}_H)$ be the language accepted by the DFA \mathcal{D}_H constructed in the previous Section. Let $M_{\mathcal{D}_H}$ be the adjacency matrix of the Moore diagram of \mathcal{D}_H and $L_H(z)$ be the growth of L_H , i.e.

$$L_H(z) = \sum_{w \in L_H} z^{|w|}.$$

Recall that \mathcal{D}_H has $|Q_{\mathcal{D}_H}| = 1 + \sum_{v \in \widehat{V}} \deg(v)$ states. In our numeration of states of \mathcal{D}_H we start

with the initial state q^* first and then the states $(v_k, a_i^{\epsilon_i})$ where $k = 0, \dots, (|\widehat{V}| - 1)$, $i \in \{1, \dots, m\}$ and $\epsilon_i \in \{1, -1\}$. Let \mathbf{i} be the standard unit row vector of dimension $1 \times |Q_{\mathcal{D}_H}|$ and let the column vector \mathbf{f} be the characteristic vector of the set of final states of dimension $|Q_{\mathcal{D}_H}| \times 1$.

Proposition 6.3.

$$H(z) = L_H(z) = \mathbf{i} \cdot (I - zM_{\mathcal{D}_H})^{-1} \cdot \mathbf{f}, \tag{6.6}$$

where I is the identity matrix of order $|Q_{\mathcal{D}_H}| \times |Q_{\mathcal{D}_H}|$.

Proof. Recall that $L(\mathcal{D}_H) = L_H =$ set of reduced elements of H which implies that $H(z) = L_H(z)$. For every $w \in L_H$ of length $n \geq 0$ we have a unique admissible path p in the Moore diagram $G_{\mathcal{D}_H}$ of \mathcal{D}_H such that $w = l(p)$. Therefore, we write the growth series $L_H(z)$ as

$$\begin{aligned} H(z) = L_H(z) &= \sum_{q \in \mathcal{F}_{\mathcal{D}_H}} \gamma_{q^*q}(G_{\mathcal{D}_H}, z) \\ &= \mathbf{i} \cdot (I - zM_{\mathcal{D}_H})^{-1} \cdot \mathbf{f}. \end{aligned}$$

□

Remark 6.4. Observe that from (5.6) and (6.2) one deduces

$$\alpha_H = e^{h(L_H)},$$

where $h(L_H)$ is the entropy of L_H .

7. EXAMPLES

Let H be a nontrivial finitely generated subgroup of F_m . In this section we shall compute the cogrowth $H(z)$ of H .

7.1. Computations of cogrowth using DFA \mathcal{D}_H . In each of the examples below, we consider DFA \mathcal{D}_H that recognize language L_H of reduced elements of H . Let $M_{\mathcal{D}_H}$ be the adjacency matrix of the Moore diagram of \mathcal{D}_H .

(1) Let H be a finite index subgroup of F_m .

- (a) $H = \langle a^2, b, c, aba^{-1}, aca^{-1} \rangle$ is a subgroup of F_3 . In this case, the core Δ_H is a complete graph. i.e. for each vertex there are three outgoing and three incoming edges labelled by a, b, c and their inverses, respectively. Hence Δ_H coincides with the Schreier graph Γ of H . The index of H in $F_3 = |\widehat{V}| = 2$. Moreover, any vertex of Δ_H can be considered as a root vertex, which implies that H in fact is a normal subgroup. The diagram of core Δ_H and the automaton \mathcal{D}_H are shown in the Figures (5a) and (5b), respectively. The adjacency matrix $M_{\mathcal{D}_H}$ of $G_{\mathcal{D}_H}$ is given below

$$M_{\mathcal{D}_H} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Applying formula (6.6) we get

$$H(z) = \frac{(1+z)(1-4z+5z^2)}{(1-5z)(1-2z+5z^2)} \text{ and } \alpha_H = 5.$$

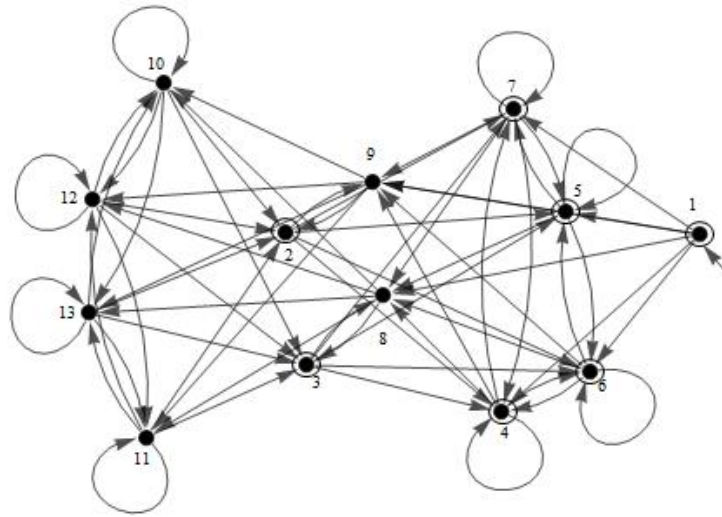
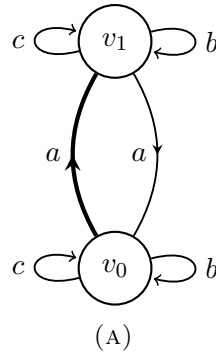


FIGURE 5. (5a) The core graph Δ_H of H . The highlighted spanning tree T in Δ_H can be used to find a free basis $\{a^2, b, c, aba^{-1}, aca^{-1}\}$ of H . (5b) is the Moore diagram of \mathcal{D}_H . The initial state has label 1 and it is indicated by an incoming arrow. The final states of \mathcal{D}_H have labels 1, 2, 3, 4, 5, 6 and 7, respectively and they are indicated by the circles around the states.

- (b) $H = \langle a^3, ab, ab^{-1}, a^{-1}ba \rangle$ is a subgroup of F_2 . In this case, the core Δ_H coincides with the Schreier graph Γ of H . The index of H in $F_2 = |\widehat{V}| = 3$. Moreover, any vertex of Δ_H can not be considered as a root vertex, and hence H is not a normal subgroup. See Figures (6a) and (6b) for the core Δ_H and the diagram of the

automaton \mathcal{D}_H , respectively. The adjacency matrix $M_{\mathcal{D}_H}$ is

$$M_{\mathcal{D}_H} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Applying formula (6.6) we get

$$H(z) = \frac{(1+z)(1-2z+5z^2-6z^3+9z^4)}{(1-3z)(1-z+3z^2)(1+3z+3z^2)} \text{ and } \alpha_H = 3.$$

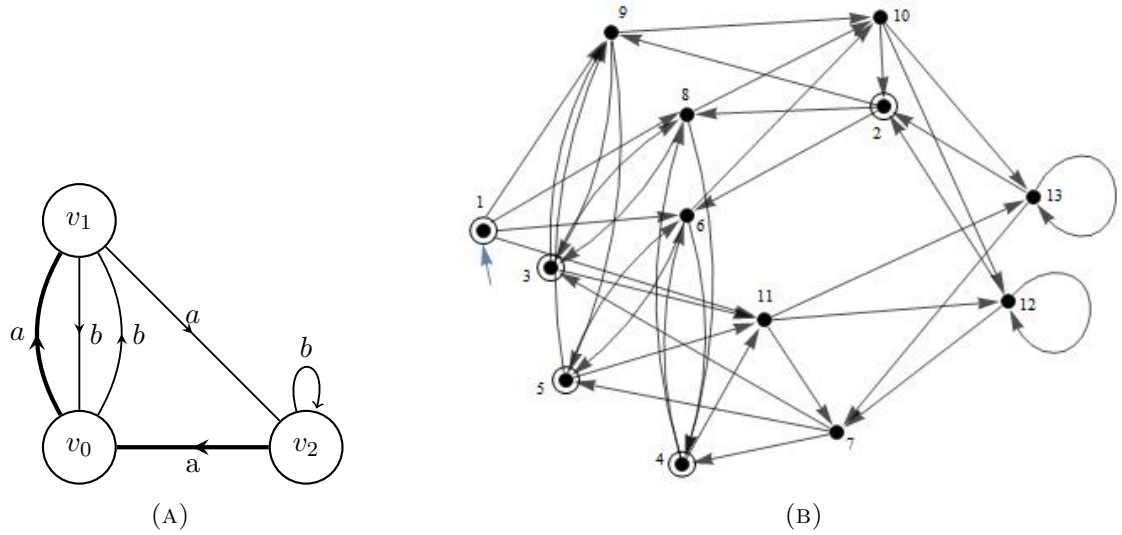


FIGURE 6. (6a) The core graph Δ_H of $H = \langle a^3, ab, ab^{-1}, a^{-1}ba \rangle$. (6b) is the Moore diagram of \mathcal{D}_H .

- (2) Let H be an infinite index subgroup of F_m .
 - (a) $H = \langle aba^{-1}, aca^{-1} \rangle$ is a subgroup of F_3 . In this case, the core Δ_H is incomplete graph and it is a subgraph of Γ . So H is an infinite index subgroup of F_3 . See Figures (7a) and (7b) for the core Δ_H and the diagram of the automaton \mathcal{D}_H ,

respectively. The $\deg(v_0) = 1$. The adjacency matrix $M_{\mathcal{D}_H}$ of $G_{\mathcal{D}_H}$ is

$$M_{\mathcal{D}_H} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}$$

Applying formula (6.6) we get

$$H(z) = 1 + \frac{4z^3}{1 - 3z} \text{ and } \alpha_H = 3.$$

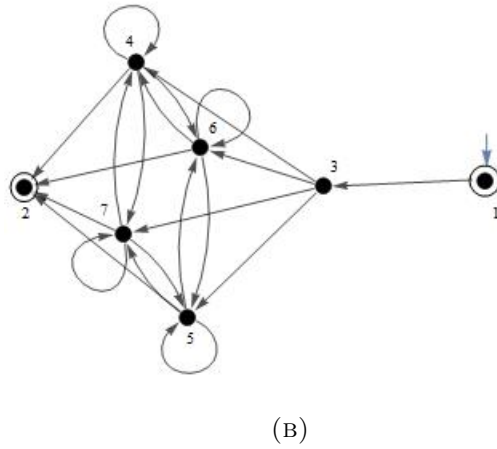
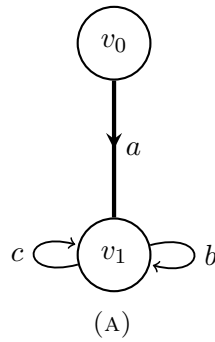


FIGURE 7. (7a) is the core Δ_H of $H = \langle aba^{-1}, aca^{-1} \rangle$. (7b) is the Moore diagram of \mathcal{D}_H .

- (b) $H = \langle b^2, bab^{-1}a^{-1}, a^3 \rangle$ is subgroup of F_2 . In this case, the core Δ_H is incomplete and it is a subgraph of Γ . So H is an infinite index subgroup of F_2 . See Figures (8a) and (8b) for the core Δ_H and the diagram of the automaton \mathcal{D}_H , respectively.

The $\deg(v_0) = 4$. The adjacency matrix $M_{\mathcal{D}_H}$ of $G_{\mathcal{A}}$ is

$$M_{\mathcal{D}_H} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Applying formula (6.6) we get

$$H(z) = \frac{1 + z + 3z^2 + 3z^3 + 5z^4 + 5z^5 + 6z^6 + 6z^7 + 4z^8 + 4z^9}{1 + z + z^2 - z^3 - 5z^4 - 13z^5 - 16z^6 - 20z^7 - 12z^8 - 12z^9}$$

and $\alpha_H = 1.88233$.

7.2. Computations of cogrowth using a Nielsen basis of H . In each of the examples below, we compute the growth series by solving the system (6.3) explained in the Section 6. We shall consider the same set of examples that we have discussed in the previous section (7.1).

(1) Let H be a finite index subgroup of F_m .

(a) $H = \langle a^2, b, c, aba^{-1}, aca^{-1} \rangle$ is a subgroup of F_3 . Solving the system

$$\begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix},$$

where

$$B_1 = \begin{pmatrix} 1 - z^2 & 0 & -z^2 & -z^2 & -z^2 \\ 0 & 1 - z^2 & -z^2 & -z^2 & -z^2 \\ -z & -z & 1 - z & 0 & -z \\ -z & -z & 0 & 1 - z & -z \\ -z & -z & -z & -z & 1 - z \end{pmatrix}, B_2 = \begin{pmatrix} -z^2 & -1 & -1 & -1 & -1 \\ -z^2 & -z^2 & -z^2 & -z^2 & -z^2 \\ -z & -z & -z & -z & -z \\ -z & -z & -z & -z & -z \\ 0 & -z & -z & -z & -z \end{pmatrix},$$

$$B_3 = \begin{pmatrix} -z & -z & -z & -z & 0 \\ -z^3 & -z & -z^3 & -z^3 & -z^3 \\ -z^3 & -z & -z^3 & -z^3 & -z^3 \\ -z^3 & -z & -z^3 & -z^3 & -z^3 \\ -z^3 & -z & -z^3 & -z^3 & -z^3 \end{pmatrix}, B_4 = \begin{pmatrix} 1 - z & -z & -z & -z & -z \\ -z^3 & 1 - z & 0 & -z & -z \\ -z^3 & 0 & 1 - z & -z & -z \\ -z^3 & -z & -z & 1 - z & 0 \\ -z^3 & -z & -z & 0 & 1 - z \end{pmatrix},$$

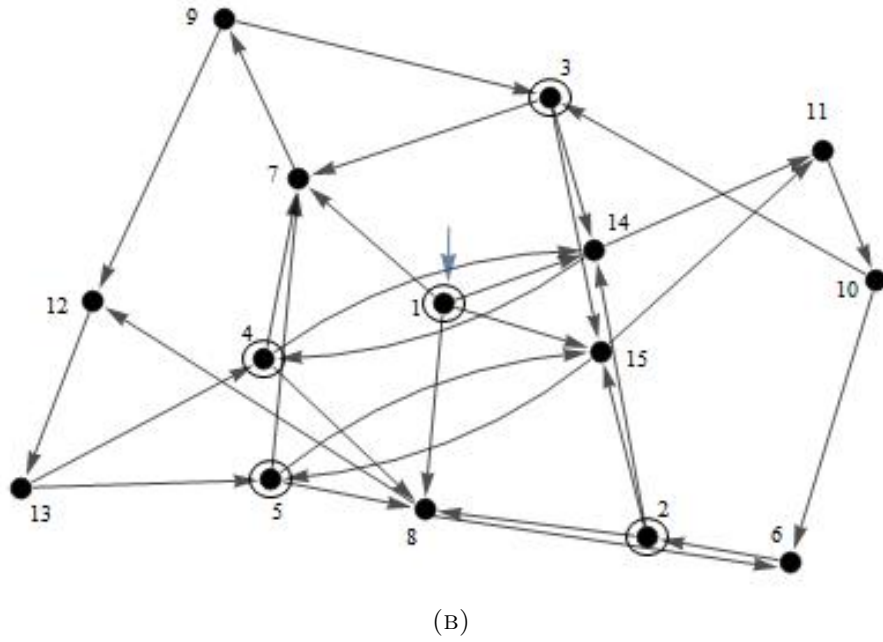
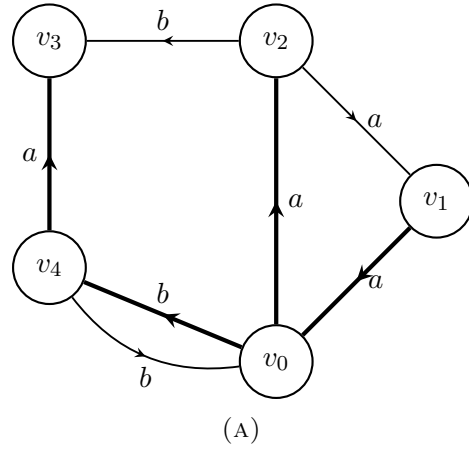


FIGURE 8. (8a) The core graph Δ_H of $H = \langle b^2, bab^{-1}a^{-1}, a^3 \rangle$. (8b) is the Moore diagram of \mathcal{D}_H .

$$Y_1 = \begin{pmatrix} H_1^1(z) \\ H_1^{-1}(z) \\ H_2^1(z) \\ H_2^{-1}(z) \\ H_3^1(z) \end{pmatrix}, Y_2 = \begin{pmatrix} H_3^{-1}(z) \\ H_4^1(z) \\ H_4^{-1}(z) \\ H_5^1(z) \\ H_5^{-1}(z) \end{pmatrix}, Z_1 = \begin{pmatrix} z^2 \\ z^2 \\ z \\ z \\ z \end{pmatrix} \text{ and } Z_2 = \begin{pmatrix} z \\ z^3 \\ z^3 \\ z^3 \\ z^3 \end{pmatrix}$$

we get

$$H(z) = \frac{(1+z)(1-4z+5z^2)}{(1-5z)(1-2z+5z^2)} \text{ and } \alpha_H = 5.$$

(b) $H = \langle a^3, ab, ab^{-1}, a^{-1}ba \rangle$ is a subgroup of F_2 . Solving the system

$$BY = Z,$$

where

$$B = \begin{pmatrix} 1 - z^3 & 0 & -z^3 & -z & -z^3 & -z & -z^3 & -z^3 \\ 0 & 1 - z^3 & -z^3 & -z^3 & -z^3 & -z^3 & -z & -z \\ -z^2 & -1 & 1 - z^2 & 0 & -z^2 & -1 & -z^2 & -z^2 \\ -z^2 & -z^2 & 0 & 1 - z^2 & -z^2 & -z^2 & -z^2 & -z^2 \\ -z^2 & -1 & -z^2 & -1 & 1 - z^2 & 0 & -z^2 & -z^2 \\ -z^2 & -z^2 & -z^2 & -z^2 & 0 & 1 - z^2 & -z^2 & -z^2 \\ -z & -z^3 & -z^3 & -z^3 & -z^3 & -z^3 & 1 - z & 0 \\ -z & -z^3 & -z^3 & -z^3 & -z^3 & -z^3 & 0 & 1 - z \end{pmatrix},$$

$$Y = \begin{pmatrix} H_1^1(z) \\ H_1^{-1}(z) \\ H_2^1(z) \\ H_2^{-1}(z) \\ H_3^1(z) \\ H_3^{-1}(z) \\ H_4^1(z) \\ H_4^{-1}(z) \end{pmatrix} \text{ and } Z = \begin{pmatrix} z^3 \\ z^3 \\ z^2 \\ z^2 \\ z^2 \\ z^2 \\ z^3 \\ z^3 \end{pmatrix}$$

we get

$$H(z) = \frac{(1+z)(1-2z+5z^2-6z^3+9z^4)}{(1-3z)(1-z+3z^2)(1+3z+3z^2)} \text{ and } \alpha_H = 3.$$

(2) Let H be an infinite index subgroup of F_m .

(a) $H = \langle aba^{-1}, aca^{-1} \rangle$ is a subgroup of F_3 . Solving the system

$$BY = Z,$$

where

$$B = \begin{pmatrix} 1 - z & 0 & -z & -z \\ 0 & 1 - z & -z & -z \\ -z & -z & 1 - z & 0 \\ -z & -z & 0 & 1 - z \end{pmatrix},$$

$$Y = \begin{pmatrix} H_1^1(z) \\ H_1^{-1}(z) \\ H_2^1(z) \\ H_2^{-1}(z) \end{pmatrix} \text{ and } Z = \begin{pmatrix} z^3 \\ z^3 \\ z^3 \\ z^3 \end{pmatrix}$$

we get

$$H(z) = 1 + \frac{4z^3}{1-3z} \text{ and } \alpha_H = 3.$$

(b) $H = \langle b^2, bab^{-1}a^{-1}, a^3 \rangle$ is a subgroup of F_2 . Solving the system

$$BY = Z,$$

where

$$B = \begin{pmatrix} 1 - z^2 & 0 & -z^2 & -1 & -z^2 & -z^2 \\ 0 & 1 - z^2 & -z^2 & -z^2 & -z^2 & -z^2 \\ -z^4 & -z^2 & 1 - z^4 & 0 & -z^4 & -z^4 \\ -z^4 & -z^4 & 0 & 1 - z^4 & -z^4 & -z^2 \\ -z^3 & -z^3 & -z & -z^3 & 1 - z^3 & 0 \\ -z^3 & -z^3 & -z^3 & -z^3 & 0 & 1 - z^3 \end{pmatrix},$$

$$Y = \begin{pmatrix} H_1^1(z) \\ H_1^{-1}(z) \\ H_2^1(z) \\ H_2^{-1}(z) \\ H_3^1(z) \\ H_3^{-1}(z) \end{pmatrix} \text{ and } Z = \begin{pmatrix} z^2 \\ z^2 \\ z^4 \\ z^4 \\ z^3 \\ z^3 \end{pmatrix}$$

we get

$$H(z) = \frac{1 + z + 3z^2 + 3z^3 + 5z^4 + 5z^5 + 6z^6 + 6z^7 + 4z^8 + 4z^9}{1 + z + z^2 - z^3 - 5z^4 - 13z^5 - 16z^6 - 20z^7 - 12z^8 - 12z^9}$$

and $\alpha_H = 1.88233$.

8. ACKNOWLEDGMENTS

RG was partially supported by Simons Foundation Collaboration Grant for Mathematicians, Award Number 527814. Also, RG acknowledges the support of the Max Planck Institute for Mathematics in Bonn and Humboldt Foundation. AS acknowledges United States-India Educational Foundation(USIEF), New Delhi and U.S. Department of State for the Fulbright Nehru Postdoctoral Research Fellowship, Award No.2479/FNPDR/2019. AS is grateful to S. P. Mandali, Pune and Government of Maharashtra for the sanction of study leave to undertake the fellowship. Also, AS acknowledges the generous hospitality of ICTS-TIFR, Bengaluru during the GGD 2017 program, where the initial idea of the project was discussed among the authors. The authors also would like to thank the anonymous referee for many suggestions that improved the exposition of the paper.

REFERENCES

- [1] Sergei I. Adyan. Random walks on free periodic groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 46(6):1139–1149, 1343, 1982.
- [2] Anatolij W. Anissimow and F. D. Seifert. Zur algebraischen charakteristik der durch kontext-freie sprachen definierten gruppen. *J. Inf. Process. Cybern.*, 11:695–702, 1975.
- [3] Jean Berstel and Christophe Reutenauer. *Noncommutative rational series with applications*, volume 137 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2011.
- [4] Tullio Ceccherini-Silberstein, Antonio Machi, and Fabio Scarabotti. On the entropy of regular languages. volume 307, pages 93–102. 2003. Words.
- [5] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

- [6] Scott Garrabrant and Igor Pak. Permutation patterns are hard to count. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 923–936. ACM, New York, 2016.
- [7] Stephen M. Gersten and Hamish B. Short. Rational subgroups of biautomatic groups. *Ann. of Math. (2)*, 134(1):125–158, 1991.
- [8] R. Grigorchuk and P. de la Harpe. On problems related to growth, entropy, and spectrum in group theory. *J. Dynam. Control Systems*, 3(1):51–89, 1997.
- [9] Rostislav I. Grigorchuk. *Banach Means on Homogeneous Spaces and Random Walks*. PhD Thesis, Lomonosov Moscow State University (in Russian), 1978.
- [10] Rostislav I. Grigorchuk. Invariant measures on homogeneous spaces. *Ukrain. Mat. Zh.*, 31(5):490–497, 618, 1979.
- [11] Rostislav I. Grigorchuk. Symmetrical random walks on discrete groups. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 285–325. Dekker, New York, 1980. English extended translation from the Russian edition of the book edited by Dobrušin, R. L. and Sinaĭ, Ja. G.
- [12] Rostislav I. Grigorchuk, Vadim A. Kaimanovich, and Tatiana Nagnibeda. Ergodic properties of boundary actions and the Nielsen-Schreier theory. *Adv. Math.*, 230(3):1340–1380, 2012.
- [13] John E. Hopcroft and Jeffrey D. Ullman. *Introduction to automata theory, languages, and computation*. Addison-Wesley Publishing Co., Reading, Mass., 1979. Addison-Wesley Series in Computer Science.
- [14] A. G. Howson. On the intersection of finitely generated free groups. *J. London Math. Soc.*, 29:428–434, 1954.
- [15] Ilya Kapovich and Alexei Myasnikov. Stallings foldings and subgroups of free groups. *J. Algebra*, 248(2):608–668, 2002.
- [16] Dmitri Kouksov. On rationality of the cogrowth series. *Proc. Amer. Math. Soc.*, 126(10):2845–2847, 1998.
- [17] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.
- [18] Roger C. Lyndon and Paul E. Schupp. *Combinatorial group theory*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1977 edition.
- [19] Wilhelm Magnus, Abraham Karrass, and Donald Solitar. *Combinatorial group theory: Presentations of groups in terms of generators and relations*. Courier Corporation, 2004.
- [20] Aleksandr Y. Ol’shanskii. On the problem of the existence of an invariant mean on a group. *Russian Mathematical Surveys*, 35(4):180, 1980.
- [21] Arto Salomaa and Matti Soittola. *Automata-theoretic aspects of formal power series*. Springer-Verlag, New York-Heidelberg, 1978. Texts and Monographs in Computer Science.
- [22] John R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.
- [23] Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.