UNIVERSITY OF CALIFORNIA

Los Angeles

Combinatorics of Finitely Generated Groups

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Mathematics

by

Anton Sergeevich Malyshev

2014

Abstract of the Dissertation

Combinatorics of Finitely Generated Groups

by

Anton Sergeevich Malyshev

Doctor of Philosophy in Mathematics

University of California, Los Angeles, 2014

Professor Igor Pak, Chair

We present some combinatorial results about finitely generated groups, particularly groups acting on rooted trees.

Given a finitely generated group G and a positive integer k, the product replacement graph $\Gamma_k(G)$ is the graph whose vertices are generating k-tuples of G, and whose edges are Nielsen transformations between these generating k-tuples. We prove that if G has polynomial growth or G has exponential growth, then $\Gamma_k(G)$ has exponential growth for sufficiently large k. We also prove with a direct combinatorial argument that $\Gamma_k(\mathbb{G}_\omega)$ has exponential growth for $k \geq 5$, where $\mathbb{G}_\omega$ is the generalized Grigorchuk group.

We prove that $\Gamma_k(G)$ is non-amenable for sufficiently large k in either of the following two cases: G is uniformly non-amenable, or G is virtually indicable. It follows that $\Gamma_k(G)$ is non-amenable whenever G is a linear group, or a hyperbolic group, or an elementary amenable group.

We describe two Mealy automata, the Aleshin automaton and the Bellaterra automaton, whose Schreier graphs are conjectured to be expanders. We verify that these Schreier graphs have polylogarithmic diameter. We describe a class of Mealy automata for which the same is true. However, some members of this class do not give rise to expander graphs.

The dissertation of Anton Sergeevich Malyshev is approved.

Rafail Ostrovsky

Ciprian Manolescu

Eli Gafni

Igor Pak, Committee Chair

University of California, Los Angeles

2014

Table of Contents

Acknowledgments

Chapter 2 is a version of

A. Malyshev and I. Pak, Growth in product replacement graphs, *J. Group Theory*, forthcoming.

Igor Pak posed the problem, and contributed the introduction and final remarks.

Chapter 3 is a version of

A. Malyshev, Non-amenabilty of product replacement graphs, *Comm. Algebra*, submitted.

Chapter 4 is a version of

A. Malyshev and I. Pak, Lifts, derandomization, and diameters of Schreier graphs of Mealy automata, *Combinatorica*, submitted.

Igor Pak posed the problem, and contributed the introduction.

Vita

2006	REU Participant, IPAM Research in Industrial Projects for Students.
2007	Counselor, Program in Mathematics for Young Scientists, Boston University.
2008	REU Participant, Trinity University.
2009	B.S. (Mathematics), Princeton University.
2009–2013	Teaching Assistant, Mathematics Department, UCLA.

Publications

A. Malyshev and I. Pak, Lifts, derandomization, and diameters of Schreier graphs of Mealy automata, *Combinatorica*, submitted.

A. Malyshev, Non-amenabilty of product replacement graphs, *Comm. Algebra*, submitted.

A. Malyshev and I. Pak, Growth in product replacement graphs, *J. Group Theory*, forthcoming.

R. Daileda, R. Krishnamoorthy and A. Malyshev, Maximal class numbers of CM number fields, *J. Number Theory* **130** (2010), 936–943.

CHAPTER 1

Introduction

We present some results in combinatorial group theory, relating to growth and expansion. There are two main problems we're concerned with. One is growth and expansion in the product replacement graph of a group. The other is constructing sequences of expander graphs via groups acting on rooted trees.

We are particularly concerned with self-similar groups, and more generally groups acting on rooted trees. This is partly because they are the only known way to construct groups of intermediate growth, which is the interesting case for some of the problems we're interested in. Additionally, the Schreier graphs of self-similar group actions on a rooted tree also have a natural description in terms of graph lifts. Random graph lifts give rise to expander graphs, and Schreier graphs of some self-similar groups are a good candidate for a derandomization of such a construction.

1.1 Product replacement graphs

Chapters 2 and 3 are concerned with product replacement graphs of finitely generated groups. The k-th *product replacement graph* of a finitely generated group G is the graph $\Gamma_k(G)$ whose vertices are k-tuples $(g_1, \ldots, g_k)$ such that $G = \langle g_1, \ldots, g_k \rangle$. The edges of this graph correspond to Nielsen transformations: for every pair of indices i, j, we have edges

$$(g_1, \ldots, g_i, \ldots, g_j, \ldots, g_k) — (g_1, \ldots, g_i, \ldots, g_i^{\pm 1} g_j, \ldots, g_k)$$
$$(g_1, \ldots, g_i, \ldots, g_j, \ldots, g_k) — (g_1, \ldots, g_i, \ldots, g_j g_i^{\pm 1}, \ldots, g_k).$$

A practical reason to be concerned with these graphs is that they give a fast way to sample random elements in a finite group given a generating set of that group. A naive way to make such a sample is to take a random walk on the Cayley graph of the group G, but the mixing time of this random walk may be large compared to $\log |G|$. A more efficient algorithm makes use of a random walk on $\Gamma_k(G)$, which can have a much shorter mixing time. For a survey on the product replacement algorithm, see [P1]. We restate some results here for motivation.

In most cases, the product replacement graphs of finite groups appear to be expander graphs. In the case that of abelian groups, this is known to be true, and follows from the fact that $SL_k(\mathbb{Z})$ has Kazhdan's property (T) for $k \geq 3$, and property (τ) for $k = 2$. For non-abelian groups, the corresponding question is whether $\mathrm{Aut}(F_k)$, the automorphism group of the free group F_k, has property (T). This is known to be false for $k = 2, 3$, and the question is open for $k \geq 4$.

We are interested in the case of infinite groups. If $\mathrm{Aut}(F_k)$ does have property (T), then for any infinite finitely generated group G, any component of $\Gamma_k(G)$ is non-amenable. That is, there is a constant h such that any finite set $S \subseteq \Gamma_k(G)$ has a boundary of size at least $h\,|S|$. In particular, $\Gamma_k(G)$ has exponential growth, i.e. balls grow exponentially with respect to their radius.

In Chapter 2, we consider which groups G have exponentially growing product replacement graphs. Group of polynomial growth have enough structure to guarantee that their product replacement graphs have exponential growth. On the other hand, groups of exponential growth also have product replacement graphs of exponential growth. The unresolved case, then, is that of groups with growth intermediate between polynomial and exponential. The first and best known example of such a group is the Grigorchuk group $\mathbb{G}$. This group is a self-similar group, defined via its action on a rooted binary tree. Using this action, we provide a direct counting argument that $\Gamma_k(\mathbb{G})$ also has exponential growth, for $k \geq 5$. The same argument applies to the generalized Grigorchuk groups $\mathbb{G}_\omega$. This technique is built around the action of the group $\mathbb{G}$ on a rooted tree. The only known way to construct groups of intermediate growth is via such an action. However there is little reason to expect that all groups of intermediate growth arise this way, so our methods cannot be applied in general.

We continue the investigation of product replacement graphs in Chapter 3, which is concerned with non-amenability of product replacement graphs. We establish similar results. If a group is "small enough" or "large enough", then its product replacement graphs are non-amenable. In this case "small enough" includes all elementary amenable groups, and "large enough" means the group is uniformly non-amenable. These two categories cover many well-known types of groups, e.g. linear groups and hyperbolic groups.

1.2 Expanding Mealy automata

In Chapter 4, we describe two group actions on a rooted binary tree, whose Schreier graphs we conjecture to be expanders. These graphs can be thought of as a derandomization of a construction of expanders via random 2-lifts [BL]. The n-th graph has size 2^n, so if they are expanders then their diameter grows linearly with respect to n. By a direct construction, we establish that their diameter grows at most quadratically.

Additionally, we describe a large class of similar actions on rooted trees whose Schreier graphs also have polynomially growing diameter. However, we demonstrate that some members of this class do not have expanding Schreier graphs.

References

[BL] Y. Bilu and N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* **26** (2006), 495–519.

[P1] I. Pak, What do we know about the product replacement algorithm?, in *Groups and Computation III*, de Gruyter, Berlin, 2001, 301–347.

CHAPTER 2

Growth in product replacement graphs of Grigorchuk groups

GROWTH IN PRODUCT REPLACEMENT GRAPHS OF GRIGORCHUK GROUPS

ANTON MALYSHEV⋆ AND IGOR PAK⋆

Abstract. The product replacement graph $\Gamma_k(G)$ is the graph on the generating k-tuples of a group G, with edges corresponding to Nielsen moves. We prove the exponential growth of product replacement graphs $\Gamma_k(\mathbb{G}_\omega)$ of Grigorchuk groups, for $k \geq 5$.

1. Introduction

The *product replacement graphs* $\Gamma_k(G)$ are the graphs on generating k-tuples of a group G, with edges corresponding to multiplications of one generator by another (see below). These graphs play an important role in computational group theory (see e.g. [BL, NP, P1]), and are related to the Andrews–Curtis conjecture in algebraic topology (see e.g. [BKM, BLM, Met]). For infinite groups, proving non-amenability of graphs $\Gamma_k(G)$ is a major open problem, closely related to Kazhdan's property (T) of $\mathrm{Aut}(F_k)$. In this paper we establish a weaker property, the exponential growth of product replacement graphs, for the Grigorchuk group $\mathbb{G}$ and its generalizations $\mathbb{G}_\omega$.

Let us begin by stating the main conjecture we address in this paper.

Conjecture 1.1 (Main Conjecture)**.** *Let G be an infinite group generated by d elements. Then the product replacement graphs $\Gamma_k(G)$ have exponential growth, for all $k \geq d+1$.*

Formally speaking, graphs $\Gamma_k(G)$ can be disconnected, in which case we conjecture that *at least one* connected component has exponential growth.

The motivation behind our Main Conjecture is rather interesting, which makes the conjecture both natural and speculative. First, recall that $\Gamma_k(G)$ are Schreier graphs of $\mathrm{Aut}(F_k)$, generated by Nielsen transformations [LP] (see also [LŻ, P1]). A well known conjecture states that $\mathrm{Aut}(F_k)$ has *Kazhdan's property (T)* for $k > 3$. If true, this would imply the following conjecture:

Conjecture 1.2. *For every infinite group G, product replacement graphs $\Gamma_k(G)$ are* non-amenable*, for k large enough.*

In particular, this conjecture implies that all connected components of $\Gamma_k(G)$ are infinite and have exponential growth, for all G and k large enough. We should mention that $\mathrm{Aut}(F_k)$ does not have (T) for $k = 2$ and 3 (see [GL, Lub]). On the other hand, the non-amenability of $\Gamma_n(G)$ follows from a weaker property (τ) for an appropriate family of subgroups (see [LŻ]).

We approach the Main Conjecture by looking at the growth of groups. The conjecture is straightforward for groups of exponential growth. It can be shown that the conjecture holds for virtually nilpotent groups (see Section 3). By Gromov's theorem, this implies the conjecture for all groups of polynomial growth.

Unfortunately, groups of intermediate growth lack the rigid structure of nilpotent groups, so much that even explicit examples are difficult to construct and analyze (see e.g. [dlH1, G3]). Even now, much remains open for the classical *Grigorchuk group* $\mathbb{G}$, the first example of a group of intermediate growth discovered by Grigorchuk (see [G1, G2]).

We present a new combinatorial technique which allows us to establish the conjecture for a large class of Grigorchuk groups $\mathbb{G}_\omega$. This is the main result of this paper:

⋆Department of Mathematics, UCLA, Los Angeles, CA, 90095. Email: {amalyshev,pak}@math.ucla.edu.

Theorem 1.3. *Let $\mathbb{G}_\omega$ be a generalized Grigorchuk group. Then $\Gamma_n(\mathbb{G}_\omega)$ is connected for each $n \geq 4$, and has exponential growth for each $n \geq 5$.*

The techniques in this paper generalize fairly easily to several other groups of intermediate growth, such as the Gupta–Sidki p-groups [GS], as well as large families of Grigorchuk p-groups. Many groups of intermediate growth, such as the groups of oscillating growth defined in [KP], have, by construction, some $\mathbb{G}_\omega$ as a subgroup or a factor group. Such groups, then, also have exponential Nielsen growth (see Proposition 3.7).

In fact, the techniques in this paper apply to a general class of branch groups defined in [Bar] called *splitter-mixer groups*. Many known group of intermediate growth appears to be based on a splitter-mixer group. (An example that does not fall into this class is given in [Nek] and analyzed in [BP], but our techniques should apply there as well). The proofs will appear in [M2].

In summary, although we have yet to find proofs in all cases, we believe the Main Conjecture holds for all *known* constructions of groups of intermediate growth. In that sense the situation is similar to the "$p_c < 1$" conjecture by Benjamini and Schramm [BS] for groups of superlinear growth. The conjecture is known to hold for groups of exponential and polynomial growth, and by an ad hoc argument for Grigorchuk groups and general self-similar groups [MP]. It remains open for general groups of intermediate growth (see [Pete]).

Let us mention that in a followup paper [M1], the first author establishes Conjecture 1.2 for several classes of groups of exponential growth, which include virtually solvable groups, linear groups, random finitely presented groups (in Gromov sense), and hyperbolic groups. He uses a technical extension of *uniform exponential growth* and *uniform non-amenability* (see [A+, BG, dlH2, Wil]).

Unfortunately, the explicit combinatorial approach in this paper, does not seem to be strong enough to establish Conjecture 1.2 for the Grigorchuk group, which we state as a separate conjecture of independent interest.

Conjecture 1.4. *Product replacement graphs $\Gamma_k(\mathbb{G})$ are non-amenable, for all $k \geq 5$.*

The rest of this paper is structured as follows. We begin with basic definitions of growth of groups and the product replacement graphs (Section 2). In Section 3 we present basic results on the growth and connectivity of graphs $\Gamma_k(G)$; we also present general tools for establishing the exponential growth results. In a technical Section 4 we describe general tools and techniques for working with subgroups $G \subset \mathrm{Aut}(\mathbf{T}_2)$ and their product replacement graphs. In the next two sections 5 and 6 we establish the main result. First, we prove the exponential growth of $\Gamma_k(\mathbb{G})$ for $k \geq 5$; in this case the (technical) argument is the most lucid. We then generalize this approach to *all* Grigorchuk groups $\mathbb{G}_\omega$. We conclude with final remarks and open problems (Section 7).

2. Background and definitions

2.1. **Notation.** Let X be a finite set. We write $\# X$ or $|X|$ to denote the size of X. Throughout the paper we use $\mathbb{Z}_n$ to denote the cyclic group $\mathbb{Z}/n\mathbb{Z}$.

Let Γ be a directed graph, which may have loops and repeated edges. We define $v \in \Gamma$ to mean that v is a vertex of Γ. Let v, w be vertices of Γ. We write $v \to w$ when there is an edge in Γ from v to w, and $v \rightsquigarrow w$ when there is a path in Γ from v to w. We say Γ is *symmetric* if for every edge $v \to w$ of Γ there is an *inverse edge* $w \to v$. Every graph considered in this paper is a symmetric directed graph, unless otherwise specified. When convenient, we think of a symmetric directed graph as an undirected graph by identifying every edge with its inverse.

Let G be a group, which may be finite or infinite. A *generating n-tuple* of G is an element $(g_1, \ldots, g_n) \in G^n$, such that $G = \langle g_1, \ldots, g_n \rangle$. Let $S = (g_1, \ldots, g_n)$ be such an n-tuple. Consider a left action of G on a set X. The *Schreier graph* $\mathrm{Schr}_S(G, X)$ of this action with respect to S, is the directed graph whose vertices are the elements of X, with edges $x \to g_i x$ and $x \to g_i^{-1} x$ for each $x \in X$, and each $0 \leq i \leq n$. Note that each vertex in $\mathrm{Schr}_S(G, X)$ has $2n$ edges leaving it, and each edge $v \to w$ in such a graph has an inverse edge $w \to v$. Thus, $\mathrm{Schr}_S(G, X)$ is a $2n$-regular symmetric directed graph.

The *Cayley graph* $\mathrm{Cay}_S(G)$ is the Schreier graph $\mathrm{Schr}_S(G, G)$ with respect to the left action of G on itself by multiplication. Clearly, the Cayley graph $\mathrm{Cay}_S(G)$ is connected. Given $g \in G$, we define $\ell_S(g)$ to be the length of the shortest path from 1 to g in the Cayley graph of G.

When the context makes it clear what the generating n-tuple S is, we drop the subscript, and simply write $\mathrm{Cay}(G)$, $\mathrm{Schr}(G, X)$, and $\ell(g)$. We write $\mathrm{Aut}(G)$ for the group of automorphisms of G. We write $H < G$ when H is a subgroup of G, and $H \lneqq G$ when H is a proper subgroup of G. For an element $g \in G$, denote by $\mathrm{ord}(g)$ the order of g. For $g_1, \ldots, g_n \in G$, denote

$$\overrightarrow{\prod_{i=1\ldots n}} g_i \, = \, g_1 \cdots g_n.$$

2.2. **Growth in graphs.** Let Γ be a symmetric directed graph, and let $v \in \Gamma$. The ball of radius r centered at v, denoted $B_\Gamma(v, r)$, is the set of vertices $w \in \Gamma$ such that there is a path of length at most r between v and w. For example, suppose $\Gamma = \mathrm{Cay}_S(G)$. Then $B_\Gamma(1, r)$ consists of the elements $g \in G$ for which $\ell_S(g) \le r$.

We say Γ has *exponential growth from* v, if there is a constant $\alpha > 1$, such that $|B_\Gamma(v, r)| \ge \alpha^r$ for all r (equivalently, for sufficiently large r). Suppose Γ has exponential growth from w, and there is a path $v \rightsquigarrow w$ in Γ. Then Γ also has exponential growth from v. Thus, if Γ is connected and has exponential growth from some $v \in \Gamma$, it also has exponential growth from any $w \in \Gamma$. In this case, we say that Γ has *exponential growth.*

2.3. **Growth in groups.** Let G be a group, Let S be a generating n-tuple of G. Define $B_{G,S}(r) = B_\Gamma(1, r)$, where $\Gamma = \mathrm{Cay}_S(G)$. When it is clear what S is, we simply write $B_G(r)$ instead. It is easy to verify that the following definitions are independent of the choice of generators S.

We say G has *exponential growth* if Γ has exponential growth. In other words, G has exponential growth if there is a constant $\alpha > 1$ such that $|B_G(r)| \ge \alpha^r$ for sufficiently large r. Equivalently G has exponential growth if and only if

$$\liminf_{r \to \infty} \frac{\log \bigl|B_G(r)\bigr|}{r} > 0.$$

Similarly, we say G has *polynomial growth* if there is a constant d with $|B_G(r)| \le r^d$ for sufficiently large r. In other words, G has polynomial growth if

$$\limsup_{r \to \infty} \frac{\log \bigl|B_G(r)\bigr|}{\log r} < \infty.$$

Example 2.1. The group $\mathbb{Z}$ has polynomial growth. With respect to the generating 1-tuple $S = (1)$, we have $B_\mathbb{Z}(r) = [-r, r]$, and hence $|B_\mathbb{Z}(r)| = 2r + 1$.

Example 2.2. The free group with two generators, $G = F_2 = \langle a, b\rangle$ has exponential growth. With respect to the generators $S = (a, b)$, we have $|B_G(r)| = 1 + 4 \cdot 3^{r-1}$ for $r \ge 1$.

We say G has *intermediate growth* if it has neither exponential nor polynomial growth. The first known example of a group of intermediate growth is the Grigorchuk group $\mathbb{G}$, which will be defined later, in Section 5. We refer to [dlH1, §VI] and [GP] for more on the growth of groups.

2.4. **Product replacement graphs.** Given a generating n-tuple of S a group G, we can take an element of S and multiply it, either on the left or the right, by another element or another element's inverse. Such an operation is called a *Nielsen move.* Formally, for each $1 \le i, j \le n$ with $i \ne j$, we define the Nielsen moves $R_{ij}^{\pm 1}$, $L_{ij}^{\pm 1}$ by

$$\begin{aligned} & R_{ij}^{\pm 1}(g_1, \ldots, g_i, \ldots, g_j, \ldots g_n) = (g_1, \ldots, g_i, \ldots, g_j g_i^{\pm 1}, \ldots, g_n), \\ \text{and} \quad & L_{ij}^{\pm 1}(g_1, \ldots, g_i, \ldots, g_j, \ldots g_n) = (g_1, \ldots, g_i, \ldots, g_i^{\pm 1} g_j, \ldots, g_n). \end{aligned}$$

Clearly, if S is a generating n-tuple of G, then $R_{ij}S$, $R_{ij}^{-1}S$, $L_{ij}S$, and $L_{ij}^{-1}S$ are also generating n-tuples of G.

We define the *product replacement graph* $\Gamma_n(G)$ to be the directed graph whose vertices are the generating n-tuples of G, where there is an edge from S to $R_{ij}S$, $R_{ij}^{-1}S$, $L_{ij}S$, and $L_{ij}^{-1}S$, for each generating n-tuple S and each pair of integers $i \neq j$ satisfying $1 \leq i, j \leq n$. This is a $4n(n-1)$-regular symmetric directed graph.

Observe that

$$\begin{aligned} R_{ij}L_{ji}^{-1}L_{ij}(g_1,\ldots,g_i,\ldots,g_j,\ldots,g_n) \,&=\, R_{ij}L_{ji}^{-1}(g_1,\ldots,g_i,\ldots,g_ig_j,\ldots,g_n) \\ =\, R_{ij}(g_1,\ldots,g_j^{-1},\ldots,g_ig_j,\ldots,g_n) \,&=\, (g_1,\ldots,g_j^{-1},\ldots,g_i,\ldots,g_n). \end{aligned}$$

Hence, a series of Nielsen moves can swap two elements in a generating n-tuple, inverting one of them. Doing this twice simply inverts both elements. This implies that Nielsen moves permit us to rearrange generators in an n-tuple, except that we may need to invert one element (see [P1]). Moreover, if $g_i = 1$ for some i, then we can use Nielsen moves invert any one element, and therefore we can rearrange the generators freely.

Example 2.3. The graph $\Gamma_2(\mathbb{Z})$ has a vertex for each pair of relatively prime integers (a, b), with two edges from (a, b) to each of $(a, b+a)$, $(a, b-a)$, $(a+b, b)$ and $(a-b, b)$. It is easy to check that this graph has exponential growth: the subgraph induced by $\big\{(a,b) \in \mathbb{Z}^2 \,\big|\, a, b > 0,\ \gcd(a,b) = 1\big\}$ is a rooted binary tree.

Example 2.4. Let $G = \mathbb{Z}_p^n$, with p prime. Then $\Gamma_n(G)$ is the set of bases of $\mathbb{Z}_p^n$ as a vector space over $\mathbb{Z}_p$. These bases are in one-to-one correspondence with matrices in $GL_n(\mathbb{Z}_p)$, and Nielsen moves correspond to elementary row operations. Row operations do not change the determinant of a matrix. It follows that there is one connected component for every value of the determinant. This implies that $\Gamma_n(\mathbb{Z}_p^n)$ has $p-1$ connected components (see [DG]).

2.5. **Growth of $\Gamma_n(G)$.** Let $S = (g_1, \ldots, g_n) \in \Gamma_n(G)$. We write

$$S^{(m)} := (g_1, \ldots, g_n, 1, \ldots, 1) \in \Gamma_{n+m}(G),$$

and define $\Gamma_{n+m}(G, S)$ to be the connected component of $\Gamma_{n+m}(G)$ containing $S^{(m)}$.

We say G has *exponential Nielsen growth* if $\Gamma_n(G, S)$ has exponential growth for some n and some generating n-tuple S of G. It is easy to show that a finitely generated group G has exponential Nielsen growth if G is either an infinite group of polynomial growth, or a group of exponential growth (see Proposition 3.10). This suggests that every infinite finitely generated group has exponential Nielsen growth:

Conjecture 2.5. *For every infinite finitely generated group G, there is an generating n-tuple $S \in \Gamma_n(G)$ such that $\Gamma_n(G, S)$ has exponential growth.*

Note that this conjecture is a weaker version of Conjecture 1.2. Here we accounted for the possibility that there can be many connected components, and are working with only one of them. Our Main Conjecture 1.1 is also stronger; implicit in it is a reference to a conjecture that every generating k-tuple is connected to a redundant generating k-tuple in $\Gamma_k(G)$. For this and stronger conjectures on connectivity of $\Gamma_k(G)$, see [P1] (see also [BKM]).

3. Basic results

3.1. **Growth of graphs.** We do not need to prove that $B_\Gamma(v, r)$ is large for every single r to conclude that Γ has exponential growth from v. As the following lemma shows, it suffices to prove it for a relatively sparse set of numbers r.

A sequence of positive integers $r_1, r_2, \ldots$ is called *log-dense* if it is increasing, and there is a constant β such that $r_{i+1} \leq \beta r_i$ for every $i \geq 1$. In other words, an increasing integer sequence (r_i) is log-dense if the gaps in the sequence $(\log r_i)$ are bounded above.

Lemma 3.1. *Let Γ be a symmetric directed graph, and let v be a vertex of Γ. Suppose that for some constant $\alpha > 1$, there is a log-dense sequence $r_1, r_2, \dots$ such that $|B(v, r_i)| \geq \alpha^{r_i}$ for every $i \geq 1$. Then Γ has exponential growth from v.*

Proof. Since r_i is an increasing sequence of positive integers, we can conclude that for sufficiently large r, there is an i with $r_i \leq r \leq r_{i+1}$. Since $r_{i+1} \leq \beta r_i$, we have $r_i \geq r/\beta$. Thus,

$$|B(v,r)| \geq |B(v,r_i)| \geq \alpha^{r_i} \geq \alpha^{r/\beta},$$

which implies the result. □

If a graph Γ is a covering of another graph Γ', and Γ' has exponential growth, then so does Γ.

Proposition 3.2. *Let Γ' and Γ be symmetric directed graphs, and suppose $\phi : \Gamma' \to \Gamma$ maps the set of neighbors of each vertex $v \in \Gamma'$ surjectively onto the neighbors of $\phi(v)$. Suppose Γ has exponential growth from $\phi(w)$. Then Γ' has exponential growth from w.*

Proof. It suffices to show that ϕ maps $B_{\Gamma'}(w, r)$ onto $B_\Gamma(\phi(w), r)$ for all $r \geq 0$, since in that case

$$|B_{\Gamma'}(w,r)| \geq |B_\Gamma(\phi(w),r)| .$$

We prove this by induction on r. The base case $r = 0$ is trivial. Suppose

$$\phi\big(B_{\Gamma'}(w,r)\big) \supseteq B_\Gamma\big(\phi(w),r\big),$$

and consider $v \in B_\Gamma(\phi(w), r+1)$. We know that v has a neighbor $u \in B_\Gamma(\phi(w), r)$, which has a preimage $u' \in B_{\Gamma'}(w, r)$. Since v is a neighbor of u, we know that some neighbor of u' is mapped to v. Therefore, $v \in \phi\big(B_{\Gamma'}(w, r+1)\big)$, as desired. □

It is easy to see that if a graph Γ is a subgraph of Γ', and Γ has exponential growth, so does Γ'. Moreover, we have the following stronger result:

Proposition 3.3. *Let Γ and Γ' be symmetric directed graphs, and suppose $\phi : \Gamma \to \Gamma'$ sends neighbors to neighbors. Suppose that there is a constant C such that $\#\, \phi^{-1}(v') \leq C$ for every vertex $v' \in \Gamma'$. Suppose that Γ has exponential growth from w. Then Γ' has exponential growth from $\phi(w)$.*

Proof. It suffices to show that ϕ maps $B_\Gamma(w, r)$ into $B_{\Gamma'}(\phi(w), r)$ for all $r \geq 0$, since in that case

$$|B_{\Gamma'}(\phi(w),r)| \geq |B_\Gamma(w,r)| /C.$$

We prove this by induction or r. The base case $r = 0$ is trivial. Suppose

$$\phi\big(B_\Gamma(w,r)\big) \subseteq B_{\Gamma'}\big(\phi(w),r\big),$$

and consider $v \in B_\Gamma(w, r+1)$. We know that v has a neighbor $u \in B_\Gamma(w, r)$, and $\phi(u) \in B_{\Gamma'}(\phi(w), r)$. Since u and v are neighbors, and ϕ sends neighbors to neighbors, we see that $\phi(v)$ is a neighbor of $\phi(u)$. It follows that $\phi(v) \in B_{\Gamma'}(\phi(w), r+1)$, as desired. □

3.2. **Growth of product replacement graphs.** Observe that if $m \geq n$ then $\Gamma_n(G, S)$ embeds into $\Gamma_m(G, S)$. Therefore, by Lemma 3.3 if $\Gamma_n(G, S)$ has exponential growth, so does $\Gamma_m(G, S)$.

Moreover, if H is a finitely generated subgroup of G, then every product replacement graph of H embeds in some product replacement graph of G. We can conclude that if a subgroup of G has a product replacement graph of exponential growth, so does G. Formally:

Proposition 3.4. *Let H and G be finitely generated groups with $H < G$. Suppose some connected component of $\Gamma_m(H)$ has exponential growth, and let $S \in \Gamma_n(G)$. Then $\Gamma_{n+m}(G, S)$ has exponential growth. In particular, if $H < G$ and H has exponential Nielsen growth, then G also has exponential Nielsen growth.*

Proof. Let $S = (g_1, \dots, g_n) \in \Gamma_n(G)$. We know that $\Gamma_m(H)$ has exponential growth from some $T \in \Gamma_m(H)$. Let $T = (h_1, \dots, h_m)$. There is a graph embedding $\phi : \Gamma_m(H) \to \Gamma_{n+m}(G)$ given by

$$\phi(h'_1, \dots, h'_m) = (g_1, \dots, g_n, h'_1, \dots, h'_m).$$

Hence, $\Gamma_{n+m}(G)$ has exponential growth from $\phi(T)$. Since the g_i's generate G, we know that each h_i is a product of g_i's and their inverses. Thus, there is a sequence of Nielsen moves $S^{(m)} \rightsquigarrow \phi(T)$, where

$$S^{(m)} = (g_1, \dots, g_n, 1, \dots, 1), \text{ and } \phi(T) = (g_1, \dots, g_n, h_1, \dots, h_m).$$

Therefore, $\Gamma_{n+m}(G, S) = \Gamma_{n+m}\big(G, \phi(T)\big)$, which implies that $\Gamma_{n+m}(G, S)$ has exponential growth. □

Similarly, we can show that if a group quotient of G has a product replacement graph of exponential growth, then so does G.

Proposition 3.5. *Let G and H be finitely generated groups, and let $f : G \to H$ be a surjective group homomorphism. Let $S \in \Gamma_n(G)$. Then the following hold.*

(1) *Suppose $\Gamma_n\big(H, f(S)\big)$ has exponential growth. Then $\Gamma_n(G, S)$ has exponential growth.*
(2) *Suppose some connected component of $\Gamma_m(H)$ has exponential growth. Then $\Gamma_{n+m}(G, S)$ has exponential growth.*
(3) *Suppose H has exponential Nielsen growth. Then G also has exponential Nielsen growth.*

Proof. For (1), we extend f to a map $\Gamma_n(G) \to \Gamma_n(H)$ by making the following definition.

$$f(g_1, \dots, g_n) = \big(f(g_1), \dots, f(g_h)\big).$$

This map f sends the neighbors of every $T \in \Gamma_n(G)$ surjectively onto the neighbors of $f(T)$. Thus, since $\Gamma_n(H)$ has exponential growth from $f(S)$, we can apply Proposition 3.2, and conclude that $\Gamma_n(G)$ has exponential growth from S.

For (2), let $S = (g_1, \dots, g_n) \in \Gamma_n(G)$, and choose

$$T = (h_1, \dots, h_m) = \big(f(\tilde{h}_1), \dots, f(\tilde{h}_m)\big) \in \Gamma_m(H)$$

such that $\Gamma_m(H, T)$ has exponential growth. Then

$$\Gamma_{n+m}\big(H, (f(g_1), \dots, f(g_n), h_1, \dots, h_m)\big)$$

also has exponential growth. Thus, by (1),

$$\Gamma_{n+m}\big(G, (g_1, \dots, g_n, \tilde{h}_1, \dots, \tilde{h}_m)\big)$$

has exponential growth. Since the g_i's generate G, we know that there is a path in $\Gamma_{n+m}(G)$

$$(g_1, \dots, g_n, \tilde{h}_1, \dots, \tilde{h}_m) \rightsquigarrow (g_1, \dots, g_n, 1, \dots, 1) = S^{(m)}.$$

Hence, $\Gamma_{n+m}(G)$ also has exponential growth from $S^{(m)}$, i.e. $\Gamma_{n+m}(G, S)$ has exponential growth. Finally, part (3) follows immediately from (2). □

In a different direction, if G has a product replacement graph of exponential growth, so does every quotient of H by a finite subgroup.

Proposition 3.6. *Let G and H be finitely generated groups, and let $f : G \to H$ be a surjective group homomorphism with finite kernel. For every $S \in \Gamma_n(G)$, if $\Gamma_n(G, S)$ has exponential growth, then $\Gamma_n\big(H, f(S)\big)$ has exponential growth. In particular, if G has exponential Nielsen growth, then H also has exponential Nielsen growth.*

Proof. We extend the map $f : G \to H$, to the map $f : \Gamma_n(G) \to \Gamma_n(H)$, given by

$$f(g_1, \dots, g_n) = \big(f(g_1), \dots, f(g_h)\big).$$

This map sends neighbors to neighbors, and the preimage of each vertex has bounded size. The graph $\Gamma_n(G)$ has exponential growth from S. Hence, by Proposition 3.3, $\Gamma_n(H)$ has exponential growth from $f(S)$. □

We summarize the previous three results in the following proposition.

Proposition 3.7. *Let G and G' be finitely generated groups, and suppose G is a subgroup, quotient, or extension by a finite group of G'. If G has exponential Nielsen growth, then G' also has exponential Nielsen growth.*

Remark 3.8. Proposition 3.7 relates the Nielsen growth of a subgroup H of G to the Nielsen growth of G. We conjecture that for any finite index subgroup H of G, if $\Gamma_n(G)$ has exponential growth, then so does $\Gamma_k(H)$ of G, for sufficiently large k. This would imply that the property of having exponential Nielsen growth respects virtual isomorphism. More generally, it would be interesting to see if this property is an invariant under quasi-isometry.

The proposition gives us an easy way to prove that a fairly large class of groups have exponential Nielsen growth.

Lemma 3.9. *Let G be a finitely generated group. Suppose G contains an element of infinite order. For every $S \in \Gamma_n(G)$ and every $m \geq n+2$, we have that $\Gamma_m(G,S)$ has exponential growth.*

Proof. By assumption, the group G contains a subgroup isomorphic to $\mathbb{Z}$. It is easy to see that $\Gamma_2(\mathbb{Z})$ has exponential growth (see Example 2.3). By Proposition 3.4, it follows that $\Gamma_{n+2}(G,S)$ has exponential growth, and hence so does $\Gamma_m(G,S)$ for every $m \geq n+2$. □

In particular, we can prove that groups of polynomial or exponential growth all have exponential Nielsen growth, which leaves Conjecture 2.5 open only for groups of intermediate growth.

Proposition 3.10. *Let G be an infinite finitely generated group. Suppose that either G has polynomial or exponential growth. Then G has exponential Nielsen growth.*

Proof. Suppose G has polynomial growth. By Gromov's theorem, G is virtually nilpotent [Gro]. It follows that some subgroup of G has infinite abelianization. Thus, G has an element of infinite order and, by Lemma 3.9, G has exponential Nielsen growth.

Now suppose G has exponential growth. Let $S = (g_1, \ldots, g_n)$ be a generating n-tuple of G and denote $\Gamma = \Gamma_{n+1}(G,S)$. Let r be any positive integer. For any $g \in B_{G,S}(r)$, the distance between $S^{(1)} = (g_1, \ldots, g_n, 1)$ and $(g_1, \ldots, g_n, g)$ in Γ is at most r, i.e. $(g_1, \ldots, g_n, g) \in B_\Gamma(S,r)$. Thus,

$$|B_\Gamma(S,r)| \geq |B_{G,S}(r)|\,.$$

But $|B_{G,S}(r)|$ grows exponentially in r, and thus so does $|B_\Gamma(S,r)|$. That is, $\Gamma = \Gamma_{n+1}(G,S)$ has exponential growth, and therefore G has exponential Nielsen growth. □

Remark 3.11. The Grigorchuk group $\mathbb{G}$ does not have an element of infinite order, so Lemma 3.9 is not enough to show that its product replacement graphs have exponential growth. It can be shown that $\Gamma_n(G)$ has exponential growth for sufficiently large n as long as there are elements of G whose order is exponential in their word length (see [M2]). The Grigorchuk group $\mathbb{G}$ does not satisfy this condition either, but some of the generalized Grigochuk groups $\mathbb{G}_\omega$ do.

3.3. **Effective results.** The Grigorchuk group has no elements of infinite order, so Lemma 3.9 is not strong enough to prove it has exponential Nielsen growth. We use a different approach. It is enough to find large cubes in G, as follows.

Let G be any group, and let $(g_1, \ldots, g_k) \in G^k$, we say the *cube spanned by* $(g_1, \ldots, g_k)$ is

$$\mathcal{C}(g_1, \ldots, g_n) := \left\{ g_1^{\varepsilon_1} \cdots g_n^{\varepsilon_n} \,\middle|\, \varepsilon_i \in \{0,1\} \right\}.$$

Observe that $\#\,\mathcal{C}(g_1, \ldots, g_n) \leq 2^n$. We say $(g_1, \ldots, g_k)$ is a *cubic* k-tuple if

$$\#\,\mathcal{C}(g_1, \ldots, g_k) = 2^k.$$

Lemma 3.12. *Let G be a finitely generated group, and fix $S \in \Gamma_n(G)$. Let $\alpha > 1$ be a constant, and (k_i) be a log-dense sequence. Suppose for each $i \geq 1$, there is a path γ of length at most αk_i in $\Gamma_n(G)$, such that γ starts at S and visits some $S_1, \ldots, S_{k_i} \in \Gamma_n(G)$ in that order. Suppose further that there is a cubic k_i-tuple $(g_1, \ldots, g_{k_i})$, where $g_j \in S_j$ for each $1 \leq j \leq k_i$. Then $\Gamma_m(G,S)$ has exponential growth for every $m \geq n+1$.*

Proof. It is enough to show that $\Gamma_{n+1}(G,S)$ has exponential growth. Let $\Gamma = \Gamma_{n+1}(G)$, and $k = k_i$. By Lemma 3.1, it suffices to show that

$$\left|B_\Gamma(S^{(1)}, (\alpha+1)k)\right| \geq 2^k.$$

Given $(\varepsilon_1, \ldots, \varepsilon_k) \in \{0,1\}^k$, we traverse the path γ in the first n coordinates of $\Gamma_{n+1}(G)$, but when we reach S_j, if $\varepsilon_j = 1$ we also apply a Nielsen transformation to multiply the last entry by g_j. This gives us a path γ' in $\Gamma_{n+1}(G)$ of length at most $\alpha k + k$. The path γ' ends at an element of $\Gamma_{n+1}(G)$ whose last entry is $g_1^{\varepsilon_1} \ldots g_n^{\varepsilon_n}$. Since $(g_1, \ldots, g_k)$ is cubic, there are 2^k distinct such elements. Thus, we have constructed 2^k distinct elements of $B_\Gamma(S^{(1)}, \alpha k + k)$, as desired. □

3.4. **Connectivity of product replacement graphs.** Recall the *Frattini subgroup* $\Phi(G)$,

$$\Phi(G) = \left\{ g \in G \,\middle|\, \text{if } H \lneqq G, \text{ then } \langle H, g\rangle \lneqq G \right\}.$$

(see e.g. [Hall, §10.4]). It is easy to see that $\Phi(G)$ is a normal subgroup of G. We need the following connectivity result by Evans (see [Eva, Theorem 4.3]).

Theorem 3.13 (Evans). *Suppose G is generated by some n-tuple. Let $m \geq n+1$, and suppose $\Gamma_m\big(G/\Phi(G)\big)$ is connected. Then $\Gamma_m(G)$ is connected.*

It is known that for any finite abelian group G with n generators, the product replacement graph $\Gamma_m(G)$ is connected for every $m > n$ [DG] (see also [P1]). We use only the following special case, which is easy to verify by hand.

Lemma 3.14. *The product replacement graph $\Gamma_m(\mathbb{Z}_2^n)$ is connected for every $m \geq n$.*

In particular, suppose $G/\Phi(G) \cong \mathbb{Z}_2^n$. Then $\Gamma_m(G)$ is connected for every $m > n$.

Remark 3.15. Theorem 3.13 is an analogue for infinite groups of the following result in [LP] (see also [P1]). Let G and H be finite groups with k generators, and $f : G \to H$ is a surjective group homomorphism, then the extension $f : \Gamma_k(G) \to \Gamma_k(H)$ is surjective. That is, every generating k-tuple of H lifts to a generating k-tuple of G. As a corollary, if $\Gamma_k(G)$ is connected, then so is $\Gamma_k(H)$. This claim is not true for infinite groups.

4. Automorphisms of the rooted binary tree

In this section, we introduce and discuss properties of the group $\mathrm{Aut}(\mathbf{T})$ of automorphisms of a binary tree.

4.1. **Definitions.** Let $\mathbf{T} = \{0,1\}^*$ denote the rooted binary tree consisting of finite strings over the alphabet $\{0,1\}$, whose root is the empty string, where the children of the string s are $s0$ and $s1$. Define $\mathrm{Aut}(\mathbf{T})$ to the group of automorphisms of this tree. Formally, $\mathrm{Aut}(\mathbf{T})$ consists of length preserving bijections g of $\mathbf{T}$ such that for any $s, t \in \mathbf{T}$, $g(st)$ begins with $g(s)$. To avoid confusion with the bit 1, we let $\mathbf{i} \in \mathrm{Aut}(\mathbf{T})$ denote the identity element. Let $g{\downarrow_s}$ denote the action of g on tails of strings beginning with s. In other words, we define it to satisfy $g(st) = g(s)g{\downarrow_s}(t)$.

Define $a \in \mathrm{Aut}(\mathbf{T})$ to be the automorphism which flips the first bit of s. Formally, $a(0s) = 1s$ and $a(1s) = 0s$ for all $s \in \mathbf{T}$. Clearly, every element of $\mathrm{Aut}(\mathbf{T})$ either fixes 0 and 1 or swaps them. Let g be an element that fixes them. Then $g(0s) = 0g{\downarrow_0}(s)$ and $g(1s) = 1g{\downarrow_1}(s)$. In this case, we write g in one of the following two forms, depending on which is more convenient:

$$g = (g{\downarrow_0}, g{\downarrow_1}) \quad \text{or} \quad g = \begin{pmatrix} g{\downarrow_0} \\ g{\downarrow_1} \end{pmatrix}.$$

On the other hand, suppose $g \in \mathrm{Aut}(\mathbf{T})$ swaps 0 and 1. Then $g = a(g{\downarrow_0}, g{\downarrow_1}) = (g{\downarrow_1}, g{\downarrow_0})a$. Thus, we can write every element $g \in \mathrm{Aut}(\mathbf{T})$ in the form $(h,k)a^\varepsilon$, for some $h, k \in \mathrm{Aut}(\mathbf{T})$ and some $\varepsilon \in \{0,1\}$. Moreover $(h,k)a = a(k,h)$ for all $h, k \in \mathrm{Aut}(\mathbf{T})$. In other words, we have $\mathrm{Aut}(\mathbf{T}) = (\mathrm{Aut}(\mathbf{T}) \times \mathrm{Aut}(\mathbf{T})) \rtimes \mathbb{Z}_2$ where $\mathbb{Z}_2$ acts on $\mathrm{Aut}(\mathbf{T}) \times \mathrm{Aut}(\mathbf{T})$ by swapping the two coordinates.

Let $s \in \mathbf{T}$ be a given binary string. We say the *stabilizer* of s is the subgroup of $\mathrm{Aut}(\mathbf{T})$ consisting of those elements $g \in \mathrm{Aut}(\mathbf{T})$ which fix s:

$$\mathrm{Stab}(s) := \left\{ g \in \mathrm{Aut}(\mathbf{T}) \ \text{ s.t. } g(s) = s \right\}.$$

The n-th *level stabilizer* is the subgroup of $\mathrm{Aut}(\mathbf{T})$ consisting of those elements which fix the n-th level of $\mathbf{T}$:

$$\mathrm{Stab}_n := \bigcap_{s \in \{0,1\}^n} \mathrm{Stab}(s).$$

Let $g \in \mathrm{Stab}_n$. The *n-support* of g is

$$\mathrm{supp}_n(g) = \left\{ s \in \{0,1\}^n \ \text{ s.t. } g{\downarrow_s} \neq \mathbf{i} \right\}.$$

Finally, given $s \in \{0,1\}^n$, we define the *rigid stabilizer* of s to be the subgroup

$$\mathrm{Rist}(s) := \left\{ g \in \mathrm{Stab}_n \ \text{ s.t. } \mathrm{supp}_n(g) \subseteq \{s\} \right\}.$$

In other words, $\mathrm{Rist}(s)$ consists of those elements of $\mathrm{Aut}(\mathbf{T})$ which fix every string that does not begin with s.

For a subgroup G of $\mathrm{Aut}(\mathbf{T})$, define

$$\mathrm{Stab}_G(s) = G \cap \mathrm{Stab}(s) \quad \text{and} \quad \mathrm{Rist}_G(s) = G \cap \mathrm{Rist}(s).$$

Note that

$$\begin{aligned} \mathrm{Rist}(0s) &= \{(g, \mathbf{i}) \mid g \in \mathrm{Rist}(s)\} = \mathrm{Rist}(s) \times \{\mathbf{i}\}, \\ \text{and} \quad \mathrm{Rist}(1s) &= \{(\mathbf{i}, g) \mid g \in \mathrm{Rist}(s)\} = \{\mathbf{i}\} \times \mathrm{Rist}(s). \end{aligned}$$

4.2. **Growth in subgroups of** $\mathrm{Aut}(\mathbf{T})$**.** For distinct $s, s' \in \{0,1\}^m$, elements of $\mathrm{Rist}(s)$ and $\mathrm{Rist}(s')$ have disjoint n-support. We use Nielsen transformations to reach many of these elements. This implies we can find a large cubic set, which lets us construct many different generating n-tuples.

Lemma 4.1. *Let $G < \mathrm{Aut}(\mathbf{T})$ be finitely generated, and fix a generating n-tuple $S \in \Gamma_n(G)$. Suppose G acts transitively on every level of $\mathbf{T}$. Suppose there is a constant α such that for every $m \geq 1$, there is a string $s \in \{0,1\}^m$ and a nontrivial element $g \in \mathrm{Rist}_G(s)$ with $\ell(g) \leq \alpha 2^m$. Then $\Gamma_k(G, S)$ has exponential growth for every $k \geq n + 2$*

Proof. Given m, define $L = \{0,1\}^m$ and $N = 2^m$. Fix $s \in L$ such that there is a nontrivial $g \in \mathrm{Rist}_G(s)$, satisfying $\ell(g) \leq \alpha N$. Since G acts transitively on L, we have that the Schreier graph $\mathrm{Schr}_S(G, L)$ is connected. Therefore, $\mathrm{Schr}_S(G, L)$ has a spanning tree $\mathcal{T}$. Consider a depth-first traversal of $\mathcal{T}$ with respect to the lexicographic order on L, starting at s. This is a path of length $2\,|L| - 2 < 2N$ which visits every element of L. Suppose it visits them in the order $s_1, \ldots, s_N$. For each $1 \leq i \leq N$, define h_i to be the group element corresponding to the walk along this path from s to s_i, so that $(s_1, \ldots, s_N) = (h_1(s), \ldots, h_N(s))$. Then we have $\ell(h_2 h_1^{-1}) + \cdots + \ell(h_N h_{N-1}^{-1}) \leq 2N = 2^{m+1}$, and $\bigl(h_1(s), \ldots, h_N(s)\bigr)$ is a permutation of the elements of L.

Since $g \in \mathrm{Rist}_G(s)$, we have $h_i g h_i^{-1} \in \mathrm{Rist}_G\left(h_i(s)\right)$, for all $1 \leq i \leq N$. We claim that

$$(h_1 g h_1^{-1}, \ldots, h_N g h_N^{-1})$$

is a cubic N-tuple, i.e.

$$\#\left\{ \overrightarrow{\prod_{i=1\ldots N}} (h_i g h_i)^{\varepsilon_i}, \ \text{ where } \ \varepsilon_i \in \{0,1\} \right\} = 2^N.$$

Indeed, the surjection $\phi : \{0,1\}^N \to \mathcal{C}(h_1 g h_1^{-1}, \ldots h_N g h_N^{-1})$ given by

$$\phi(\varepsilon) := \overrightarrow{\prod_{i=1\ldots N}} (h_i g h_i^{-1})^{\varepsilon_i}$$

is also injective, since $\varepsilon_i = 1$ if and only if $s_i \in \mathrm{supp}_n \phi(\varepsilon)$. Hence $\#\,\mathcal{C}(h_1 g h_1^{-1}, \ldots h_N g h_N^{-1}) = 2^N$, as desired.

Since $\ell(g) \leq \alpha 2^m$, there is a path γ_1 in $\Gamma_{n+1}(G,S)$ of length at most $\alpha 2^m$

$$S^{(1)} = (g_1, \ldots, g_n, 1) \rightsquigarrow (g_1, \ldots, g_n, g) = (g_1, \ldots, g_n, h_1 g h_1^{-1}).$$

Observe that the distance in $\Gamma_{n+1}(G,S)$ between $(g_1, \ldots, g_n, h_i g h_i^{-1})$ and $(g_1, \ldots, g_n, h_{i+1} g h_{i+1}^{-1})$ is at most $2\ell(h_{i+1} h_i^{-1})$. Since $\ell(h_2 h_1^{-1}) + \cdots + \ell(h_N h_{N-1}^{-1}) \leq 2^{m+1}$, there is a path γ_2 in Γ_{n+1} of length at most 2^{m+2} which starts at $(g_1, \ldots, g_n, g)$ and visits each $(g_1, \ldots, g_n, h_i g h_i^{-1})$, in that order.

Composing γ_1 and γ_2, we see that there is a path in $\Gamma_{n+1}(G,S)$ of length at most $(\alpha+4)2^m$ which starts at $S^{(1)}$ and visits generating $(n+1)$-tuples containing $h_1 g h_1^{-1}, \ldots, h_N g h_N^{-1}$, in that order. These elements of G form a cubic 2^m-tuple. Applying Lemma 3.12 with $k_m = 2^m$, then, tells us that $\Gamma_k(G,S)$ has exponential growth for all $k \geq n+2$. □

Remark 4.2. We cannot replace $\ell(g) \leq \alpha 2^n$ in the hypotheses of this lemma with $\ell(g) \leq \alpha^n$ with some $\alpha > 2$. Roughly speaking, that would only let us reach a cubic 2^n-tuple in α^n steps. Thus, we can only generate an $r^{1/d}$-cube in $B_\Gamma(S,r)$, where $d = \log_2 \alpha$, which is not sufficient to guarantee exponential growth. We can, however, replace the assumption that $\ell(g) \leq \alpha 2^n$ with the assumption that we can reach a generating $(n+1)$-tuple containing g in $\alpha 2^n$ Nielsen moves.

5. The Grigorchuk group

5.1. **Definition.** The Grigorchuk group $\mathbb{G} < \mathrm{Aut}(\mathbf{T})$ is defined as $\mathbb{G} = \langle a, b, c, d \rangle$, where a flips the first bit of a string, and b, c, and d are defined recursively by the relations

$$\begin{aligned} b &:= (a, c) \\ c &:= (a, d) \\ d &:= (\mathbf{i}, b). \end{aligned}$$

It is easy to check that $a^2 = b^2 = c^2 = d^2 = bcd = \mathbf{i}$. Thus, $\mathbb{G}$ is actually generated by just three elements: $\mathbb{G} = \langle a, b, c \rangle$.

Here is an explicit description of the action of these involutions on $\mathbf{T}$.

$$\begin{aligned} d(1^n) &= 1^n \\ d(1^n 0 s) &= \begin{cases} 1^n 0 s, & n \equiv 0 \pmod 3 \\ 1^n 0 a(s), & n \equiv 1, 2 \pmod 3 \end{cases} \end{aligned}$$

In other words, d changes at most one bit in a string – the bit after the first 0. Specifically, d flips that bit if and only if the number n of 1's in the string up to that point is 1 or 2 (mod 3). Similarly, c flips it when $n \equiv 0, 2 \pmod 3$, and b flips it when $n \equiv 0, 1 \pmod 3$.

Theorem 5.1 (Gigorchuk)**.** *The group $\mathbb{G}$ has intermediate growth.*

The theorem was first proved by Grigorchuk in [G1] (see also [GP, dlH1]).

5.2. **Connectivity of $\Gamma_n(\mathbb{G})$.** We prove the following result:

Proposition 5.2. *For each $n \geq 4$, the product replacement graph $\Gamma_n(\mathbb{G})$ is connected (see also §7.1).*[1]

Proof. Fix $n \geq 4$. It is known that $\mathbb{G}/\Phi(\mathbb{G}) \cong \mathbb{Z}_2^3$ (see [Per] and [G2, §6]). The graph $\Gamma_n(\mathbb{Z}_2^3)$ is connected by Lemma 3.14. Thus, by Lemma 3.13, $\Gamma_n(\mathbb{G})$ is connected. □

[1] After this paper was written, we learned that the proposition was independently derived in [Myr].

5.3. Exponential growth in $\Gamma_n(\mathbb{G})$.

The goal of this section is to prove the following result:

Theorem 5.3. *For each $n \geq 5$, the product replacement graph $\Gamma_n(\mathbb{G})$ of the Grigorchuk group has exponential growth.*

The proof is based on Lemma 4.1. Roughly, our strategy is to find an element g of $\mathrm{Rist}_{\mathbb{G}}(1^n)$ with length $O(2^n)$. In $O(2^n)$ more steps, we conjugate g to reach an element of $\mathrm{Rist}_{\mathbb{G}}(s)$ for each s on the same level of $\mathbf{T}$. Then we can construct every product of these conjugates in $O(2^n)$ steps. There are 2^{2^n} such products, which gives us exponential growth.

Proof of Theorem 5.3. Fix $n \geq 5$. It is easy to check that $\mathbb{G}$ acts transitively on the levels of $\mathbf{T}$ (see e.g. [dlH1, §VIII] or Lemma 6.1, below). By Lemma 4.1, it suffices to show that for every $m \geq 0$, there is a nontrivial element of $\mathrm{Rist}(1^m)$ of length at most 2^{m+4} with respect to the generating 3-tuple (a, b, c).

Define $t_0 = abab$. Observe that $t_0^2(111) = 110$, and therefore $t_0^2 \neq \mathbf{i}$. We prove by induction on m that there is a $t_m \in \mathbb{G}$ of the form

$$(*) \qquad t_m = \overrightarrow{\prod_{i=1\ldots N}} abax_i,$$

where $N = 2^m$, $x_i \in \{b, c, d\}$ for each $1 \leq i \leq 2^m$, such that $t_m^2 \in \mathrm{Rist}_{\mathbb{G}}(1^m)$ and $t_m{\downarrow_{1^m}} = t_0$. The base case $m = 0$ is trivial.

Given t_m and (x_i) related by $(*)$, for each $0 \leq i \leq N$ we define $x_i' \in \{b, c, d\}$ by $x_i' = (a^{\varepsilon_i}, x_i)$ where $\varepsilon_i \in \{0, 1\}$. We define t_{m+1} by applying the rewriting rules $a \mapsto aba$, $b \mapsto d$, $c \mapsto b$, $d \mapsto c$ to t_m. Then we have

$$t_{m+1} = \left[\overrightarrow{\prod_{i=1\ldots N}} (aba)d(aba)x_i'\right] = \left[\overrightarrow{\prod_{i=1\ldots N}} \begin{pmatrix} c \\ a \end{pmatrix}\begin{pmatrix} \mathbf{i} \\ b \end{pmatrix}\begin{pmatrix} c \\ a \end{pmatrix}\begin{pmatrix} a^{\varepsilon_i} \\ x_i \end{pmatrix}\right] = \begin{pmatrix} a^{\varepsilon} \\ t_m \end{pmatrix},$$

Thus,

$$t_{m+1}^2 = (\mathbf{i}, t_m^2) \in \{\mathbf{i}\} \times \mathrm{Rist}(1^m) = \mathrm{Rist}(1^{m+1}),$$

and $t_{m+1}^2{\downarrow_{1^{m+1}}} = t_m^2{\downarrow_{1^m}} = t_0^2$.

Since $t_m^2{\downarrow_{1^m}} = t_0^2 \neq \mathbf{i}$, we can conclude that $t_m^2 \neq \mathbf{i}$. Hence, for every $m \geq 0$, we have that t_m^2 is a nontrivial element of $\mathrm{Rist}_{\mathbb{G}}(1^m)$, with $\ell_{\langle a,b,c\rangle}(t_m^2) \leq 2\ell_{\langle a,b,c,d\rangle}(t_m^2) \leq 2^{m+4}$, which concludes the proof. □

6. The generalized Grigorchuk groups

In this section, we use the same approach to analyze growth in the product replacement graph of $\mathbb{G}_\omega$. The same techniques apply, but the technical details are more involved.

6.1. Definition.

Let ω be an infinite string in the alphabet[2] $\{b, c, d\}$. The *generalized Grigorchuk group* $\mathbb{G}_\omega$ is the group of automorphisms of $\{0, 1\}^n$ given by $\mathbb{G}_\omega = \langle a, b_0, c_0, d_0\rangle$. Here, the element a flips the first digit of a string, and for each $x \in \{b, c, d\}$, the elements x_n are defined recursively by

$$x_n := (a^{\varepsilon}, x_{n+1}), \qquad \text{where } \varepsilon = \begin{cases} 0, & x = \omega_n \\ 1, & \text{otherwise.} \end{cases}$$

For convenience, we write $b = b_0$, $c = c_0$, and $d = d_0$. As with $\mathbb{G}$, we have $a^2 = b^2 = c^2 = d^2 = bcd = 1$.

[2] The usual definition uses the alphabet $\{0, 1, 2\}$ but for our purposes it is more convenient to use $\{b, c, d\}$.

As before, we give a more explicit description of the action of $\mathbb{G}_\omega$ on $\mathbf{T}$. Given $x \in \{b, c, d\}$ and $s \in \mathbf{T}$,

$$x(1^n) = 1^n, \text{ and}$$

$$x(1^n 0 s) = \begin{cases} 1^n 0 s, & \omega_n = x \\ 1^n 0 a(s), & \text{otherwise.} \end{cases}$$

Taking $\omega = dcbdcbdcbdcb\ldots$ gives the usual Grigorchuk group. The following fact is well-known, but we include a proof here for completeness.

Lemma 6.1. *The generalized Grigorchuk group $\mathbb{G}_\omega$ acts transitively on every level of $\mathbf{T}$.*

Proof. We prove that $\mathbb{G}_\omega$ acts transitively on the n-th level by induction on n. This is trivial for $n = 0$, and true for $n = 1$ because $a \in \mathbb{G}_\omega$. For $n > 1$, note that it suffices to show that for each $s \in \{0,1\}^n$, there is a $g \in \mathbb{G}_\omega$ such that $g(s) = 1^{n-2}00$. Consider $s \in \{0,1\}^n$. We know that $s = s'd$, for some $s' \in \{0,1\}^{n-1}$ and $d \in \{0,1\}$. By the induction hypothesis, $\mathbb{G}_\omega$ acts transitively on $\{0,1\}^{n-1}$. Thus there is a $g \in \mathbb{G}_\omega$ with $g(s') = 1^{n-2}0$. Then either $g(s) = 1^{n-2}00$ or $g(s) = 1^{n-2}01$. In the latter case, there is an $x \in \{b, c, d\}$ such that $\omega_{n-2} \neq x$, and then $x(g(s)) = 1^{n-2}00$. In both cases, there is an $h \in \mathbb{G}_\omega$ with $h(s) = 1^{n-2}00$. □

6.2. **Exponential growth in $\Gamma_n(\mathbb{G}_\omega)$.** To prove Theorem 1.3, we first need some lemmas about $\mathbb{G}_\omega$. A standard computation shows that, under some weak assumptions on ω, every element of $\mathbb{G}_\omega$ has finite order. We will use the following more specialized result.

Lemma 6.2. *Suppose $\omega_{n-1} = d$. Then in $\mathbb{G}_\omega$, we have $(ad_k)^{2^{n-k+1}} = \mathbf{i}$ for every $0 \le k < n$.*

Proof. Since $\omega_{n-1} = d$, we have $d_{n-1} = (\mathbf{i}, d_n)$ and $ad_{n-1}a = (d_n, \mathbf{i})$. We prove the lemma by induction on $j = n - k$. When $j = 1$, i.e. $k = n - 1$, we have

$$(ad_k)^4 = [(ad_{n-1}a)d_{n-1}]^2 = \left[\begin{pmatrix} d_n \\ \mathbf{i} \end{pmatrix}\begin{pmatrix} \mathbf{i} \\ d_n \end{pmatrix}\right]^2 = \begin{pmatrix} d_n^2 \\ d_n^2 \end{pmatrix} = \mathbf{i}.$$

When $j > 1$, i.e. $k < n - 1$, the induction hypothesis tells us $(ad_{k+1})^{2^j} = \mathbf{i}$. Note that also $(d_{k+1})^{2^j} = \mathbf{i}$, since d_{k+1} has order 2. Then, for some $\varepsilon \in \{0,1\}$, we have

$$(ad_k)^{2^{j+1}} = [(ad_k a)d_k]^{2^j} = \left[\begin{pmatrix} d_{k+1} \\ a^\varepsilon \end{pmatrix}\begin{pmatrix} a^\varepsilon \\ d_{k+1} \end{pmatrix}\right]^{2^j} = \begin{pmatrix} (a^\varepsilon d_{k+1})^{-2^j} \\ (a^\varepsilon d_{k+1})^{2^j} \end{pmatrix} = \mathbf{i}.$$

□

Lemma 6.3. *Suppose $\omega \in \{b, c, d\}^*$ is not eventually constant. Then for each $n \ge 0$, there is a nontrivial $t \in \mathrm{Rist}_{\mathbb{G}_\omega}(1^n)$ with $\ell(t) \le 2^{n+2}$.*

Proof. This is trivial if $n = 0$. If $n > 1$, then by relabeling b, c, and d if necessary, we may assume $\omega_{n-1} = d$.

By induction on $j = n - k$ we show that for every $0 \le k \le n$, there is a t_k of the form

$$t_k = \overrightarrow{\prod_{i=1\ldots 2^{n-k}}} a x_i,$$

where $x_i \in \{b_k, d_k\}$ for each i, and there is an odd number of i's with $x_i = d_k$, such that $t_k^2 \in \mathrm{Rist}(1^{n-k})$, and $t_k^2 \neq \mathbf{i}$.

For $j = 0$, i.e. $k = n$, we define $t_n = ad_n$. We know that $d_n = (a^\varepsilon, d_{n+1})$ for some $\varepsilon \in \{0,1\}$, and therefore we have

$$t_n^2 = (ad_n a)d_n = \begin{pmatrix} d_{n+1} \\ a^\varepsilon \end{pmatrix}\begin{pmatrix} a^\varepsilon \\ d_{n+1} \end{pmatrix} = \begin{pmatrix} d_{n+1}a^\varepsilon \\ (d_{n+1}a^\varepsilon)^{-1} \end{pmatrix}.$$

Since ω is not eventually constant, we know that there is an $m \geq n+1$ with $\omega_m \neq d$. Therefore we have

$$d_{n+1}(1^{m-n-1}00) = 1^{m-n-1}d_m(00) = 1^{m-n-1}0a(0) = 1^{m-n-1}01.$$

Hence, $d_{n+1} \neq \mathbf{i}$. It follows that $d_{n+1}a^{\varepsilon}$ is nontrivial whether $\varepsilon = 0$ or $\varepsilon = 1$. Therefore, $t_n^2 \neq \mathbf{i}$.

For $j = 1$, i.e. $k = n-1$, we define $t_k = ab_{n-1}ad_{n-1}$. We have $\omega_k = d$, hence $d_k = (\mathbf{i}, d_{k+1})$ and $b_k = (a, b_{k+1})$. Therefore, we have:

$$t_{n-1}^2 = [(ab_{n-1}a)d_{n-1}]^2 = \left[\begin{pmatrix} b_n \\ a \end{pmatrix}\begin{pmatrix} \mathbf{i} \\ d_n \end{pmatrix}\right]^2 = \begin{pmatrix} \mathbf{i} \\ (ad_n)^2 \end{pmatrix} = \begin{pmatrix} \mathbf{i} \\ t_n^2 \end{pmatrix}.$$

For $j > 1$, i.e. $k < n-1$, let $N = 2^{n-k-1}$. We have

$$t_{k+1} = \overrightarrow{\prod_{i=1\ldots N}} ax_i$$

from the previous step. For each $1 \leq i \leq N$, we know that $x_i = b_{k+1}$ or d_{k+1}, and we define

$$x_i' = \begin{cases} b_k, & x_i = b_{k+1} \\ d_k, & x_i = d_{k+1}. \end{cases}$$

Then $x_i' = (a^{\varepsilon_i}, x_i)$ for some $\varepsilon_i \in \{0, 1\}$. We have three possibilities:

Case (i): $\omega_k = b$. Then $d_k = (a, d_{k+1})$ and $b_k = (\mathbf{i}, b_{k+1})$. Define

$$t_k = \overrightarrow{\prod_{i=1\ldots N}} ad_kax_i'.$$

The product has an even number of terms, thus, we have not changed the parity of the number of d_k's in the product, which implies it is still odd.

$$t_k = \overrightarrow{\prod_{i=1\ldots N}} (ad_ka)x_i' = \overrightarrow{\prod_{i=1\ldots N}} \begin{pmatrix} d_{k+1} \\ a \end{pmatrix}\begin{pmatrix} a^{\varepsilon_i} \\ x_i \end{pmatrix} = \begin{pmatrix} \overrightarrow{\prod} d_{k+1}a^{\varepsilon_i} \\ t_{k+1} \end{pmatrix},$$

where the final product runs over $i = 1\ldots N$. Observe that $\varepsilon_i = 1$ if and only if $x_i = d_i$. There are an odd number of such i, therefore $\overrightarrow{\prod} d_{k+1}a^{\varepsilon_i}$ is a product containing an odd number of a's and an even number of d_{k+1}'s. The elements d_{k+1} and a have order 2, so the group $\langle d_{k+1}, a\rangle$ is a dihedral group in which they are both reflections. Hence, the product $\overrightarrow{\prod} d_{k+1}a^{\varepsilon_i}$ is also a reflection in that dihedral group, and thus it has order 2. Therefore, $t_k^2 = (\mathbf{i}, t_{k+1}^2)$.

Case (ii): $\omega_k = d$. Then $d_k = (\mathbf{i}, d_{k+1})$ and $b_k = (a, b_{k+1})$. Define

$$t_k = \overrightarrow{\prod_{i=1\ldots N}} ab_kax_i',$$

and argue as in case (i).

Case (iii): $\omega_k = c$. Then $d_k = (a, d_{k+1})$ and $b_k = (a, b_{k+1})$. Hence $x_i' = (a, x_i)$ for each $0 \leq i \leq N$. We can again define

$$t_k = \overrightarrow{\prod_{i=1\ldots N}} ad_kax_i'.$$

Then

$$t_k = \overrightarrow{\prod_{i=1\ldots N}} (ad_ka)x_i' = \overrightarrow{\prod_{i=1\ldots N}} \begin{pmatrix} d_{k+1} \\ a \end{pmatrix}\begin{pmatrix} a \\ x_i \end{pmatrix} = \begin{pmatrix} (d_{k+1}a)^N \\ t_{k+1} \end{pmatrix}.$$

By Lemma 6.2, we have $(d_{k+1}a)^{2N} = (d_{k+1}a)^{2^{n-k}} = \mathbf{i}$. Hence, $t_k^2 = (\mathbf{i}, t_{k+1}^2)$.

In all three cases, $t_k^2 = (\mathbf{i}, t_{k+1}^2)$. It follows that t_k^2 is nontrivial and

$$t_k^2 \in \{\mathbf{i}\} \times \operatorname{Rist}(1^{n-k-1}) = \operatorname{Rist}(1^{n-k}).$$

Thus, we have a nontrivial $t_0^2 \in \operatorname{Rist}(1^n) \cap \mathbb{G}_\omega$, with $\ell(t_0^2) \le 2^{n+2}$, as desired. □

Proof of Theorem 1.3. It is known that $\mathbb{G}_\omega/\Phi(\mathbb{G}_\omega) \cong \mathbb{Z}_2^k$ for some $k \le 3$ [Per] (see also [G2, §6]). Recall from Lemma 3.14 that $\Gamma_n(\mathbb{Z}_2^k)$ is connected. Lemma 3.13 tells us that $\Gamma_n(\mathbb{G}_\omega)$ is connected for each $n \ge 4$.

Assume that ω is eventually constant. Then it is not hard to check that $\mathbb{G}_\omega$ has polynomial growth. In fact, G_ω is virtually abelian [G2, §2]. It follows that G_ω has an element of infinite order. The group $\mathbb{G}_\omega$ is generated by three elements, $\mathbb{G}_\omega = \langle a, b, c\rangle$. By Lemma 3.9, this implies that the product replacement graph $\Gamma_n(\mathbb{G}_\omega)$ has exponential growth for each $n \ge 5$.

Otherwise, if ω is not eventually constant, for every $m \ge 0$, Lemma 6.3 gives a nontrival $t \in \operatorname{Rist}_{\mathbb{G}_\omega}(1^m)$ of length at most 2^{m+2}. Since $\mathbb{G}_\omega$ acts transitively on the levels of $\mathbf{T}$, we can apply Lemma 4.1 to conclude that $\Gamma_6(\mathbb{G}_\omega)$ has exponential growth from $(a, b, c, d, 1, 1)$.

Moreover, note that the group $\mathbb{G}_\omega$ is generated by (a, b, c), and rewriting t as a word in these generators at most doubles its length. Thus, we also have that $\Gamma_5(\mathbb{G}_\omega)$ has exponential growth from $(a, b, c, 1, 1)$. It follows that $\Gamma_n(\mathbb{G}_\omega)$ has exponential growth for each $n \ge 5$. □

7. Final remarks

7.1. There are several other directions in which our Theorem 5.3 can be extended. First, there is the problem of smaller k: we believe that that $\Gamma_3(\mathbb{G})$ is connected (cf. Lemma 3.14 and Proposition 5.2).[3] Moreover, it is conceivable that both $\Gamma_3(\mathbb{G})$ and $\Gamma_4(\mathbb{G})$ have exponential growth, the cases missing from Theorem 5.3.

Similarly, in case Conjecture 1.4 proves too difficult, there is a weaker and perhaps more accessible open problem.

Conjecture 7.1. *The nearest neighbor random walk on $\Gamma_k(\mathbb{G})$ has positive speed, for all $k \ge 5$.*

The speed of r.w. is defined as the limit of $\mathbb{E}[\operatorname{dist}(t)/t]$ as $t \to \infty$, where $\operatorname{dist}(t)$ is the distance of the r.w. after t steps, from the starting vertex. It is known that non-amenable graphs have positive speed, but so do some amenable graphs, such as the standard Cayley graph of the lamplighter group $\mathbb{Z}_2 \wr \mathbb{Z}^3$ (see e.g. [Pete, Woe]). We believe it might be possible to extend our approach to establish the positive speed of r.w. on $\Gamma_k(\mathbb{G})$, and we intend to return to this problem.

7.2. The connectivity of product replacement graphs is delicate already for finite groups. For example, Dunwoody showed in [Dun], that if G is a finite solvable group with d generators, then $\Gamma_k(G)$ is connected, for every $k > d$ (see also [P1]). This property is conjectured to hold for all finite groups, but fails for infinite groups, even for metabelian groups (see [P1] and references therein).

As of now, is unknown whether for any finitely generated group G, graphs $\Gamma_k(G)$ are connected for all sufficiently large k. It is not even known that if $\Gamma_k(G)$ is connected then $\Gamma_{k+1}(G)$ is connected. The difficulty arises from the possibility that $\Gamma_{k+1}(G)$ has a connected component which consists of non-redundant generating $(k+1)$-tuples. However, it is not hard to check that in $\Gamma_{2d}(G)$ every element of the form $(g_1, \ldots, g_d, 1, \ldots, 1)$ lies in the same connected component, which we may call $\Gamma^\star_{2d}(G)$. Then if we know that some connected component of $\Gamma_d(G)$ has exponential growth, we know that $\Gamma^\star_{2d}(G)$ has exponential growth.

[3] See also Corollary 1.2 and Question 1 in [Myr].

7.3. Finally, let us mention that the notion of exponential Nielsen growth may be applicable to sequences of finite groups, which stabilize in a certain sense. Proving such a result would be a step towards proving expansion of product replacement graphs of general finite groups (see [P1, P2]). We refer to [Bla] for the notion of growth of finite groups, and to [Ell] for a recent conceptual approach.

Acknowledgements. The authors are grateful to Tatiana Nagnibeda who brought to our attention a question on connectivity and exponential growth of $\Gamma_k(\mathbb{G})$, and help with the references. We are also very thankful to Slava Grigorchuk and Martin Kassabov for interesting conversations on groups of intermediate growth, and to Yehuda Shalom for helpful remarks on uniform growth. The second author was partially supported by the BSF and the NSF.

References

[Adi] S. Adian, *The Burnside problem and identities in groups*, Springer, Berlin, 1979.

[A+] G. N. Arzhantseva, J. Burillo, M. Lustig, L. Reeves, H. Short and E. Ventura, Uniform non-amenability, *Adv. Math.* **197** (2005), 499–522

[Bar] L. Bartholdi, Groups of intermediate growth, `arXiv:0201293`.

[BL] H. Bäärnhielm and C. R. Leedham-Green, The product replacement prospector, *J. Symbolic Comput.* **47** (2012), 64–75.

[BS] I. Benjamini and O. Schramm, Percolation beyond $\mathbb{Z}^d$, many questions and a few answers, *Electron. Comm. Probab.* **1** (1996), 71–82.

[Bla] S. Black, Asymptotic growth of finite groups, *J. Algebra* **209** (1998), 402–426.

[BKM] A. V. Borovik, E. I. Khukhro and A. G. Myasnikov, The Andrews-Curtis Conjecture and Black Box Groups, *Int. J. Algebra Comput.* **13** (2003), 415–436.

[BLM] A. V. Borovik, A. Lubotzky and A. G. Myasnikov, The finitary Andrews-Curtis conjecture, in *Progr. Math.* **248**, Birkhäuser, Basel, 2005, 15–30.

[BG] E. Breuillard and T. Gelander, Uniform independence in linear groups, *Invent. Math.* **173** (2008), 225–263.

[BP] K. Bux and R. Pérez On the growth of iterated monodromy groups, *Contemp. Math.* **394** (2006), 61–76.

[dlH1] P. de la Harpe, *Topics in Geometric Group Theory*, University of Chicago Press, Chicago, 2000.

[dlH2] P. de la Harpe, Uniform growth in groups of exponential growth, *Geom. Dedicata* **95** (2002), 1–17.

[DG] P. Diaconis and R. Graham, The graph of generating sets of an abelian group, *Colloq. Math* **80** (1999), 31–38.

[Dun] M. J. Dunwoody, Nielsen transformations, in *Computational Problems in Abstract Algebra*, Pergamon Press, Oxford, 1970, 45–46.

[Ell] J. S. Ellenberg, Superstrong approximation for monodromy groups, `arXiv:1210.3757`.

[Eva] M. J. Evans, Presentations of groups involving more generators than are necessary, *Proc. LMS* **67** (1993), 106–126.

[G1] R. I. Grigorchuk, Degrees of growth of finitely generated groups, and the theory of invariant means, *Izv. Akad. Nauk SSSR Ser. Mat.* **48** (1984), 939–985.

[G2] R. I. Grigorchuk, Solved and unsolved problems around one group, in *Infinite Groups: Geometric, Combinatorial and Dynamical Aspects*, Birkhäuser, Basel, 2005, 117–218.

[G3] R. I. Grigorchuk, Some problems of the dynamics of group actions on rooted trees, *Proc. Steklov Inst. Math.* **273** (2011), 64-175.

[GP] R. I. Grigorchuk and I. Pak, Groups of intermediate growth, an introduction, *L'Ens. Math.* **54** (2008), 251–272.

[GS] R. I. Grigorchuk and S. N. Sidki, The group of automorphisms of a 3-generated 2-group of intermediate growth, *Internat. J. Algebra Comput.* **14** (2004), 667–676.

[Gro] M. Gromov, Groups of polynomial growth and expanding maps, *IHES Publ. Math.* **53** (1981), 53–78.

[GL] F. Grunewald and A. Lubotzky, Linear representations of the automorphism group of a free group, *Geom. Funct. Anal.* **18** (2009), 1564–1608.

[GS] N. Gupta and S. Sidki, On the Burnside problem for periodic groups, *Math. Z.* **182** (1983), 385–388.

[Hall] M. Hall, *The Theory of Groups*, Chelsea, New York, 1976.

[KP] M. Kassabov and I. Pak, Groups of oscillating intermediate growth, *Ann. Math.* **177** (2013), 1113–1145.

[Lub] A. Lubotzky, *Discrete Groups, Expanding Graphs, and Invariant Measures*, Birkhäuser, Basel, 1994.

[LP] A. Lubotzky and I. Pak, The product replacement algorithm and Kazhdan's property (T), *J. AMS* **14** (2001), 347–363.

[LŻ] A. Lubotzky and A. Żuk, *On property* (τ), monograph in preparation.

[M1] A. Malyshev, Non-amenability of product replacement graphs, in preparation.

[M2] A. Malyshev, *Combinatorics of finitely generated groups*, Ph.D. thesis, UCLA, in preparation.

[Met] W. Metzler, On the Andrews-Curtis conjecture and related problems, in *Contemp. Math.* **44**, AMS, Providence, RI, 1985, 35–50.

[Mil] J. Milnor, Growth of finitely generated solvable groups, *J. Diff. Geom.* **2** (1968), 447–449.

[MP] R. Muchnik and I. Pak, Percolation on Grigorchuk groups, *Comm. Algebra* **29** (2001), 661–671.

[Myr] A. Myropolska, Andrews–Curtis and Nielsen equivalence relations on some infinite groups, `arXiv:`1304.2668.

[Nek] V. Nekrashevych, Iterated monodromy groups, *London Math. Soc. Lecture Note Ser.* **387** (2011), 41–93.

[NP] A. C. Niemeyer and C. E. Praeger, Complexity and computation in matrix groups, in *Aspects of Complexity*, de Gruyter, Berlin, 2001, 87–113.

[P1] I. Pak, What do we know about the product replacement algorithm?, in *Groups and Computation III*, de Gruyter, Berlin, 2001, 301–347.

[P2] I. Pak, The product replacement algorithm is polynomial, in *Proc. FOCS 2000*, IEEE Comput. Soc. Press, Los Alamitos, CA, 2000, 476–485.

[Per] E. L. Pervova, Everywhere dense subgroups of one group of tree automorphisms, *Tr. Mat. Inst. Stekl.* **231** (2000), 356–367.

[Pete] G. Pete, *Probability and Geometry on Groups*, Lecture notes for a graduate course, 2013, 203 pp.; available at `http://www.math.bme.hu/~gabor/PGG.pdf`

[Wil] J. S. Wilson, On exponential growth and uniformly exponential growth for groups, *Invent. Math.* **155** (2004), 287–303.

[Woe] W. Woess, *Random walks on infinite graphs and groups*, Cambridge U. Press, Cambridge, 2000.

CHAPTER 3

Non-amenabilty of product replacement graphs

NON-AMENABILTY OF PRODUCT REPLACEMENT GRAPHS

ANTON MALYSHEV*

Abstract. We prove non-amenability of the product replacement graphs $\Gamma_n(G)$ for uniformly non-amenable groups. We also prove it for virtually indicable groups, when n is sufficiently large. It follows that $\Gamma_n(G)$ is non-amenable when n is sufficiently large for hyperbolic groups, linear groups, and elementary amenable groups.

1. Introduction

The *product replacement graph* of a group G is the graph of generating n-tuples of G, connected by Nielsen moves. These graphs play an important role in computational group theory (see e.g. [BL]), and have been considered in connection with the Andrews-Curtis conjecture (see e.g. [BLM, Met]). Relatively little is known about these graphs. Even their connectivity is a major open problem (see e.g. [Eva, Myr]). In this paper, we continue the investigation in [MP].

The main subject of this paper is the non-amenability of product replacement graphs. This is related to the well-known problem of whether the automorphism group $\mathrm{Aut}(F_n)$ of the free group has Kazhdan property (T) for $n > 3$, (see [LZ, LP], see also Subsection 7.2). If so, product replacement graphs of infinite groups would be *non-amenable*. This motivates the following conjecture.

Conjecture 1.1. [MP] *The product replacement graph $\Gamma_n(G)$ of an infinite finitely generated group G is non-amenable for sufficiently large n.*

The purpose of this note is to prove that Conjecture 1.1 holds for several classes of groups. We prove and use the following two theorems.

Theorem 1.2. *If G is uniformly non-amenable, then the product replacement graph $\Gamma_n(G)$ is non-amenable for every $n \geq d(G)$.*

We say G is *virtually indicable* if G has a finite index subgroup which has $\mathbb{Z}$ as a quotient.

Theorem 1.3. *If G is virtually indicable, then the product replacement graph $\Gamma_n(G)$ is non-amenable for sufficiently large n.*

We combine these theorems with the results on uniform non-amenability in [A+] to show that several classes of infinite groups have non-amenable product replacement graphs. In particular, hyperbolic groups, linear groups, elementary amenable groups, and free Burnside groups all satisfy Conjecture 1.1. Note that elementary amenable groups include virtually solvable groups and virtually amenable groups.

The paper is structured as follows. In Section 2, we define our terms and recall some basic facts about non-amenability. In Sections 4 and 3, we prove our main theorems. In Section 5 we discuss the consequences of these theorems to several general classes of groups. In Section 6, we prove the lemmas we used in the previous sections. Finally, in Section 7 we discuss the relationship of this problem to unsolved problems, and indicate further directions.

[redacted]

2. Definitions

The *product replacement graph* $\Gamma_n(G)$ of a finitely generated group G is the undirected graph whose vertices are n-tuples $S = (s_1, \ldots, s_n) \in G^n$ with $G = \langle s_1, \ldots, s_n \rangle$, and edges

$$(s_1, \ldots, s_i, \ldots, s_j, \ldots s_n) \leftrightarrow (s_1, \ldots, s_i, \ldots, s_j s_i^{\pm 1}, \ldots, s_n)$$
$$(s_1, \ldots, s_i, \ldots, s_j, \ldots s_n) \leftrightarrow (s_1, \ldots, s_i, \ldots, s_i^{\pm 1} s_j, \ldots, s_n)$$

for each pair $1 \leq i, j \leq n$, with $i \neq j$. We call a step from a vertex in this graph to one of its neighbors a *Nielsen move*. We denote the minimum number of generators of G with $d(G)$, so $\Gamma_n(G)$ is nonempty only for $n \geq d(G)$.

Here and elsewhere, by a *graph* $\Gamma = (V, E)$ we mean a possibly infinite undirected graph of bounded degree, which may have loops and repeated edges. By an abuse of notation, we often write Γ to mean the vertex set of Γ. The *distance* $d(v, w)$ between two vertices $v, w \in \Gamma$ is defined to be the length of the shortest path connecting them (or ∞ if there is no such path). We define $d(v, X)$ to be the distance from a vertex $v \in \Gamma$ to a set of vertices $X \subset \Gamma$, i.e. $d(v, X) = \min\{d(v, x) \mid x \in X\}$. For any two sets of vertices $X, Y \subset \Gamma$, we write $E(X, Y)$ for the set of edges between X and Y.

Let G be a finitely generated group, and fix a generating tuple $S = (s_1, \ldots, s_n) \in \Gamma_n(G)$. The *Cayley graph* $\mathrm{Cay}(G, S) = (V, E)$ is the graph with vertex set $V = G$, and edges $g \leftrightarrow g s_i^{\pm 1}$ for each $1 \leq i \leq n$.

We define the *Cheeger constant* of a nonempty graph $\Gamma = (V, E)$ by

$$h(\Gamma) := \inf_X \frac{|\partial X|}{|X|},$$

where the infimum runs over all finite sets of vertices $X \subset \Gamma$, and $\partial X = E(X, \overline{X})$ denotes the set of edges leaving X. We say Γ is *non-amenable*[1] if $h(\Gamma) > 0$. When $\Gamma = \mathrm{Cay}(G, S)$, we denote the Cheeger constant $h(\mathrm{Cay}(G, S))$ with $h(G, S)$. It is easy to check that the property $h(G, S) > 0$ depends only on the group G and not on the choice of generators S. We say a finitely generated group G is *non-amenable* if $h(G, S) > 0$, i.e. if $\mathrm{Cay}(G, S)$ is non-amenable. However, the Cheeger constant $h(G, S)$ itself may depend on S. We say G is *uniformly non-amenable*[2] if for every $n \geq d(G)$, there is a constant $h(G, n) > 0$ such that

$$\inf_{S \in \Gamma_n(G)} h(G, S) > h(G, n).$$

We say a map $f : \Gamma' \to \Gamma$ between two graphs is a *Lipschitz* map if there is a constant $C > 0$ such that for every pair of neighbors $v, w \in \Gamma'$, we have $d(f(v), f(w)) \leq C$. We say a subset $W \subset \Gamma$ is *dense* if there is a constant $D > 0$ such that for every $v \in \Gamma$ we have $d(v, W) \leq D$.

3. Uniformly non-amenable groups

A natural special case of Conjecture 1.1 is the case where G is assumed to be a non-amenable group. One obstacle to proving this is that the Cheeger constant of the Cayley graph may be arbitrarily small, depending on the generating set of G (see e.g. [A+, O2]). Theorem 1.2 asserts that if this is not the case, i.e. if the group is uniformly non-amenable, then $\Gamma_n(G)$ is non-amenable.

3.1. **Lemmas.**

To prove this theorem, we need the following two lemmas. The first lemma is a variation on the fact that *quasi-isometry* of graphs preserves non-amenability. For more on quasi-isometry, see e.g. [Woe, §3,4], [dlH, §IV.B].

Lemma 3.1. *Let Γ and Γ' be infinite graphs, with Γ non-amenable. Let $f : \Gamma \to \Gamma'$ be an injective Lipschitz map such that $f(\Gamma)$ is dense in Γ'. Then Γ' is also non-amenable.*

[1] Many authors require a non-amenable graph to be connected, but we omit this requirement. Our definition still forbids finite connected components in Γ.

[2] This is a weaker definition than usually appears in the literature. See remark 7.7.

The second lemma is related to the following standard fact: The class of amenable groups is closed under the operation of group extensions. That is, if G is non-amenable, then for every normal subgroup H of G, either H or G/H is non-amenable. In our argument, we only consider the case where H is the center of G, but we need an explicit lower bound on the Cheeger constant of G/H:

Lemma 3.2. *If G is a non-amenable group, then $G/Z(G)$ is also non-amenable. Moreover, for every $S = (s_1, \dots, s_n) \in \Gamma_n(G)$, we have*

$$h(G/Z(G), \widetilde{S}) \geq h(G, S),$$

where $\widetilde{S} = (s_1 Z(G), \dots, s_n Z(G)) \in \Gamma_n(G/Z(G))$.

We prove these lemmas in Section 6.

3.2. **Proof of Theorem 1.2.**

Let G be a uniformly non-amenable group. Given any $S = (s_1, \dots, s_n) \in \Gamma_n(G)$, define a map $f_S : G \to \Gamma_n(G)$ by

$$f_S(g) = gSg^{-1} = (gs_1g^{-1}, \dots, gs_ng^{-1})$$

Observe that

$$\begin{aligned} f_S(gs_1) &= \Big(gs_1s_1s_1^{-1}g^{-1},\, gs_1s_2s_1^{-1}g^{-1},\, \dots\, ,\, gs_1s_ns_1^{-1}g^{-1}\Big) \\ &= \Big(gs_1g^{-1},\, (gs_1g^{-1})(gs_2g^{-1})(gs_1g^{-1})^{-1},\, \dots\, ,\, (gs_1g^{-1})(gs_ng^{-1})(gs_1g^{-1})^{-1}\Big) \end{aligned}$$

is within $2n-2$ Nielsen moves of $f_S(g)$. The same is true of $f_S(gs_i)$ for every $1 \leq i \leq n$, so f_S is a Lipschitz map from $\mathrm{Cay}(G, S)$ to $\Gamma_n(G)$, with Lipschitz constant $2n-2$.

Let $\widetilde{G} = G/Z(G)$, where $Z(G)$ denotes the center of G. For each $g \in G$, denote by $\widetilde{g}$ the projection of g into $G/Z(G)$. Let $\widetilde{S}$ denote the corresponding generating n-tuple of $\widetilde{G}$, i.e. $\widetilde{S} = (\widetilde{s}_1, \dots, \widetilde{s}_n)$. Observe that $f_S(g) = f_S(h)$ if and only if $gs_ig^{-1} = hs_ih^{-1}$ for every $1 \leq i \leq n$. This occurs precisely when $g^{-1}h$ commutes with every s_i, i.e. when $hg^{-1} \in Z(G)$. Thus we have a well-defined and injective induced map $\widetilde{f}_S : \widetilde{G} \to \Gamma_n(G)$ given by $\widetilde{f}_S(\widetilde{g}) = f_S(g)$. We also have that $\widetilde{f}_S$ is a Lipschitz map from $\mathrm{Cay}(\widetilde{G}, \widetilde{S})$ to $\Gamma_n(G)$, with Lipschitz constant $2n-2$.

Now let S vary. Given $S' \in \Gamma_n(G)$, note that $S' \in \operatorname{image} \widetilde{f}_S$ if and only if $S' = gSg^{-1}$ for some $g \in G$. This is an equivalence relation, so the images of the maps $\widetilde{f}_S$ form a partition of $\Gamma_n(G)$ into equivalence classes. Let $\mathcal{S}$ be a set of representatives of these equivalence classes, and consider the disjoint union of graphs

$$\Delta := \coprod_{S \in \mathcal{S}} \mathrm{Cay}(\widetilde{G}, \widetilde{S}).$$

Using Lemma 3.2 and the fact that G is uniformly non-amenable, we can conclude that the graph Δ is non-amenable, since

$$h(\Delta) \geq \inf_{S \in \mathcal{S}} h(\widetilde{G}, \widetilde{S}) \geq \inf_{S \in \mathcal{S}} h(G, S) > 0$$

The maps $\widetilde{f}_S$ for $S \in \mathcal{S}$ combine into one map $\widetilde{f} : \Delta \to \Gamma_n(G)$. This map is injective, surjective, and Lipschitz. Therefore, by Lemma 3.1, $\Gamma_n(G)$ is also non-amenable. □

4. Virtually indicable groups

It was shown in [MP] that a group containing an element of infinite order must have exponentially growing product replacement graphs $\Gamma_n(G)$, for sufficiently large n. In order to guarantee non-amenability, we require a stronger property: we say a group G is *virtually indicable* if it contains a finite index subgroup which has $\mathbb{Z}$ as a quotient. Theorem 1.3 asserts that virtually indicable groups have non-amenable product replacement graphs $\Gamma_n(G)$, for sufficiently large n.

Note that in this case there may exist $n \geq d(G)$ for which $\Gamma_n(G)$ fails to be non-amenable. For example, the infinite dihedral group D_∞ is virtually indicable, but $\Gamma_2(D_\infty)$ is an amenable infinite graph. However, $\Gamma_n(D_\infty)$ is non-amenable for every $n \geq 3$.

4.1. **Lemmas.**

To prove Theorem 1.3, we again use Lemma 3.1, as well as two additional lemmas. If H is a quotient of G, then $\Gamma_n(G)$ is a lift of $\Gamma_n(H)$. A lift of a non-amenable graph is non-amenable, so we have:

Lemma 4.1. *Let G be a finitely generated group. If $\Gamma_n(H)$ is non-amenable for some quotient H of G, then $\Gamma_n(G)$ is non-amenable.*

The following fact is well-known. It is related to the fact that $\Gamma_2(\mathbb{Z}^k)$ is a Schreier graph of $SL(2,\mathbb{Z})$, which has property (τ) with respect to its congruence subgroups (see [LZ]).

Lemma 4.2. *The product replacement graph $\Gamma_2(\mathbb{Z}^k)$ is non-amenable for every $k > 0$.*

We refer the reader to Section 6 for the proofs.

4.2. **Proof of Theorem 1.3.**

Let G be a finitely generated virtually indicable group. Let H be a finite index subgroup of G which has $\mathbb{Z}$ as a quotient. Then H is also finitely generated. Thus, the statement that H has $\mathbb{Z}$ as a quotient is equivalent to the statement that $[H : H'] = \infty$, where $H' = [H,H]$ is the commutator subgroup of H. This, in turn, is equivalent to the statement that $[G : H'] = \infty$. Let $H^\circ = \bigcap_{g\in G} gHg^{-1}$ be the normal core of H in G. Then we also have $[G : H^\circ] < \infty$ and $[G : (H^\circ)'] = \infty$. That is, H° satisfies the same hypotheses as H, so by replacing H with H° if necessary, we may assume that H is normal in G.

We have that $H/H' \cong \mathbb{Z}^r \times A$ for some positive integer r and some finite abelian group A. Let N be the kernel of the corresponding homomorphism $H \to \mathbb{Z}^r$. Then N is a characteristic subgroup of H, and therefore N is a normal subgroup of G.

By Lemma 4.1, it is enough to show that $\Gamma_n(G/N)$ is non-amenable for sufficiently large n. Thus, by replacing G with G/N and H with H/N, we may assume N is trivial, and hence $H \cong \mathbb{Z}^r$.

Fix $n \geq \log_2 |G/H|$. Consider a generating tuple $S = (s_1, \ldots, s_{n+2}) \in \Gamma_{n+2}(G)$, and the corresponding tuple $\widetilde{S} = (\widetilde{s}_1, \ldots, \widetilde{s}_{n+2}) \in \Gamma_n(G/H)$. Every generating $(n+1)$-tuple of G/H is redundant (see e.g. [P1, 2.2]), so a bounded number of Nielsen moves in $\Gamma_n(G/H)$ sends

$$(\widetilde{s}_1, \ldots, \widetilde{s}_{n+2}) \rightsquigarrow (\widetilde{t}_1, \ldots, \widetilde{t}_n, 1, 1).$$

The same Nielsen moves in $\Gamma_n(G)$, then, send

$$(s_1, \ldots, s_{n+2}) \rightsquigarrow (t_1, \ldots, t_n, h_1, h_2),$$

where $h_1, h_2 \in H$. If h_1 and h_2 are both trivial, then $G = \langle t_1, \ldots, t_n \rangle$, so in at most $[G : H]$ more Nielsen moves, we can reach an element of that form with h_1, h_2 not both trivial.

For every nontrivial subgroup $K < H$ and every $T = (t_1, \ldots, t_n) \in G^n$ with $G = \langle T, K \rangle$, there is a graph embedding $\Gamma_2(K) \to \Gamma_{n+2}(G)$ given by

$$(h_1, h_2) \mapsto (t_1, \ldots, t_n, h_1, h_2),$$

and the images of these embedding are disjoint. Let Δ denote the union of these embeddings. In the previous paragraph, we showed that every vertex of $\Gamma_{n+2}(G)$ is a bounded distance away from Δ. Since each K satisfies $K \cong \mathbb{Z}^k$ for some $1 \leq k \leq r$, we have

$$h(\Delta) \geq \inf_{T,K} h(\Gamma_2(K)) \geq \min_{1\leq k\leq r} h\left(\Gamma_2(\mathbb{Z}^k)\right) > 0.$$

Thus, Δ is non-amenable. Lemma 3.1 implies that $\Gamma_{n+2}(G)$ is also non-amenable. □

5. Examples

Theorems 1.2 and 1.3 combine to show that several nice classes of groups satisfy Conjecture 1.1. First of all, there are classes of groups known to be uniformly non-amenable:

Theorem 5.1. *The following classes of groups are uniformly non-amenable:*

(i) *non-elementary word-hyperbolic groups* [A+],
(ii) *non-elementary relatively hyperbolic groups* [Xia],
(iii) *large groups (i.e. groups with a finite index subgroup which has F_2 as a quotient)* [A+],
(iv) *groups which act acylindrically on a simplicial tree without fixed points, and are not virtually cyclic* [A+],
(v) *free Burnside groups $B(m,n)$ with $m \geq 2$ and sufficiently large odd n* [O1],
(vi) *finitely generated groups with positive first ℓ^2-Betti number* [LPV].

Combining this result with Theorem 1.2, we obtain:

Corollary 5.2. *The product replacement graph $\Gamma_n(G)$ is non-amenable for every $n \geq d(G)$, if G belongs to one of the classes* (i)-(vi) *of Theorem 5.1.*

Using Theorem 1.3, we can extend this result to a larger class of groups, at the cost of a somewhat weaker conclusion. First, we make the following observation, which we prove in Section 6:

Lemma 5.3. *Let G be an infinite finitely generated group. If G is elementary amenable, then G is virtually indicable.*

By Gromov's Theorem [Gro], every infinite group of polynomial growth is virtually nilpotent, and therefore elementary amenable. By definition, infinite elementary hyperbolic groups contain $\mathbb{Z}$ as a finite index subgroup. Finally, every virtually solvable group is elementary amenable.

We also have the following theorem, which follows from a stronger version of the Tits alternative proven in [BG].

Theorem 5.4 ([BG, Theorem 1.5])**.** *If G is a linear group, then either G is virtually solvable, or G is uniformly non-amenable.*

Combining these observations with Corollary 5.2, we obtain:

Corollary 5.5. *The product replacement graph $\Gamma_n(G)$ is non-amenable for sufficiently large n, if G is an infinite finitely generated group which belongs to one of the following classes:*

(i) *elementary amenable groups,*
(ii) *groups of polynomial growth,*
(iii) *word-hyperbolic groups,*
(iv) *relatively hyperbolic groups,*
(v) *linear groups.*

6. Proofs of Lemmas

We now prove the lemmas used in the previous sections. The arguments in this section are standard, but we need the results in a specific form.

6.1. Proof of Lemma 3.1.

Let Γ and Γ' be any infinite graphs, where Γ is non-amenable. Let $f : \Gamma \to \Gamma'$ be an injective Lipschitz map with Lipschitz constant C. Suppose that $d(x, f(\Gamma)) \leq D$ for every $x \in \Gamma'$.

Given a finite set of vertices $X \subset \Gamma'$, define the *r-neighborhood* of X to be

$$X^{(r)} = \{v \in \Gamma' \mid d(v, X) \leq r\}.$$

Let $d \geq 2$ be an upper bound on the degrees of vertices in Γ' and Γ. Suppose $\left|X^{(r)}\right| \geq \alpha\,|X|$ for some $\alpha > 1$. Then there are at least $(\alpha - 1)\,|X|$ paths of length r or less from X to $\overline{X}$, each of which contains at least one edge leaving X. Each such edge occurs in at most $rd^{r-1} + (r-1)d^{r-2} + ... + 1 \leq r^2 d^{r-1}$ of these paths, so

$$|\partial X| \geq \frac{\alpha - 1}{r^2 d^{r-1}}\,|X|\,.$$

Thus, it is enough to show that there is positive integer r and a constant $\alpha > 1$ such that $\left|X^{(r)}\right| \geq \alpha\left|X\right|$ for every finite subset $X \subset \Gamma'$.

Let C and D be as above. Given a finite $X \subset \Gamma'$, every vertex of X is within D steps of some $v \in f(\Gamma)$, and for each $v \in f(\Gamma)$ there are at most $d^D + d^{D-1} + \cdots + 1 \leq d^{D+1}$ vertices within D steps of v. It follows that

$$\left|f^{-1}\big(X^{(D)}\big)\right| \geq \left|X^{(D)} \cap f(\Gamma)\right| \geq \left|X\right| / d^{D+1}.$$

Since Γ is non-amenable, there are at least $h(\Gamma)\left|X\right|/d^{D+1}$ edges leaving $f^{-1}(X^{(D)})$, and therefore at least $h(\Gamma)\left|X\right|/d^{D+2}$ vertices $v \in \Gamma$ with $d(v, f^{-1}(X^{(D)})) = 1$. Each such v maps to a unique $v' \in \Gamma'$ with $v' \notin X^{(D)}$ and $d(v', X^{(D)}) \leq C$. Hence,

$$\left|X^{(D+C)}\right| \geq \left|X^{(D)}\right| + h(\Gamma)\left|X\right|/d^{D+2} \geq (1 + h(\Gamma)/d^{D+2})\left|X\right|,$$

as desired. □

6.2. **Proof of Lemma 3.2.**

Let G be a non-amenable group, and fix a generating n-tuple $S = (s_1, \ldots, s_n)$ of G. Let $H = Z(G)$. Define $\widetilde{G} = G/H$, and $\widetilde{S} = (\widetilde{s}_1, \ldots, \widetilde{s}_n) = (s_1H, \ldots, s_nH)$. Let $\pi : G \to \widetilde{G}$ denote the usual projection.

By picking representatives for each coset of H, we have a bijection between G and $\widetilde{G} \times H$. That is, elements of G can be represented in the form $(g, h) \in \widetilde{G} \times H$, where group operation is given by $(g_1, h_1)(g_2, h_2) = (g_1g_2, \star)$. In fact, because elements of H commute with everything, we must have

$$(g_1, h_1)(g_2, h_2) = (g_1g_2, \phi(g_1, g_2)h_1h_2),$$

for some function $\phi : \widetilde{G} \times \widetilde{G} \to H$. Then we can write the original generators s_i in this form: $s_i = (\widetilde{s}_i, t_i)$, for some $t_i \in H$.

It is enough to show that for every finite subset $X \subset \mathrm{Cay}(\widetilde{G}, \widetilde{S})$, we have $\left|\partial X\right| / \left|X\right| \geq h(G, S)$. Let X be a finite subset of $\mathrm{Cay}(\widetilde{G}, \widetilde{S})$, and let $Y = \pi^{-1}(X)$. Let K be the subgroup of H generated by

$$T = \Big(\phi(x, \widetilde{s}_i)t_i\Big)_{\substack{x \in X \\ 1 \leq i \leq n.}}$$

Then K is an abelian group, which implies it is amenable. Therefore, we have a sequence of finite subsets $B_1, B_2, \dots \subseteq \mathrm{Cay}(K, T)$ with $\left|\partial B_k\right| / \left|B_k\right| \to 0$ as $k \to \infty$. Let $C_k = \{(x, h) \,|\, x \in X, h \in B_k\}$.

Partition the set ∂C_k of edges leaving C_k into two sets:

$$\partial_{\mathrm{out}} C_k = E(C_k, G \setminus Y) \qquad \text{and} \qquad \partial_{\mathrm{in}} C_k = E(C_k, Y \setminus C_k).$$

We have that $\left|\partial_{\mathrm{out}}(C_k)\right| = \left|B_k\right|\left|\partial X\right|$ and $\left|\partial_{\mathrm{in}}(C_k)\right| \leq \left|X\right|\left|\partial B_k\right|$. Thus

$$h(G, S) \leq \frac{\left|\partial C_k\right|}{\left|C_k\right|} = \frac{\left|\partial_{\mathrm{out}}(C_k)\right| + \left|\partial_{\mathrm{in}}(C_k)\right|}{\left|C_k\right|} \leq \frac{\left|B_k\right|\left|\partial X\right| + \left|X\right|\left|\partial B_k\right|}{\left|X\right|\left|B_k\right|} = \frac{\left|\partial X\right|}{\left|X\right|} + \frac{\left|\partial B_k\right|}{\left|B_k\right|}.$$

Since the second term goes to 0 as $k \to \infty$, we have $h \leq \left|\partial X\right| / \left|X\right|$, as desired.

6.3. **Proof of Lemma 4.1.**

There is a characterization of non-amenability in terms of recurrent walks. Let $\Gamma = (V, E)$ be a nonempty d-regular graph. Let $p_\Gamma^{(k)}(v, v)$ denote the probability that the nearest neighbor random walk on Γ starting at v returns to v at time k. That is, $d^k p_\Gamma^{(k)}(v, v)$ is the number of walks of length k in Γ from v to v. We define the *spectral radius* of Γ to be

$$\rho(\Gamma) := \sup_{v \in V} \limsup_{k \to \infty} (p_\Gamma^{(k)}(v, v))^{1/k}.$$

Then Γ is non-amenable if and only if $\rho(\Gamma) < 1$ (see e.g. [Woe, §10]).

Let $\pi : G \to H$ be a surjective group homomorphism, where G is some finitely generated group. We extend π to a graph homomorphism $\pi : \Gamma_n(G) \to \Gamma_n(H)$ given by

$$\pi(s_1, \dots, s_n) = \big(\pi(s_1), \dots, \pi(s_n)\big).$$

This is a local graph isomorphism, in other words for each $S \in \Gamma_n(G)$, the map π induces a bijection between the edges leaving S and the edges leaving $\pi(S)$. It follows that walks in $\Gamma_n(H)$ starting at $\pi(S)$ lift uniquely to walks in $\Gamma_n(G)$ starting at S. Thus,

$$p^{(n)}_{\Gamma_n(G)}\big(S, S\big) \le p^{(n)}_{\Gamma_n(H)}\big(\pi(S), \pi(S)\big),$$

and therefore

$$\rho\big(\Gamma_n(G)\big) \le \rho\big(\Gamma_n(H)\big) < 1.$$

□

6.4. **Proof of Lemma 4.2.**

The subgraph of $\Gamma_2(\mathbb{Z})$ induced by $\{(a, b) \in \Gamma_2(\mathbb{Z}) \mid a, b > 0\}$ is a rooted binary tree, which has positive Cheeger constant. The same holds for the other three quadrants, so $\Gamma_2(\mathbb{Z})$ has a subgraph Δ which is a disjoint union of four binary rooted trees. The only vertices that don't lie in Δ are $(\pm 1, 0)$ and $(0, \pm 1)$, and Δ is non-amenable, so by Lemma 3.1, $\Gamma_2(\mathbb{Z})$ is non-amenable. By Lemma 4.1, it follows that $\Gamma_2(\mathbb{Z}^k)$ is non-amenable for every $k \ge 1$. □

6.5. **Proof of Lemma 5.3.**

Let $\mathcal{L}$ be the class of groups which are either virtually indicable, or finite, or not finitely generated. We want to show that every elementary amenable group belongs to $\mathcal{L}$. Clearly finite groups are in $\mathcal{L}$. Every finitely generated infinite abelian groups has $\mathbb{Z}$ as a quotient, so abelian groups are also in $\mathcal{L}$. Thus by the characterization of elementary amenable groups in [Chou] it is enough to show that $\mathcal{L}$ is closed under direct unions and extensions.

If G is finite or G is not finitely generated, then G belongs to $\mathcal{L}$, so we may suppose G is infinite and finitely generated. Suppose G is a direct union of groups $G_i \in \mathcal{L}$. Since G is finitely generated, $G = G_i$ for some i, so $G \in \mathcal{L}$. Now suppose G is an extension of G'' by G' where $G', G'' \in \mathcal{L}$. Let $\pi : G \to G''$ be the projection with kernel G'. Since G is finitely generated, so is G''. If G'' is infinite, then it must be virtually indicable, so there is a finite index subgroup $H < G''$ which has $\mathbb{Z}$ is a quotient. Then $\pi^{-1}(H)$ is a finite index subgroup of G which has $\mathbb{Z}$ as a quotient, so G is virtually indicable. On the other hand, if G'' is finite, then G' is a finite index subgroup of G, so it is infinite and finitely generated. Thus it is virtually indicable, and therefore so is G. □

7. Final remarks

7.1. Our arguments can be followed through to give explicit bounds on Cheeger constants of $\Gamma_n(G)$, and on how large n must be in order for $\Gamma_n(G)$ to be non-amenable. To ease exposition, we did not track these bounds, and we did not present the arguments that would result in tight bounds. More detailed arguments with improved bounds will be presented in [Mal].

7.2. The elements of $\Gamma_n(G)$ can be identified with epimorphisms $F_n \to G$. The Nielsen moves then correspond to precomposition by one of the *Nielsen automorphisms* $R^{\pm 1}_{ij}, L^{\pm 1}_{ij}$ of $F_n = \langle x_1, \dots, x_n \rangle$ given by

$$L^{\pm 1}_{ij}(x_k) = \begin{cases} x_k & k \ne j \\ x_i^{\pm 1} x_k & k = j, \end{cases} \qquad R^{\pm 1}_{ij}(x_k) = \begin{cases} x_k & k \ne j \\ x_k x_i^{\pm 1} & k = j \end{cases}$$

with $i \ne j$ and $1 \le i, j \le n$. These automorphisms generate an index 2 subgroup of $\mathrm{Aut}(F_n)$, which we call $\mathrm{Aut}^+(F_n)$ (see e.g. [P1, LP]). Thus, every product replacement graph is a Schreier graph of $\mathrm{Aut}^+(F_n)$.

A well-known open question is whether $\mathrm{Aut}(F_n)$ with $n \geq 3$ has Kazhdan property (T)[3]. If $\mathrm{Aut}(F_n)$ has property (T) for a particular value of n, then $\Gamma_n(G)$ is non-amenable for every infinite n-generated group G. In fact, there is a uniform lower bound on the Cheeger constants of these graphs.

7.3. The *product replacement algorithm* is a well-known method for generating random elements of a finite group G. It begins with a generating n-tuple $S \in \Gamma_n(G)$, and takes a random walk on $\Gamma_n(G)$, outputting a random element of the resulting generating n-tuple. The running time of this algorithm depends on the mixing time of the random walk on $\Gamma_n(G)$ of G. The analysis of this mixing time is also related to the question in Subsection 7.2: if $\mathrm{Aut}(F_n)$ has property (T), then the finite product replacement graphs $\Gamma_n(G)$ form a family of expanders for any fixed n, and the random walk on such a graph has mixing time $O(\log|G|)$. It is known that for an appropriate value of n, the mixing time is polynomial in $\log|G|$ (see [P2]). For a survey on the product replacement algorithm, see [P1].

7.4. For a finite group G and a fixed number n, there is a lower bound on the coefficient of expansion of $\Gamma_n(G)$ in terms of the coefficients of expansion of $\mathrm{Cay}(G,S)$, ranging over all generating n-tuples S (see [GP]). Theorem 1.2 can be thought of as an analogue of this result for infinite groups, though the proofs differ.

7.5. A simple consequence of Conjecture 1.1 is the following.

Conjecture 7.1. [MP] *The product replacement graph $\Gamma_n(G)$ of an infinite finitely generated group G has exponential growth for sufficiently large n.*

Some progress on this conjecture is made in [MP]. Specifically, it is shown that it holds for all groups of polynomial growth, and all groups of exponential growth. It is also shown to hold for some groups of intermediate growth, including the Grigorchuk group.

7.6. In [MP] it was shown that if G has exponential growth, then $\Gamma_n(G)$ has exponential growth for every $n \geq d(G)+1$. The proof of Theorem 1.2 is easily modified to prove a slight improvement of this result: if G has exponential growth, then $\Gamma_n(G)$ has exponential growth for every $n \geq d(G)$. Moreover, if G has uniform exponential growth, then $\Gamma_n(G)$ also has uniform exponential growth.

7.7. The reader should note that there are several definitions of uniform non-amenability in the literature. A priori, our definition is weaker than the one in [A+], as we do not compare Cheeger constants with respect to generating tuples of different lengths. In turn, the definition in [A+] is weaker than the notion implicit in [O2] and called uniform non-amenability in [O1].

7.8. We have shown that both infinite elementary amenable groups and uniformly non-amenable groups have non-amenable product replacement graphs. A natural next step is to look at Conjecture 1.1 for groups in between those two classes. This includes every group of non-uniform exponential growth: such a group clearly cannot be uniformly non-amenable, and it has been shown that it cannot be elementary amenable either [O3]. In particular, the groups of non-uniform exponential growth constructed by Wilson in [W1, W2] fall between these two classes.

7.9. Groups which are neither elementary amenable nor uniformly non-amenable belong to one of two types: amenable groups which are not elementary amenable, and non-amenable groups which are not uniformly non-amenable.

An example of the first type is the Grigorchuk group $\mathbb{G}$ (see [Gri, dlH, GP]). It was shown in [MP] that its product replacement graphs $\Gamma_n(\mathbb{G})$ have exponential growth for $n \geq 5$, but the techniques do not appear to be strong enough to show non-amenability.

An example of the second type is the Baumslag-Solitar group $B(p,q)$ where p and q are relatively prime [O2, A+]. However, this group has $\mathbb{Z}$ as a quotient, so $\Gamma_n(B(p,q))$ is non-amenable for every

[3] The answer is known to be negative for $n \leq 3$ [Mc, GL].

$n \geq 2$. A more interesting example is the torsion group Q constructed in [O2]. Neither Theorem 1.2 nor Theorem 1.3 is enough to show that $\Gamma_n(Q)$ is non-amenable for some n.

7.10. Another example of interest is Thompson's group F (see [CFP]). Whether F is amenable is a well-known open problem, but it is known that it is not elementary amenable. Thompson's group F has $\mathbb{Z}^2$ as a quotient, and therefore $\Gamma_n(F)$ is non-amenable for every $n \geq 2$. However, the related groups T and V are both simple groups. Thus, they cannot be virtually indicable, and Theorem 1.3 does not apply. Note that T and V both have exponential growth, so by the results in [MP] they satisfy Conjecture 7.1.

8. Acknowledgements

I am grateful to Igor Pak for suggesting the problem of non-amenability of infinite product replacement graphs. I would like to thank Ana Khukhro, Denis Osin, and Yehuda Shalom for feedback on the paper, particularly for pointing out what is known about uniform non-amenability.

References

[A+] G. N. Arzhantseva, J. Burillo, M. Lustig, L. Reeves, H. Short and E. Ventura, Uniform non-amenability, *Adv. Math.* **197** (2005), 499–522.

[BL] H. Bäärnhielm and C. R. Leedham-Green, The product replacement prospector, *J. Symbolic Comput.* **47** (2012), 64–75.

[BLM] A. V. Borovik, A. Lubotzky and A. G. Myasnikov, The finitary Andrews-Curtis conjecture, in *Progr. Math.* **248**, Birkhäuser, Basel, 2005, 15–30.

[BG] E. Breuillard and T. Gelander, Uniform independence in linear groups, *Invent. Math.* **173** (2008), 225–263.

[CFP] J. W. Cannon, W. J. Floyd and W. R. Parry, Introductory notes on Richard Thompson's groups. *Enseign. Math.* **42** (1996), 215–256.

[Chou] C. Chou, Elementary amenable groups, *Illinois J. Math.* **24** (1980), 396–407.

[dlH] P. de la Harpe, *Topics in Geometric Group Theory*, University of Chicago Press, Chicago, 2000.

[Eva] M.J. Evans, Nielsen equivalence classes and stability graphs of finitely generated groups, in *Ischia group theory 2006*, World Sci. Publ., Hackensack, 2007, 103–119.

[GP] A. Gamburd and I. Pak, Expansion of product replacement graphs, *Combinatorica* **26** (2006), 411–429.

[Gri] R. I. Grigorchuk, Solved and unsolved problems around one group, in *Infinite Groups: Geometric, Combinatorial and Dynamical Aspects*, Birkhäuser, Basel, 2005, 117–218.

[GP] R. I. Grigorchuk and I. Pak, Groups of intermediate growth, an introduction, *L'Ens. Math.* **54** (2008), 251–272.

[Gro] M. Gromov, Groups of polynomial growth and expanding maps, *IHES Publ. Math.* **53** (1981), 53–78.

[GL] F. Grunewald and A. Lubotzky, Linear representations of the automorphism group of a free group, *Geom. Funct. Anal.* **18** (2009), 1564–1608.

[LP] A. Lubotzky and I. Pak, The product replacement algorithm and Kazhdan's property (T), *J. AMS* **14** (2001), 347–363.

[LZ] A. Lubotzky and A. Żuk, *On property* (τ), monograph in preparation.

[LPV] R. Lyons, M. Pichot and S. Vassout, Uniform non-amenability, cost, and the first ℓ^2-Betti number, *Groups Geom. Dyn.* **2** (2008), 595–617.

[Mal] A. Malyshev, *Combinatorics of finitely generated groups*, Ph.D. thesis, UCLA, in preparation.

[MP] A. Malyshev and I. Pak, Growth in product replacement graphs, `arXiv:1304.5320`.

[Mc] J. McCool, A faithful polynomial representation of Out F_3. *Math. Proc. Cambridge Philos. Soc.* **106** (1989), 207–213.

[Met] W. Metzler, On the Andrews-Curtis conjecture and related problems, in *Contemp. Math.* **44**, AMS, Providence, RI, 1985, 35–50.

[Myr] A. Myropolska, Andrews–Curtis and Nielsen equivalence relations on some infinite groups, `arXiv:1304.2668`.

[O1] D. V. Osin, Uniform non-amenability of free Burnside groups, *Arch. Math.* **88** (2007), 403–412.

[O2] D. V. Osin, Weakly amenable groups. *Contemp. Math.* **298** (2002), 105–113.

[O3] D. V. Osin, Algebraic entropy of elementary amenable groups, *Geom. Dedicata* **107** (2004), 133–151.

[P1] I. Pak, What do we know about the product replacement algorithm?, in *Groups and Computation III*, de Gruyter, Berlin, 2001, 301–347.

[P2] I. Pak, The product replacement algorithm is polynomial, in *Proc. FOCS 2000*, IEEE Comput. Soc. Press, Los Alamitos, CA, 2000, 476–485.

[W1] J. S. Wilson, On exponential growth and uniformly exponential growth for groups, *Invent. Math.* **155** (2004), 287–303.

[W2] J. S. Wilson, Further groups that do not have uniformly exponential growth, *J. Algebra* **279** (2004), 292–301.

[Woe] W. Woess, *Random walks on infinite graphs and groups*, Cambridge U. Press, Cambridge, 2000.

[Xia] X. Xie, Growth of relatively hyperbolic groups, *Proc. Amer. Math. Soc.* **135** (2007), 695–704.

CHAPTER 4

Lifts, derandomization, and diameters of Schreier graphs of Mealy automata

LIFTS, DERANDOMIZATION, AND DIAMETERS OF SCHREIER GRAPHS OF MEALY AUTOMATA

ANTON MALYSHEV* AND IGOR PAK*

Abstract. It is known that random 2-*lifts* of graphs give rise to expander graphs. We present a new conjectured derandomization of this construction based on certain *Mealy automata*. We verify that these graphs have polylogarithmic diameter, and present a class of automata for which the same is true. However, we also show that some automata in this class do not give rise to expander graphs.

1. Introduction

In [BL], Bilu and Linial showed that random 2-*lifts* of expanding graphs remain expanding with high probability. This gives a probabilistic construction of expander families. Several ways to derandomize this procedure are also given in [BL], but none of them give a *strongly explicit* description of a family of expander graphs. That is, a description in which the actual graph is much larger than working memory, but a computer can list neighbors of a vertex in polylogarithmic (in the size of the graph) time.

We consider the following two families of 2-lifts of graphs. The *Aleshin graphs* $A_0, A_1, A_2, \ldots$ are a sequence of 3-regular edge-labeled directed graphs. The first graph A_0 is defined to be a single vertex with three self-loops labeled a, b, and c. Given the graph A_n, the next graph A_{n+1} is defined as a certain *graph lift* of A_n: Each vertex $v \in A_n$ lifts to two vertices $v_0, v_1 \in A_{n+1}$, and the edges transform as follows:

$$\begin{tikzcd} v \arrow[r, "a"] & w & \text{lifts to} & v_0 \arrow[r, dotted, "c"] & w_0 \\ & & & v_1 \arrow[r, dotted, "c"] & w_1 \end{tikzcd}$$

$$\begin{tikzcd} v \arrow[r, thick, "b"] & w & \text{lifts to} & v_0 \arrow[dr, "a"] & w_0 \\ & & & v_1 \arrow[ur, thick, "b"'] & w_1 \end{tikzcd}$$

$$\begin{tikzcd} v \arrow[r, dotted, "c"] & w & \text{lifts to} & v_0 \arrow[dr, thick, "b"] & w_0 \\ & & & v_1 \arrow[ur, "a"'] & w_1. \end{tikzcd}$$

That is, e.g., if A_n has an edge labeled c from v to w, then A_{n+1} has an edge labeled b from v_0 to w_1, and an edge labeled a from v_1 to w_0.

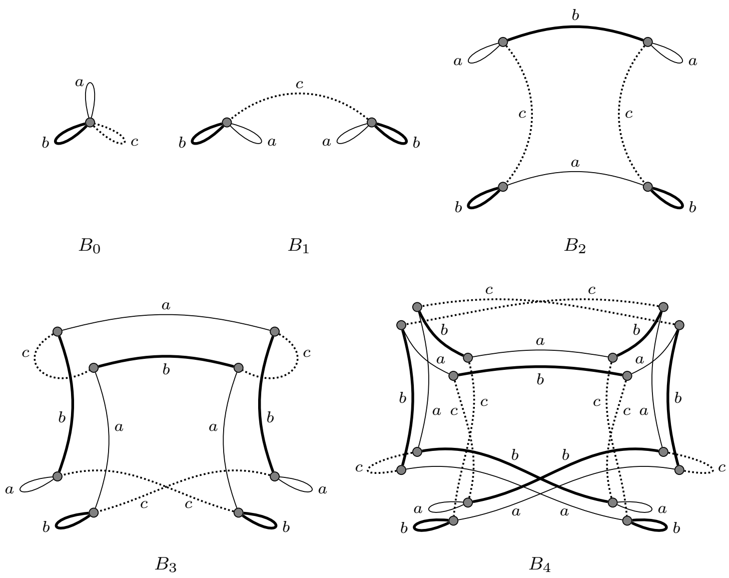


Figure 1. The Bellaterra graphs

Another family, the *Bellaterra graphs* $B_0, B_1, B_2, \ldots$ is defined the same way, except with transformation rules

$$v \xrightarrow{a} w \qquad \text{lifts to} \qquad \begin{tikzcd} v_0 \arrow[dr, dotted, "c"] & w_0 \\ v_1 \arrow[ur, dotted, "c"'] & w_1 \end{tikzcd}$$

$$v \xrightarrow{b} w \qquad \text{lifts to} \qquad \begin{matrix} v_0 \xrightarrow{a} w_0 \\ v_1 \xrightarrow{b} w_1 \end{matrix}$$

$$v \xrightarrow{c} w \qquad \text{lifts to} \qquad \begin{matrix} v_0 \xrightarrow{b} w_0 \\ v_1 \xrightarrow{a} w_1. \end{matrix}$$

It is not hard to check that the reverse of every edge in B_n is also in B_n, so these can be thought of as undirected graphs. The first few graphs in this family are pictured in Figure 1.

The main result of this paper is the following theorem:

Theorem 1.1. *The diameter of the Aleshin graphs* $\{A_i\}_{i=1}^{\infty}$ *and Bellaterra graphs* $\{B_i\}_{i=1}^{\infty}$ *grows at most quadratically in* n, *i.e.,*

$$\operatorname{diam}(A_n) = O(n^2) \quad \textit{and} \quad \operatorname{diam}(B_n) = O(n^2) \quad \textit{as} \quad n \to \infty.$$

Prior to this paper, there were no nontrivial bounds on the diameter of A_n; even subexponential bounds remained out of reach. Note also that in principle we can start with *any* 3-labeled graph in place of $A_0 = B_0$, and proceed making lifts as

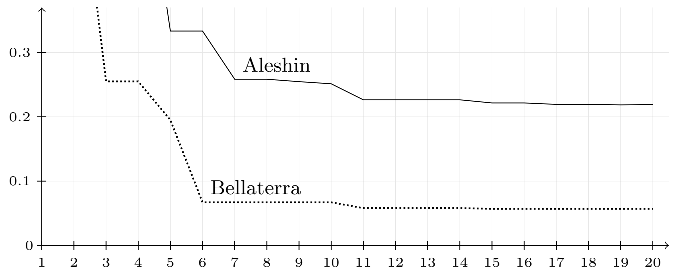


Figure 2. Eigenvalue gaps of the Bellaterra and Aleshin graphs.

above. We do not consider these in the paper, and our algebraic techniques do not apply.

Observe that both families of graphs are *very explicit* in the following sense: there is a polynomial time algorithm which, given a number n and $v \in \Gamma_n$, lists the neighbors of n. "Polynomial time" here refers to a runtime which is polynomial in the number of bits necessary to describe the input. It takes n bits to describe a vertex of B_n or A_n, so the algorithm should run in time $O(n^d)$, for some d.

In particular, it follows that they are *strongly explicit* in the sense of [BL]: There is a polynomial time (in the size of the inputs) algorithm which, given a number n, and vertices $v, w \in \Gamma_n$, decides whether v and w are adjacent in Γ_n.

As we will see below, these graphs can be described in terms of invertible *Mealy automata*. The associated automata are small: they act on binary strings and have only 3 states. A detailed study of all such small automata was performed in [B+]. The Bellaterra and Aleshin automata are numbered 846 and 2240 in that article. They are the only nontrivial *bireversible* ones. Spectra of the first few associated graphs are also computed in [B+], and the data suggest that the Aleshin graphs are a family of expanders with eigenvalue gap roughly 0.2.

Conjecture 1.2. *The Aleshin graphs $\{A_i\}_{i=1}^{\infty}$ are a family of two-sided expanders.*

Here by *two-sided* we mean that both the second largest and the smallest eigenvalues of 3-regular graphs A_n are bounded away as follows: $\lambda_2 < 3 - \varepsilon$ and $\lambda_n > -3 + \varepsilon$ (see e.g. [Tao]).

Though it is less clear from the data in [B+], our own computations (see Figure 2) suggest that the Bellaterra graphs are also expanders, with eigenvalue gap roughly 0.05, so we make the stronger conjecture:[1]

Conjecture 1.3. *The Bellaterra graphs $\{B_i\}_{i=1}^{\infty}$ are a family of two-sided expanders.*

If so, they are a strongly explicit derandomization of the probabilistic construction in [BL]. One consequence of being an expander family is logarithmic diameter growth with respect to the size of the graph, so if Conjecture 1.2 holds then $\operatorname{diam}(A_n)$ grows linearly in n, stronger claim than in the theorem.

Unfortunately, we are not near proving either conjectures and in fact our tools are too weak to prove them. Later in the paper, we state and prove general conditions

[1] See Remark 2.

on an automaton which guarantee polynomial diameter growth in the associated graphs (Section 8). We then prove that for some automata which satisfy those conditions, we do not get expanders (see Section 10). In other words, a different, perhaps combinatorial technique is needed to prove the expansion.

2. Mealy automata

The Bellaterra graphs $\{B_n\}_{n=1}^{\infty}$ are *very explicit* in the sense of [HLW].[2] That is, there is a polynomial time algorithm which, given a number n and a vertex $v \in B_n$, lists the neighbors of v in B_n. It takes n bits to describe a vertex in B_n, so the runtime of the algorithm should be polynomial in n.

In fact, there is a linear time algorithm. Even more strongly, the computation can be implemented with a Mealy automaton, i.e., a finite state automaton which outputs a letter each time it reads a letter.

Definition 2.1. A *Mealy automaton* $\mathcal{M} = (Q, A, \tau, \sigma)$ is a pair of finite sets Q, A, together with functions $\sigma : Q \times A \to A$, and $\tau : Q \times A \to Q$.

The sets Q and A are called the *states* and *alphabet*, respectively. The functions σ and τ are called the *output* and *transition* functions, respectively. When $|Q| = q$ and $|A| = a$, we call $\mathcal{M}$ a (q, a)-automaton. We adopt the following notations:

$$\begin{aligned} {}^{q}x &= \sigma_q(x) = \sigma(q, x) \\ q^x &= \tau_x(q) = \tau(q, x). \end{aligned}$$

Let A^* and A^∞ denote the set of finite and infinite words in the alphabet A, respectively, and let $A^{*,\infty} = A^* \cup A^\infty$ denote the set of all words in A. A Mealy automaton in the state $q \in Q$ acts in a length-preserving way on words in $A^{*,\infty}$ by reading the first letter x, outputting the letter $\sigma(q, x)$, and acting on the rest of the word from the state $\tau(q, x)$. That is, each $q \in Q$ has a corresponding length-preserving map $A^{*,\infty} \to A^{*,\infty}$ defined recursively by

$$\begin{aligned} {}^{q}(x_0x_1 \dots x_n) &= y_0 \, {}^{r}(x_1 \dots x_n), \\ \text{and} \quad {}^{q}(x_0x_1x_2 \dots) &= y_0 \, {}^{r}(x_1x_2 \dots), \end{aligned}$$

where $y_0 = \sigma(q, x_0)$ and $r = \tau(q, x_0)$. This extends to a left action of finite words Q^* on words in $A^{*,\infty}$ via, e.g.,

$${}^{qr}s = {}^{q}({}^{r}s).$$

So we've defined an extension of $\sigma : Q \times A \to A$ to a map $\boldsymbol{\sigma} : Q^* \times A^{*,\infty} \to A^{*,\infty}$ given by

$$\boldsymbol{\sigma}(w, s) = \boldsymbol{\sigma}_w(s) = {}^{w}s.$$

A Mealy automaton can be depicted with a Moore diagram: a directed graph with a vertex for each state $q \in Q$ and a labeled edge

$$q \xrightarrow{x\,:\,y} r$$

for every $q \in Q$ and every $x \in A$, where $y = \sigma(q, x)$ and $r = \tau(q, x)$. That is, an edge $q \xrightarrow{x\,:\,y} r$ denotes that if the Mealy automaton is in state q and reads the letter x, then it outputs the letter y and transitions to the state r. We will sometimes simply write $q \xrightarrow{x\,:\,y} r$ to denote that $y = \sigma(q, x)$ and $r = \tau(q, x)$.

[2] Sometimes, these are called *fully explicit*, see e.g. [Vad].

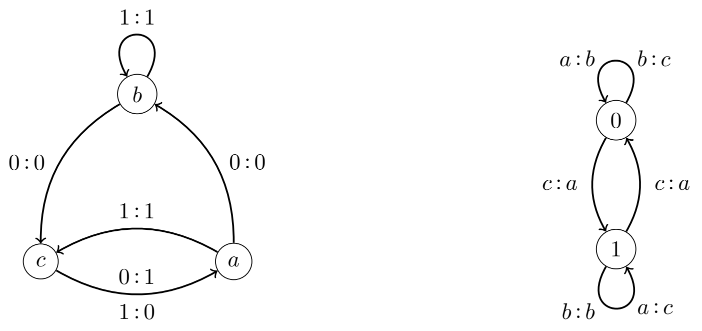


Figure 3. The Bellaterra automaton $\mathcal{B}$ and its dual $\overline{\mathcal{B}}$

Example 2.2. Consider the *Bellaterra automaton* $\mathcal{B}$ pictured in Figure 3. More formally, $\mathcal{B} = (Q, A, \tau, \sigma)$ is defined by

$$A = \{0, 1\}, \quad Q = \{a, b, c\}$$
$$\sigma_a = \sigma_b = (0)(1), \quad \sigma_c = (0\ 1),$$
$$\text{and } \tau_0 = (a\ b\ c), \quad \tau_1 = (a\ c)(b),$$

where we use the usual cycle notation for permutations, so e.g., $\tau_0(a) = b$, $\tau_0(b) = c$, $\tau_0(c) = a$.

Then given a number n, the Bellaterra graph B_n can be described as the graph whose vertices are length n binary strings, with an edge

$$s \xrightarrow{q} ({}^{q}s)$$

for each vertex $s \in A^n$ and each state $q \in Q$. For example, we have

$$^{c}(0000) = 1\,{}^{a}(000) = 10\,{}^{b}(00) = 100\,{}^{c}(0) = 1001,$$
$$\text{so} \quad 0000 \xrightarrow{c} 1001.$$

Some symmetry between states and letters of a Mealy automaton is already apparent in the definition. The nature of this symmetry becomes more clear if we consider computing compositions of maps associated to the states of an automaton, we have, e.g.,

$$^{q_1q_0}(x_0x_1 \dots x_n) = {}^{q_1}(y_0\,{}^{r_0}(x_1 \dots x_n)) = z_0\,{}^{r_1r_0}(x_1 \dots x_n) = \dots,$$

where $q_0 \xrightarrow{x_0\,:\,y_0} r_0$, and $q_1 \xrightarrow{y_0\,:\,z_0} r_1$. The computation proceeds by taking any instance of ${}^{q}(x \dots)$ in the expression, and replacing it with $y\,{}^{r}(\dots)$, where $q \xrightarrow{x\,:\,y} r$.

If we ignore parentheses, states in Q and letters in A play a symmetric role in this process, except that letters in Q are written higher and disappear when they are at the right side of the expression. Taking this symmetry into account, the automaton also naturally defines an action of the letters in A on finite words in Q^*:

$$(q_n \dots q_1q_0)^x = (q_n \dots q_1)^y r_0,$$

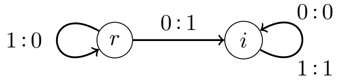


Figure 4. The binary adding machine

where $q_0 \xrightarrow{x\,:\,y} r_0$. Letters in A also act on the set of *left*-infinite words in the alphabet Q:

$$(\dots q_2 q_1 q_0)^x = (\dots q_2 q_1)^y r_0.$$

We let $Q^{-\infty}$ denote this set of left-infinite words, and let $Q^{-\infty,*}$ denote $Q^* \cup Q^{-\infty}$, so we have an action of A on $Q^{-\infty,*}$. This naturally extends to a *right* action of A^* on $Q^{-\infty,*}$, via, e.g.

$$w^{xy} = (w^x)^y.$$

So we have defined a map $\boldsymbol{\tau} : Q^{-\infty,*} \times A^* \to Q^{-\infty,*}$, given by

$$\boldsymbol{\tau}(w,s) = \boldsymbol{\tau}_s(w) = w^s.$$

It is straightforward to check that for any $s \in A^*, t \in A^{*,\infty}, w \in Q^*, v \in Q^{-\infty,*}$, the actions we have defined satisfy the following relations:

$$\begin{aligned} {}^{w}(st) &= \widetilde{s}\;{}^{\widetilde{w}}(t), \quad \text{and} \\ (vw)^s &= (v)^{\widetilde{s}}\,\widetilde{w}, \end{aligned}$$

where $\widetilde{s} = {}^{w}s$ and $\widetilde{w} = w^s$.

If we need to specify the automaton, we will write $\boldsymbol{\sigma}_{\mathcal{M}} = \boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_{\mathcal{M},w} = \boldsymbol{\sigma}_w$, and similarly for $\boldsymbol{\tau}$.

With this symmetry in mind, it is sensible to define the *dual* of an automaton $\mathcal{M} = (Q, A, \tau, \sigma)$ to be the automaton $\overline{\mathcal{M}} = (\widehat{Q}, \widehat{A}, \widehat{\tau}, \widehat{\sigma})$ given by interchanging the roles of the states and alphabet. That is, we take

$$\widehat{A} = Q, \quad \widehat{Q} = A, \quad \widehat{\sigma}(a,q) = \tau(q,a), \quad \text{and} \quad \widehat{\tau}(a,q) = \sigma(q,a).$$

In other words, for $q, r \in Q$ and $x, y \in A$, we have $x \xrightarrow{q\,:\,r} y$ in $\overline{\mathcal{M}}$ if and only if $q \xrightarrow{x\,:\,y} r$ in $\mathcal{M}$.

Computations in the dual automaton are computations in the original automaton, with each step written backwards. It follows that, e.g., for every $s \in A^*$ and $w \in Q^*$ we have

$$\boldsymbol{\sigma}_{\overline{\mathcal{M}}}(s,w) = \overline{\boldsymbol{\tau}_{\mathcal{M}}(\overline{w},\overline{s})},$$

where $\overline{u}$ denotes the reversal of u.

Example 2.3. The dual of the Bellaterra automaton is also pictured in Figure 3.

Example 2.4. Let $A = \{0,1\}$. Consider the Mealy automaton pictured in Figure 4. The map $\boldsymbol{\sigma}_r : \{0,1\}^* \to \{0,1\}^*$ is simply addition of 1, where length n words in A^* are interpreted as binary representations of numbers modulo 2^n, with the least significant digit on the left.

We say a Mealy automaton is *invertible* if σ_q is invertible for every $q \in Q$. This occurs if and only if the endomorphism $\boldsymbol{\sigma}_w : A^* \to A^*$ is invertible for every $w \in Q^*$. We are primarily interested in invertible automata, though our results can be generalized to the non-invertible case.

The *inverse* of an invertible automaton $\mathcal{M} = (Q, A, \tau, \sigma)$ is the automaton $\mathcal{M}^{-1} = (Q', a, \tau', \sigma')$ given by

$$Q = \{q' \mid q \in Q\}\,, \quad \sigma'_{q'} = \sigma_q^{-1}, \quad \tau'(q', a) = \tau(q, \sigma_q^{-1}(a)).$$

It is straightforward to check that $\boldsymbol{\sigma}_{\mathcal{M}^{-1},q'} = \boldsymbol{\sigma}_{\mathcal{M},q}^{-1}$ for every $q \in Q$.

Consider two automata $\mathcal{M} = (Q, A, \tau, \sigma)$, $\mathcal{M}' = (Q', A, \tau', \sigma')$ acting on the same alphabet, with Q, Q' disjoint. Their *union* is the automaton $\mathcal{M} \cup \mathcal{M}' = (Q \cup Q', A, \tau'', \sigma'')$, where

$$\tau''(q,a) = \begin{cases} \tau(q,a) & q \in Q \\ \tau'(q,a) & q \in Q' \end{cases} \qquad \text{and} \qquad \sigma''(q,a) = \begin{cases} \sigma(q,a) & q \in Q \\ \sigma'(q,a) & q \in Q' \end{cases}$$

For example, $\mathcal{M} \cup \mathcal{M}^{-1}$ is an automaton with twice as many states as Q, in which every state q has an inverse state q' with $\boldsymbol{\sigma}_{q'} = \boldsymbol{\sigma}_q^{-1}$.

We say an automaton is *reversible* if its dual is invertible.

We say an automaton is *bireversible* if it is invertible, reversible, and its inverse is reversible. Note that the last condition does not follow from the other two. For example, the three-state automaton in Figure 8 is reversible and invertible, but not bireversible.

3. Schreier graphs

For our purposes, graphs are locally finite, directed, and may have self-loops and repeated edges. A graph is *regular* if the indegree and outdegree are the same across all vertices.

Let Γ be a graph. Given vertices $v, w \in \Gamma$, we write $v \longrightarrow_\Gamma w$ if there is an edge in Γ from v to w. We write $d_\Gamma(v, w)$ for the distance between v and w, i.e. the length of the shortest undirected path between v and w. When there is no such path, we take $d_\Gamma(v, w) = \infty$. Given a nonnegative integer r, the *ball* of radius r centered at v is the set

$$B_\Gamma(v, r) = \{w \in \Gamma : d(v, w) \leq r\}.$$

The *diameter* of Γ is defined to be

$$\operatorname{diam}(\Gamma) = \max_{v,w \in \Gamma} d_\Gamma(v, w).$$

When it is clear from context what graph we are discussing, we will drop the subscripts and simply write $v \longrightarrow w$, $d(v, w)$, and $B(v, r)$.

In Example 2.2 we described the Bellaterra graphs in terms of a Mealy automaton. In the same way, we can associate a sequence of graphs to any Mealy automaton. Since we are primarily concerned with regular graphs, we require the automaton to be invertible.

Definition 3.1. Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be an invertible Mealy automaton. Given $n \in \{1, 2, \dots\} \cup \{\infty\}$, the *Schreier graph* $\Gamma_{\mathcal{M},n}$ is a directed graph, defined as follows: The vertices of $\Gamma_{\mathcal{M},n}$ are length n words in $A^{*,\infty}$, i.e. elements of A^n. For each vertex $s \in \Gamma_{\mathcal{M},n}$ and each state $q \in Q$, the Schreier graph $\Gamma_{\mathcal{M},n}$ has an edge

$$s \longrightarrow {}^q s.$$

Clearly, the number of edges leaving a vertex is $|Q|$. The Schreier graph of the inverse automaton, $\Gamma_{\mathcal{M}^{-1},n}$, is simply $\Gamma_{\mathcal{M},n}$ with the edges reversed. So, the number of edges entering a given vertex in $\Gamma_{\mathcal{M}^{-1},n}$ is also $|Q|$, and $\Gamma_{\mathcal{M},n}$ is regular.

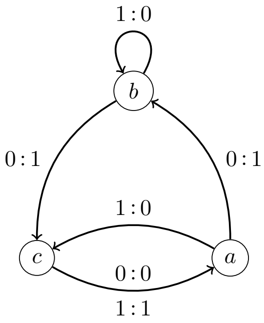


Figure 5. The Aleshin automaton $\mathcal{A}$

Example 3.2. The n-th Bellaterra graph B_n is the Schreier graph $\Gamma_{\mathcal{B},n}$, where $\mathcal{B}$ is the Bellaterra automaton, pictured in Figure 3.

Example 3.3. The n-th Aleshin graph A_n is the Schreier graph $\Gamma_{\mathcal{A},n}$, where $\mathcal{A}$ is the Aleshin automaton, first considered in [A], pictured in Figure 5.

4. Automaton groups

Let $\mathcal{M} = (Q, A, \tau, \sigma)$ be an invertible Mealy automaton. As seen in Section 2, we have invertible maps $\boldsymbol{\sigma}_q : A^{*,\infty} \to A^{*,\infty}$ for each $q \in Q$. This gives an action of the free group F_Q on $A^{*,\infty}$. We can extend the definition of $\boldsymbol{\sigma}$ as follows: For $w \in F_Q$, we can define $\boldsymbol{\sigma}_w$ in the natural way, e.g.,

$$\boldsymbol{\sigma}_{qr^{-1}} = \boldsymbol{\sigma}_q \, \boldsymbol{\sigma}_r^{-1}.$$

As usual, we will adopt the notational convention

$$\boldsymbol{\sigma}_w(s) = \boldsymbol{\sigma}(w, s) = {}^{w}s.$$

The *automaton group* associated to $\mathcal{M}$ is the group $G_{\mathcal{M}}$ generated by the automorphisms $\boldsymbol{\sigma}_q$.

For example, letting $\mathcal{A}$ denote the Aleshin automaton, it was shown in [VV] that $\boldsymbol{\sigma}_{\mathcal{A},a}$, $\boldsymbol{\sigma}_{\mathcal{A},b}$, and $\boldsymbol{\sigma}_{\mathcal{A},c}$ satisfy no nontrivial relation, so $G_{\mathcal{A}}$ is the free group F_3. However, it is straightforward to check that $\boldsymbol{\sigma}^2_{\mathcal{B},a} = \boldsymbol{\sigma}^2_{\mathcal{B},b} = \boldsymbol{\sigma}^2_{\mathcal{B},c} = \mathrm{id}$, where $\mathcal{B}$ is the Bellaterra automaton. It can be shown (see, e.g., [Nek, B+]) that the words satisfy no other relation, so we say $G_{\mathcal{B}} \cong \langle a, b, c \mid a^2, b^2, c^2 \rangle = C_2 * C_2 * C_2$.

Information about $G_{\mathcal{M}}$ as an abstract group can be used to obtain information about the Schreier graphs $\Gamma_{\mathcal{M},n}$. See Remarks 7 and 6.

5. Trees and automorphisms

In this context it is natural to think of the set of finite words A^* as vertices in a regular rooted tree, where the empty word is the root and the children of the word s are the words sx for $x \in A$. We will need to talk about rooted trees more generally, so we make the following definitions.

Definition 5.1. A *rooted tree* (or simply *tree*) is a graph $\mathbf{T}$ with a distinguished root vertex $r \in \mathbf{T}$ such that for each $v \in \mathbf{T}$ there is exactly one directed path from r to v. The *level* of v, denoted $\ell(v)$ is the length of this path. The *n-th level* of $\mathbf{T}$, denoted $\mathbf{T}_n$ is the set of all vertices $v \in \mathbf{T}$ such that $\ell(v) = n$. A *subtree* of $\mathbf{T}$ is a subgraph containing r which is itself a rooted tree. A *tree isomorphism* between two trees $\mathbf{S}$ and $\mathbf{T}$ is a graph isomorphism which sends the root of $\mathbf{S}$ to the root of $\mathbf{T}$. An *automorphism* of $\mathbf{T}$ is an isomorphism from $\mathbf{T}$ to $\mathbf{T}$. The automorphisms of $\mathbf{T}$ form a group, and we denote it $\mathrm{Aut}(\mathbf{T})$.

Then, given a Mealy automaton $\mathcal{M} = (Q, A, \sigma, \tau)$, for any $q \in Q$ the map $\boldsymbol{\sigma}_q : A^* \to A_*$ is a tree automorphism. That is, $\boldsymbol{\sigma}_q$ is a bijection which fixes the empty word, and sends children of x to children of $\boldsymbol{\sigma}_q(x)$. In other words, for every $s \in A_*$ and $x \in A$ there is some $y \in A$ such that

$$\boldsymbol{\sigma}_q(sx) = \boldsymbol{\sigma}_q(s)y.$$

Of course, it follows that for any $w \in Q^*$, the map $\boldsymbol{\sigma}_w : A^* \to A_*$ is a composition of tree automorphisms and is itself a tree automorphism.

Infinite words, i.e., elements of A^∞, can be thought of as rays in the tree A^*, and $\boldsymbol{\sigma}_w$ acts on them in the natural way.

Note that in order to think of $\tau_a : Q^* \to Q^*$ as a tree automorphism, we must think of Q^* as a tree in the reverse way, i.e. the children of w are of the form qw for $q \in Q$, rather than of the form wq.

Given a tree automorphism $g : A^* \to A^*$ and a word $s \in A^*$, the *section* of g at s, is the tree automorphism $g|_s : A^* \to A^*$ defined by

$$g(st) = g(s)g|_s(t).$$

Note that we are using a canonical identification between branches of the tree A^*. There need not be such an identification in a general tree, so this definition of sections is specific to trees of words.

We call a tree automorphism $\alpha : A^* \to A^*$ *automatic* if it $\alpha = \boldsymbol{\sigma}_{\mathcal{M},q}$ for state q of some Mealy automaton $\mathcal{M}$. Equivalently, α is automatic if and only if it has finitely many sections. The set of automatic automorphisms forms a subgroup $\mathrm{FAut}(A^*) < \mathrm{Aut}(A^*)$.

An automorphism $g : A^* \to A^*$ is determined by its action on the first level, $(A^*)_1 = A^1 = A$, and its sections $g|_x$ at all $x \in A$. If $\rho : A \to A$ is a permutation, then for notational convenience we can extend ρ as an automorphism $A^* \to A^*$ via

$$\rho(xs) = \rho(x)s.$$

If A is equipped with an ordering of its elements, say, $A = \{x_1, \ldots, x_k\}$, then we write

$$(g_1, \ldots, g_k)$$

for the automorphism $g : A_* \to A_*$ which acts trivially on A, and for which $g|_{x_i} = g_i$ for all i. Then every automorphism can be uniquely decomposed into

$$g = \rho\,(g_1, \ldots, g_k),$$

for some permutation $\rho : A \to A$ and some automorphisms $g_i : A^* \to A^*$. Specifically, ρ is the restriction of g to A, and $g_i = g|_{x_i}$. Then, given an invertible Mealy automaton $\mathcal{M} = (Q, A, \tau, \sigma)$, the definition of the automorphisms $\boldsymbol{\sigma}_q$ can be phrased recursively as

$$\boldsymbol{\sigma}_q = \sigma_q\,(\boldsymbol{\sigma}_{q_1}, \ldots, \boldsymbol{\sigma}_{q_k}),$$

where $q_i = q^{x_i}$. Such a recursive definition is called a *wreath recursion.* For example, if $\mathcal{B} = (Q, A, \tau, \sigma)$ is the Bellaterra automaton, then we have the wreath recursion

$$\begin{aligned}\boldsymbol{\sigma}_a &= (\boldsymbol{\sigma}_b, \boldsymbol{\sigma}_c)\\ \boldsymbol{\sigma}_b &= (\boldsymbol{\sigma}_c, \boldsymbol{\sigma}_b)\\ \boldsymbol{\sigma}_c &= \rho(\boldsymbol{\sigma}_a, \boldsymbol{\sigma}_a),\end{aligned}$$

where $\rho : \{0, 1\} \to \{0, 1\}$ swaps 0 and 1.

Definition 5.2. Let $\mathbf{T}$ be a rooted tree. We say a tree automorphism $g : \mathbf{T} \to \mathbf{T}$ is *spherically transitive* (or just *transitive*) if its restriction to every level of $\mathbf{T}$ is a transitive map.

For example, if $\mathcal{M}$ is the adding automaton pictured in Figure 4, then σ_r : $\{0, 1\}^* \to \{0, 1\}^*$ is spherically transitive, because its action on the n-th level is addition of 1 modulo 2^n.

6. The Bellaterra automaton

Consider the Bellaterra automaton $\mathcal{B} = (Q, A, \sigma, \tau)$, pictured in Figure 3. We want to show that the graphs $B_n = \Gamma_{\mathcal{B},n}$ have small diameter. Our approach is to find short words in Q^* which change only the last digit of the word $1^n = 11 \dots 1$. So, we are looking for words which do not fix the infinite word $1^\infty = 111\dots$, but do preserve the first n of its letters. It turns out there are enough of these words because $\boldsymbol{\tau}_1$ acts "transitively enough" on Q^*, so that almost every orbit under its action contains some word $w \in Q^*$ which swaps 0 and 1.

It is straightforward to check that a^2, b^2, and c^2 act trivially on A^*, (i.e. $\boldsymbol{\sigma}_{aa} = \boldsymbol{\sigma}_{bb} = \boldsymbol{\sigma}_{cc} = \mathrm{id}$) so we are primarily interested in *reduced words* in $\{a, b, c\}$, i.e. those which do not repeat the same letter twice in a row. Note that these words form a subtree of Q^*, which is nearly a binary tree: every vertex has two children, except the root.

We will need a simple result on the transitivity of automorphisms of a binary tree. $A = \{0, 1\}$. Define a group homomorphism $\chi : \mathrm{Aut}(A^*) \to \mathbb{Z}_2[[t]]$, by

$$\chi(g) = \sum_{n=1}^{\infty} c_n t^{n-1},$$

where $(-1)^{c_n}$ is the sign of the permutation given by the action of χ the n-th level of A^*. Values of this homomorphism can be computed recursively via

$$\chi(g) = c_1 + t\big(\chi(g|_x) + \chi(g|_y)\big),$$

where c_1 is 0 if g fixes the two elements of A, and $c_1 = 1$ if g swaps them. We call $\chi(g)$ the *characteristic function of g.* Of course, this definition makes sense when A is any two-element set, so we will state the lemma more generally:

Lemma 6.1. *Let $A = \{x, y\}$. An automorphism $g \in \mathrm{Aut}(A^*)$ is spherically transitive if and only if $\chi(g) = 1/(1 - t)$.*

Proof. If g is spherically transitive, then its action on the n-th level of A^n is a (2^n)-cycle, which is an odd permutation for all $n \geq 1$. Hence, $c_n = 1$ for all $n \geq 1$, and

$$\chi(g) = \sum_{n=1}^{\infty} t^{n-1} = \frac{1}{1-t}\,.$$

In the other direction, suppose $\chi(g) = 1/(1-t)$, i.e. g acts as an odd permutation on A^n for every $n \geq 1$. We will show by induction on n that the action of g on A^n is a (2^n)-cycle for all $n \geq 0$. This is trivial for $n = 0$.

For the inductive step, suppose g acts as a (2^n)-cycle on A^n. Given a word $s \in A^n$, we either have $g^{2^n}(sx) = sx$ or $g^{2^n}(sx) = sy$. In the first case sx belongs to a (2^n)-cycle of g, in the second case, sx belongs to a (2^{n+1})-cycle. So, any word in A^{n+1} ending in x belongs to either a (2^n)-cycle or a (2^{n+1})-cycle, and similarly for words ending in y. So, the action of g on A^{n+1} decomposes into either two (2^n)-cycles or a single (2^{n+1})-cycle. But the former is an even permutation, so g must act as a (2^{n+1})-cycle on A^{n+1}, as desired. □

Lemma 6.2. *Let $\mathcal{B} = (Q, A, \tau, \sigma)$ denote the Bellaterra automaton. Then for every natural number n, the map $\boldsymbol{\tau}_{\mathcal{B},1}$ acts transitively on the set of reduced words of length n ending with a or c.*

Proof. We will write the argument down in terms of the dual automaton $\overline{\mathcal{B}} = (A, Q, \overline{\tau}, \overline{\sigma})$, pictured in Figure 3. Since taking the dual reverses words, we want to show that $\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_{\overline{\mathcal{B}},1}$ acts transitively on the binary subtree $\mathbf{T} \subset Q^*$ of reduced words which *begin* with a or c.

It is convenient to put $\mathbf{T}$ into bijection with a binary tree of words $\mathbf{R} = \{\uparrow, \downarrow\}^*$. We define the maps $\phi_a, \phi_b, \phi_c : \mathbf{R} \to Q^*$ recursively by

$$\begin{aligned} \phi_a(\uparrow w) &= b\,\phi_b(w) & \qquad \phi_a(\downarrow w) &= c\,\phi_c(w) \\ \phi_b(\uparrow w) &= c\,\phi_c(w) & \qquad \phi_b(\downarrow w) &= a\,\phi_a(w) \\ \phi_c(\uparrow w) &= a\,\phi_a(w) & \qquad \phi_c(\downarrow w) &= b\,\phi_b(w) \end{aligned}$$

It is straightforward to check by induction on word length that for each $x \in Q$, ϕ_x defines a tree isomorphism between $\mathbf{R}$ and the reduced words in Q^* which do not begin with x. In particular, ϕ_b is a bijection between $\mathbf{R}$ and $\mathbf{T}$.

Now consider the dual $\overline{\mathcal{B}}$ of the Bellaterra automaton, and in particular the corresponding automorphisms $\boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1 \in \mathrm{Aut}(Q^*)$. Given $x, y \in Q$, $d \in A$, define

$$\boldsymbol{\sigma}_{x,d,y} = \phi_x^{-1} \boldsymbol{\sigma}_d \phi_y \in \mathrm{Aut}(\mathbf{R}).$$

Note that, a priori, the domain of ϕ_x^{-1} may not coincide with the image of $\boldsymbol{\sigma}_d \phi_y$, so $\boldsymbol{\sigma}_{x,d,y}$ may be ill-defined for some values of x, d, y. However, the computations below give an explicit recursion for computing $\boldsymbol{\sigma}_{1,b,1}$, which also demonstrates that it is well-defined.

We can compute that, e.g.,

$$\begin{aligned} \boldsymbol{\sigma}_{b,1,b}(\uparrow w) &= \phi_b^{-1}(\boldsymbol{\sigma}_1(\phi_b(\uparrow w))) \\ &= \phi_b^{-1}(\boldsymbol{\sigma}_1(c\,\phi_c(w))) \\ &= \phi_b^{-1}(a\,\boldsymbol{\sigma}_0(\phi_c(w))) \\ &= \downarrow \phi_a^{-1}(\boldsymbol{\sigma}_0(\phi_c(w))) \\ &= \downarrow \boldsymbol{\sigma}_{a,0,c} \end{aligned}$$

In particular, $\boldsymbol{\sigma}_{b,1,b}|_{\uparrow} = \boldsymbol{\sigma}_{a,0,c}$.

Similar computations give the complete recursive description of $\boldsymbol{\sigma}_{b,1,b}$, which we write down using the usual wreath recursion notation $g = \rho^{\varepsilon}(g|_{\uparrow}, g|_{\downarrow})$, where ρ

swaps ↑ and ↓:

$$\begin{aligned}
\boldsymbol{\sigma}_{b,1,b} &= \rho\,(\boldsymbol{\sigma}_{a,0,c}, \boldsymbol{\sigma}_{c,1,a})\\
\boldsymbol{\sigma}_{a,0,c} &= (\boldsymbol{\sigma}_{b,0,a}, \boldsymbol{\sigma}_{c,0,b})\\
\boldsymbol{\sigma}_{c,1,a} &= \rho\,(\boldsymbol{\sigma}_{b,1,b}, \boldsymbol{\sigma}_{a,0,c})\\
\boldsymbol{\sigma}_{b,0,a} &= (\boldsymbol{\sigma}_{c,0,b}, \boldsymbol{\sigma}_{a,1,c})\\
\boldsymbol{\sigma}_{c,0,b} &= (\boldsymbol{\sigma}_{a,1,c}, \boldsymbol{\sigma}_{b,0,a})\\
\boldsymbol{\sigma}_{a,1,c} &= \rho\,(\boldsymbol{\sigma}_{c,1,a}, \boldsymbol{\sigma}_{b,1,b})
\end{aligned}$$

Defining $F_{x,d,y} = \chi(\boldsymbol{\sigma}_{x,d,y})$, this gives us the following linear equations in the ring $\mathbb{Z}_2[[t]]$:

$$\begin{aligned}
F_{b,1,b} &= 1 + t(F_{a,0,c} + F_{c,1,a})\\
F_{a,0,c} &= t(F_{b,0,a} + F_{c,0,b})\\
F_{c,1,a} &= 1 + t(F_{b,1,b} + F_{a,0,c})\\
F_{b,0,a} &= t(F_{c,0,b} + F_{a,1,c})\\
F_{c,0,b} &= t(F_{a,1,c} + F_{b,0,a})\\
F_{a,1,c} &= 1 + t(F_{c,1,a} + F_{b,1,b})
\end{aligned}$$

Solving this system of equations yields

$$\begin{aligned}
F_{b,1,b} &= 1/(1-t)\\
F_{a,0,c} &= 0\\
F_{c,1,a} &= 1/(1-t)\\
F_{b,0,a} &= t/(1-t)\\
F_{c,0,b} &= t/(1-t)\\
F_{a,1,c} &= 1,
\end{aligned}$$

So we have

$$\chi(\phi_b^{-1}\boldsymbol{\sigma}_{\overline{\mathcal{B}},1}\phi_b) = 1/(1-t).$$

By Lemma 6.1, the automorphism $\phi_b^{-1}\boldsymbol{\sigma}_1\phi_b$ acts transitively on $\mathbf{R}$. Hence $\boldsymbol{\sigma}_1$ acts transitively on $\mathbf{T}$, that is, for each n it acts transitively on the set of length n reduced words in $\{a, b, c\}$ which begin with a or c. Hence, in the unreversed Bellaterra automaton $\mathcal{B}$, we have that $\boldsymbol{\tau}_{\mathcal{B},1}$ acts transitively on the words of a given length which end with a or c. □

Lemma 6.3. *Let $\mathcal{M} = (Q, A, \tau, \sigma)$ be a Mealy automaton, let x be a letter in A, and let w be a word in Q^*. Then w stabilizes the infinite word $xxx\ldots = x^\infty$ if and only if every element of the orbit of w under $\boldsymbol{\tau}_x$ stabilizes x. That is,*

$${}^{w}(xxx\ldots) = xxx\ldots \qquad \textit{if and only if} \qquad \boldsymbol{\sigma}(\boldsymbol{\tau}_x^n(w), x) = x \textit{ for all } n \geq 0.$$

Proof. Say

$${}^{w}(xxx\ldots) = y_0y_1y_2\ldots.$$

Then y_n is the last letter of ${}^{w}(x^{n+1})$. Letting $X = x^n$, we have

$${}^{w}(x^{n+1}) = {}^{w}(Xx) = {}^{w}(X)\ {}^{\tau(w,X)}x,$$

So,

$$y_n = {}^{\tau(w,X)}x = \boldsymbol{\sigma}(\tau(w,X),x) = \boldsymbol{\sigma}(\tau_X(w),x) = \boldsymbol{\sigma}(\tau_x^n(w),x),$$

and therefore ${}^{w}xxx\ldots = xxx\ldots$ if and only if $\boldsymbol{\sigma}(\boldsymbol{\tau}_x^n(w),x) = x$ for every $n \geq 0$, as desired. □

Theorem 6.4. *Let B_n denote the n-th Bellaterra graph. Then* $\operatorname{diam}(B_n) = O(n^2)$.

Proof. Let $\mathcal{B} = (Q, A, \tau, \sigma)$ denote the Bellaterra automaton, so that $B_n = \Gamma_{\mathcal{B},n}$. It is enough to show that for some C the ball of radius Cn^2 around the vertex $1^n = 11\ldots1$ covers all of B_n. That is, we will show that for every number n, and every $v \in \mathcal{B}_n$,

$$d(1^n, v) \leq Cn^2.$$

The only letter in $Q = \{a, b, c\}$ which swaps the elements of A is c. The other two letters fix 0 and 1. Hence, a word $w \in Q^*$ fixes 1 if and only if it has an even number of c's.

For each $n > 0$, there is a reduced word ending in a or c which contains an odd number of c's. We can take, e.g. *abab...abc* or *baba...abc*. By Lemma 6.2, if w is any reduced word of length n ending in a or c, then its orbit under $\boldsymbol{\tau}_1$ contains some word which does not fix 1. Hence, by Lemma 6.3, w does not fix the infinite word $111\ldots = x^\infty$.

Given a number $n \geq 1$, we have $|A^n| = 2^n$, and there are $2^{n+1}-1$ reduced words of length n or less which end in a or c. By the pigeonhole principle, there must be two such words, v, w with ${}^{v}(1^n) = {}^{w}(1^n)$. We may assume $\ell(v) \leq \ell(w)$. Since a^2, b^2, and c^2 all act trivially on A^*, reversing a word inverts its action on A^*. Let u be the reduced word formed by cancelling pairs of repeated letters in $\overline{v}w$. Then,

$$\begin{aligned} &{}^{u}(1^n) = {}^{\overline{v}w}(1^n) = {}^{\overline{v}v}(1^n) = 1^n, \\ \text{and}\quad &\ell(u) \leq \ell(\overline{v}w) = \ell(v) + \ell(w) \leq 2n. \end{aligned}$$

Since $v \neq w$, u is not the empty word. We assumed that $\ell(v) \leq \ell(w)$, so the last letter of w is not cancelled. Hence u also ends in in a or c, and therefore

$${}^{u}(111\ldots) \neq 111\ldots.$$

Let k be the maximal integer such that ${}^{u}(1^k) = 1^k$. We know $k \geq n$ and ${}^{u}(1^{k+1}) = 1^k0$. So, letting $s = 1^{k-n}$, $t = 1^{n+1}$, and $t' = 1^n0$, we have

$$st' = {}^{u}(st) = {}^{u}s\ {}^{u'}(t) = s\ {}^{u'}(t),$$

where $u' = u^s$. So we have

$$\begin{aligned} &{}^{u'}(1^{n+1}) = 1^n0, \\ \text{and}\quad &\ell(u') = \ell(u) \leq 2n. \end{aligned}$$

This construction works for all $n \geq 1$. That is, for every $n \geq 1$, there exists a $u_n \in Q^*$ with $\ell(u_n) \leq 2n$ and ${}^{u_n}(1^n0) = 1^{n+1}$.

We now prove by induction on n that for every $s \in A^n$, there is a $w \in Q^*$ with $\ell(w) \leq n^2$ such that ${}^{w}s = 1^n$. The base cases $n = 0$ and $n = 1$ are trivial. For the inductive step, consider any $n \geq 1$. Given $s \in A^{n+1}$, let s' be s with the last digit removed. By the induction hypothesis know there is a word w with $\ell(w) \leq n^2$ such that ${}^{w}s' = 1^n$. Then either ${}^{w}s = 1^{n+1}$ or ${}^{w}s = 1^n0$. In the first case, we are done. In the second case, ${}^{u_n w}s = 1^{n+1}$, and $\ell(u_n w) \leq 2n + n^2 \leq (n+1)^2$, so we are done.

So, we have shown that in the graph $B_n = \Gamma_{\mathcal{B},n}$, we have $d(1^n, s) \leq n^2$ for every $s \in A^n$. It follows that for any $s, t \in B_n$,

$$d(s,t) \leq d(s,1^n) + d(1^n,t) \leq 2n^2,$$

i.e. $\operatorname{diam}(B_n) \leq 2n^2$. □

7. The Aleshin automaton

The Aleshin automaton $\mathcal{A}$ and the Bellaterra automaton $\mathcal{B}$ are closely related. Indeed, let $\tau_d : \{0,1\}^* \to \{0,1\}^*$ denote map which swaps every digit of a binary word. Then it is straightforward to check by induction that

$$\begin{aligned} \boldsymbol{\tau}_{\mathcal{A},a} &= \boldsymbol{\tau}_d\, \boldsymbol{\tau}_{\mathcal{B},a}, \\ \boldsymbol{\tau}_{\mathcal{A},b} &= \boldsymbol{\tau}_d\, \boldsymbol{\tau}_{\mathcal{B},c}, \\ \text{and} \quad \boldsymbol{\tau}_{\mathcal{A},c} &= \boldsymbol{\tau}_d\, \boldsymbol{\tau}_{\mathcal{B},c}. \end{aligned}$$

With this observation, Theorem 6.4 has the following corollary.

Corollary 7.1. *Let A_n denote the n-th Aleshin graph. Then* $\operatorname{diam}(A_n) = O(n^2)$.

Proof. For every pair $q, r \in \{a, b, c\}$, we have

$$\boldsymbol{\tau}_{\mathcal{A},q}^{-1}\, \boldsymbol{\tau}_{\mathcal{A},r} = \boldsymbol{\tau}_{\mathcal{B},q}^{-1}\, \boldsymbol{\tau}_d^{-1}\, \boldsymbol{\tau}_d\, \boldsymbol{\tau}_{\mathcal{B},r} = \boldsymbol{\tau}_{\mathcal{B},q} \boldsymbol{\tau}_{\mathcal{B},r}.$$

So, if two words in $\{0,1\}^n$ are separated by a path of length 2 in the Bellaterra graph B_n, they are also separated by a path of length 2 in the Aleshin graph A_n. It follows that two endpoints of an even-length path in B_n are endpoints of a path in A_n of the same length.

For any word $s \in \{0,1\}^n$ there is a path in B_n of length $O(n^2)$ from 1^n to s. We may assume that this path has even length since 1^n has an edge in B_n from itself to itself. This corresponds to a path in A_n of the same length, so for any $s \in \{0,1\}^n$, there is a path in A_n of length $O(n^2)$ from 1^n to s. Therefore, $\operatorname{diam}(\Gamma_{\mathcal{A},n}) = O(n^2)$, as desired. □

8. Generalizations

The proof of Theorem 6.4 can be adapted to prove a more general result. In order to generalize to automata with larger alphabets, we need to consider a restricted type of automaton. We say an Mealy automaton $\mathcal{M} = (Q, A, \tau, \sigma)$ is *cyclic* if it is invertible, and $\langle \sigma_q \mid q \in Q \rangle = \langle (x_1\, x_2\, \ldots\, x_n) \rangle$, where $\{x_1, x_2, \ldots, x_n\} = A$. That is, if its action on A is a cyclic permutation group. In particular, any automaton with $|A| = 2$ is cyclic. This will enable us to reach any word of the form $x^n y$ from x^{n+1} in a short time, as long as we can reach some such word.

We first state and prove the general result with the weakest assumptions under which our argument guarantees polynomial growth of $\operatorname{diam}(\Gamma_{\mathcal{M},n})$.

Theorem 8.1. *Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be a cyclic Mealy automaton with $|A|$ prime, and let $\Gamma = \Gamma_{\mathcal{M},\infty}$. Suppose there is a letter $x \in A$ and constants $\alpha > 0$, $K > 1$ such that, for sufficiently large r,*

$$|B_\Gamma(xxx\ldots, r)| \geq K^{r^\alpha}.$$

Then there is a constant $C > 0$, such that for all n,

$$\mathrm{diam}(\Gamma_{\mathcal{M},n}) \leq Cn^{1+1/\alpha}.$$

Proof. Let $p = |A|$. By replacing $\mathcal{M}$ with $\mathcal{M} \cup \mathcal{M}^{-1}$ if necessary, we may assume that for every $q \in Q$, there is a $q' \in Q$ with $\boldsymbol{\sigma}_{q'} = \boldsymbol{\sigma}_q^{-1}$. This replacement adds edges to the Schreier graphs $\Gamma_{\mathcal{M},n}$, but only reverses of edges which were already there, so $\mathrm{diam}(\Gamma_{\mathcal{M},n})$ and $B_{\Gamma_{\mathcal{M},n}}(s,r)$ are unaffected. Then for a word $w \in Q^*$, we define w^{-1} to be w, reversed, with each letter q replaced by q', so that $\boldsymbol{\sigma}_{w^{-1}} = \boldsymbol{\sigma}_w^{-1}$.

Given sufficiently large n, pick r such that

$$((\log_K p)n)^{1/\alpha} < r < 2((\log_K p)n)^{1/\alpha}.$$

Then $|B_\Gamma(xxx\ldots, r)| > p^n$. By the pigeonhole principle, some two elements of $B_\Gamma(xxx\ldots, r)$ have the same first n digits. That is, there are $v, w \in (Q \cup Q^{-1})^*$ with

$$\ell(v), \ell(w) \leq r, \quad {}^{v}(x^n) = {}^{w}(x^n), \quad \text{and} \quad {}^{v}(xxx\ldots) \neq {}^{w}(xxx\ldots).$$

So, there is a $u_0 = v^{-1}w \in (Q \cup Q^{-1})^*$ with

$$\begin{aligned} \ell(u_0) &\leq 2r, \\ {}^{u_0}(x^n) &= x^n, \\ \text{and} \quad {}^{u_0}(xxx\ldots) &\neq xxx\ldots. \end{aligned}$$

There is some smallest value of $k \geq n+1$ such that ${}^{u_0}(x^k) \neq x^k$. Let $X_0 = x^{k-n-1}$ and $X = x^n$, so that $x^k = X_0 X x$ and ${}^{u_0}(x^k) = X_0 X y$ for some $y \in A$ with $y \neq x$. Let $u = u_0^{X_0}$ so in particular, $\ell(u) = \ell(u_0)$. Then,

$$X_0 X y = {}^{u_0}(X_0 X x) = {}^{u_0}X_0 \ {}^{u}(Xx) = X_0 \ {}^{u}(Xx),$$

so

$${}^{u}(Xx) = Xy.$$

Similarly, if $u' = u^X$, we have

$${}^{u}(Xz) = X \ {}^{u'}z$$

for any $z \in A$. Since ${}^{u'}x = y \neq x$ and $\mathcal{M}$ is cyclic, the action of u' on A is a nontrivial cyclic permutation. Since $p = |A|$ is prime, u' acts transitively on A, and therefore u acts transitively on $\{Xz \mid z \in Aa\}$. It follows that for any $z, z' \in A$,

$$d(Xz, Xz') \leq p\,\ell(u) \leq 2pr \leq 4p((\log_K p)n)^{1/\alpha}.$$

Thus, there is a constant C such that for sufficiently large n, we have

$$d(x^n z, x^{n+1}) \leq Cn^{1/\alpha}, \quad \text{for all } b \in A.$$

By increasing the constant if necessary, we can make this true for all n.

Now let us show by induction on n that for all $s \in A^n$, we have $d(s, x^n) < Cn^{1+1/\alpha}$. The base case $n = 0$ is trivial. For the inductive step, take any $s \in A^{n+1}$, and let s' be its first n letters. We know $d(s', x^n) < Cn^{1+1/\alpha}$. There is some word

$w \in (Q \cup Q^{-1})^*$ with $\ell(w) = d(s', x^n)$ and ${}^w(s') = x^n$. Then ${}^w s = x^n z$ for some $z \in B$. Thus,

$$\begin{aligned} d(s, x^{n+1}) &\leq d(s, x^n z) + d(x^n z, x^{n+1}) \\ &\leq Cn^{1+1/\alpha} + Cn^{1/\alpha} \\ &\leq C(n+1)^{1+1/\alpha}, \end{aligned}$$

which completes the induction.

It follows that for any $s, t \in A^n$, $d(s,t) \leq d(s, x^n) + d(x^n, t) \leq 2Cn^{1+1/\alpha}$, i.e.,

$$\operatorname{diam}(\Gamma_{\mathcal{M},n}) \leq 2Cn^{1+1/\alpha}.$$

□

In all the cases where we apply this, $|B_\Gamma(xxx\ldots, r)|$ will have exponential growth, so we state that case separately.

Corollary 8.2. *Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be a cyclic Mealy automaton with $|A|$ prime. Let $\Gamma = \Gamma_{\mathcal{M},\infty}$. If there is an $x \in A$ and a constant $K > 1$ such that*

$$|B_\Gamma(xxx\ldots, r)| \geq K^r$$

for sufficiently large r, then there is a constant $C > 0$, such that

$$\operatorname{diam}(\Gamma_{\mathcal{M},n}) \leq Cn^2.$$

It is not always easy to guarantee that $|B_\Gamma(xxx\ldots, r)| \geq K^r$ grows quickly, so we prove an additional result based on the size of orbits of $\boldsymbol{\tau}_x$ in Q^n. Loosely, if the orbits grow quickly enough, it must be because there are enough distinct images of words of the form x^m.

Theorem 8.3. *Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be a reversible*[3] *cyclic Mealy automaton with $|A|$ prime. Suppose there is a letter $x \in A$ and constants $K > 1$, $\alpha > 0$ such that for sufficiently large n, there is a $w \in Q^n$ with*

$$\left|\{\boldsymbol{\tau}_x^k(w) \mid k \in \mathbb{Z}\}\right| \geq K^{n^\alpha}.$$

Then there is a constant $C > 0$, such that for all n, we have

$$\operatorname{diam}(\Gamma_{\mathcal{M},n}) \;\leq\; C\, n^{1+1/\alpha}.$$

Proof. Let $P = \{p \in \mathbb{N} \mid p \text{ prime}, p \leq |Q|\}$. It is easy to see by induction on length that for each $w \in Q^*$, the sequence $w, \tau_x(w), \tau_x^2(w), \ldots$ is periodic with period m, where m is a product of some powers of primes in P. Define letters $q_{i,k}$ via

$$\tau_x^k(w) = q_{0,k} q_{1,k} \ldots q_{n,k}.$$

If m_i is the period of the sequence $q_{i,0}, q_{i,1}, \ldots$, then $m = \gcd(m_0, m_1, \ldots, m_n)$. Let $M = \max_i m_i$. Then each prime power in the prime factorization of m is a factor of some m_i, so it is at most M. The period m is the product of these prime powers, so $m \leq M^{|P|}$. That is, there is some i such that $m_i \geq m^{1/|P|}$.

Fix that i for the rest of the proof, and let v be the first i letters of w. Consider the infinite word $s = {}^v(xxx\ldots)$, and let x_k be it's k-th letter. Let l be the period of the word s. Note that $q_{i,k+1} = q_{i,k}^{x_k}$ and therefore

$$q_{i,k+l} = q_{i,k}^X,$$

[3] The assumption that $\mathcal{M}$ is reversible may be lifted, if we replace $\left|\{\boldsymbol{\tau}_x^k(w) \mid k \in \mathbb{Z}\}\right|$ with the length of the (eventual) period of $w, \tau_x(w), \tau_x^2(w), \ldots$.

where

$$X = x_k x_{k+1} \dots x_{k+l-1}.$$

Since the x_k repeat every l letters, we have $q_{i,k+l} = q_{i,k}^X$, and $q_{i,k+2l} = q_{i,k+l}^X$, and so on. Let $F = |Q|!$. Then X^F acts trivially on Q, and hence $q_{i,k+Fl} = q_{i,k+Fl}$. This is true for each k, so the $q_{i,k}$ have period $m_i \leq Fl$. Thus, the word $s = {}^v(xxx\dots)$ has period $l \geq m^{1/|P|}/F$.

Now let n be sufficiently large, so that there is a word $w \in Q^n$ whose orbit under $\boldsymbol{\tau}_x$ has size $m \geq K^{n^\alpha}$. Then, from the above, for some $v \in Q^*$ with $\ell(v) \leq n$, the word $s = {}^v xxx\dots$ has period

$$l \geq \frac{1}{F} K^{n^\alpha/|P|} \geq \widetilde{K}^{n^\alpha},$$

where we fix some $1 < \widetilde{K} < K^{1/|P|}$, and the last inequality holds for sufficiently large n.

Let $v_k = \tau_x^k(v)$. Then ${}^{v_k}xxx\dots$ is a shift of s, and since s has period l there are l distinct such shifts. So, since each v_k satisfies $\ell(v_k) \leq n$, the set $\{{}^w xxx\cdots \mid w \in Q^*, \ell(w) \leq n\}$ has at least $l \geq \widetilde{K}^{n^\alpha}$ elements. It follows that $|B_\Gamma(xxx\dots, n)| \geq \widetilde{K}^{n^\alpha}$, where $\Gamma = \Gamma_{\mathcal{M},\infty}$

So Theorem 8.1 applies, and there is a constant C such that $\operatorname{diam}(\Gamma_{\mathcal{M},n}) \leq Cn^{1+1/\alpha}$. □

We also state the following special case, which is a simple way to apply the theorem.

Corollary 8.4. *Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be a reversible cyclic Mealy automaton with $|A|$ prime. Suppose there is some $a \in A$, and some $d \geq 2$, such that there is a d-regular subtree $\mathbf{T} \subseteq Q^*$ such that $\boldsymbol{\tau}_a$ acts spherically transitively on $\mathbf{T}$. Then there is a constant $C > 0$, such that for all n, we have*

$$\operatorname{diam}(\Gamma_{\mathcal{M},n}) \leq Cn^2.$$

9. Cotransitive cyclic automata

The simplest way for the conditions in Corollary 8.4 to be satisfied is when some $\boldsymbol{\tau}_a$ acts spherically transitively on the entire tree Q^*. With that in mind, we make the following definitions.

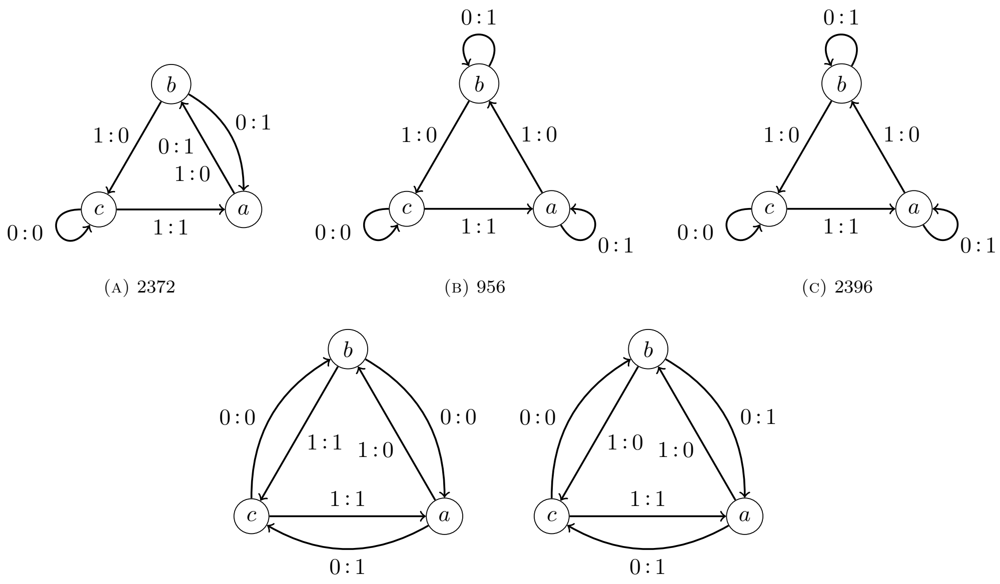
We say an invertible Mealy automaton $\mathcal{M} = (Q, A, \sigma, \tau)$ is *q-transitive* if the tree automorphism $\boldsymbol{\sigma}_q : A^* \to A^*$ is spherically transitive. We say $\mathcal{M}$ is *transitive*[4] if it is q-transitive for some $q \in Q$. We say $\mathcal{M}$ is *cotransitive* if its dual is transitive.

Then, according to Corollary 8.4, we have

Corollary 9.1. *Let $\mathcal{M} = (Q, A, \sigma, \tau)$ be a reversible cyclic cotransitive Mealy automaton with $|A|$ prime. Then* $\operatorname{diam}(\Gamma_{\mathcal{M},n}) = O(n^2)$.

We do not know a general method for determining whether a tree automorphism given by an automaton is transitive, but there are special cases where checking it is easier. For example, [St] gives a generalization of Lemma 6.1 to all cyclic automata:

[4] A more natural definition of this term might be that the $\boldsymbol{\sigma}_q$ together act transitively on each level of A^*, but that is too general for our purposes



(A) 2372 (B) 956 (C) 2396

(D) 870 (E) 2294

Figure 6. The five contransitive 3-state automata on a binary alphabet.

Lemma 9.2. *Let $\mathcal{M} = (Q, A, \tau, \sigma)$ be a cyclic automaton, with $|A| = m$. Then there is a cyclic permutation ρ of A, such that for each $q \in Q$ there is a k_q s.t. $\sigma_q = \rho^k$. Recursively define*

$$\chi(q) = k_q + t \sum_{x \in A} \chi(\tau_x(q)) \in \mathbb{Z}_m[[t]]$$

Then $\boldsymbol{\sigma}_q$ acts transitively on A^* if and only if each coefficient of $\chi(q)$ is a generator of $\mathbb{Z}_m$.

An automaton is called *cocyclic* if its dual is cyclic. Now observe that the power series $\chi(q)$ for $q \in Q$ satisfy a recursive linear relation, which can be solved to write each $\chi(q)$ as a rational function. This implies:

Corollary 9.3. *Given a (co)cyclic Mealy automaton $\mathcal{M} = (Q, A, \tau, \sigma)$, there is an algorithm to determine whether it is (co)transitive.*

For example, it is straightforward to check that, there are 16 cocyclic invertible $(3, 2)$-automata, and only four are cotransitive. These four are the automata pictured in Figures 6b–6e, i.e., automata number 956, 2396, 870, and 2294 in [B+].[5]

[5] Note that [B+] does not distinguish between an automaton and its inverse. We do, so some of our automata are actually inverses of the automata described there.

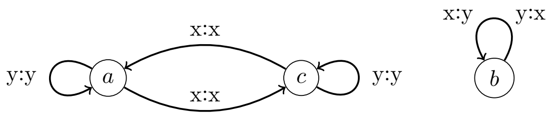


FIGURE 7. An automaton to conjugate $\mathcal{M}_{2372}$ into a cocyclic automaton.

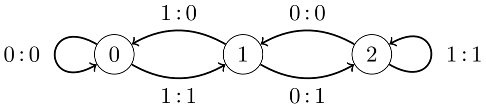


FIGURE 8

10. Further examples

Example 10.1. It can be verified that, except for the automata pictured in Figure 6, every invertible $(3, 2)$-automaton is not cotransitive: for each automaton, simply find orbits of each $\boldsymbol{\tau}_a$ which are proper subsets of A^n for some n. In this case, it suffices to take $n = 4$.

The final automaton $\mathcal{M}$ pictured in Figure 6a, which is automaton number 2372 in [B+], is not cocyclic. So, we do not have a mechanical procedure to prove it is cotransitive. It turns out, however, that there is an automorphism $\kappa : \{a, b, c\}^* \to \{a, b, c\}^*$ such that $\kappa^{-1}\boldsymbol{\tau}_{\mathcal{M},1}\kappa$ can be computed by a cyclic automaton. Indeed, one can take $\kappa = \boldsymbol{\tau}_{\mathcal{C},x}$, where $\mathcal{C}$ is the automaton in Figure 7. Then we can compute the power series $\chi(\kappa^{-1}\boldsymbol{\tau}_{\mathcal{M},1}\kappa)$, and see directly that its coefficients are nonzero. At that point, Corollary 9.2 implies that $\kappa^{-1}\boldsymbol{\tau}_{\mathcal{M},1}\kappa$ acts transitively on Q^*, and hence so does $\boldsymbol{\tau}_{\mathcal{M},1}$.

So, we have sketched a proof of the following:

Proposition 10.2. *The cotransitive invertible* $(3, 2)$*-automata are precisely the five automata pictured in Figure 6, up to relabeling of* A *and* Q*.*

Example 10.3. Of course, there are automata which are not cotransitive, but still satisfy the conditions of Corollary 8.4. As we saw, one example is the Bellaterra automaton. A natural and easy to analyze example is the automaton $\mathcal{M} = (Q, A, \tau, \sigma)$ that implements division by 3 modulo 2^n. (This is automaton number 924 in [B+]. See [BŠ] for more on this construction and related ones.) We will also see that its Schreier graphs do not form a family of expanders.

A quick way to define this automaton is that for $a, b \in Q = \{0, 1, 2\}$ and $x, y \in A = \{0, 1\}$, we have $a \xrightarrow{x\,:\,y} b$ if and only if

$$a + 3y = x + 2b.$$

This automaton is pictured in Figure 8. Note that for convenience we abuse notation slightly and call two of the states, 0 and 1, by the same name as the letters in the alphabet.

By assumption, if $a \xrightarrow{x:y} b$, then for any $x' \in \mathbb{Z}/2^{n-1}\mathbb{Z}$, we have the following equalities in $\mathbb{Z}/2^n\mathbb{Z}$:

$$\begin{aligned} x + 2b &= a + 3y \\ x + 2x' - a &= 3y + 2x' - 2b \\ \frac{(x + 2x') - a}{3} &= y + 2\frac{x' - b}{3}. \end{aligned}$$

That is, if x is the least significant binary digit of a number $X \in \mathbb{Z}/2^n\mathbb{Z}$, and $x' \in \mathbb{Z}/2^{n-1}\mathbb{Z}$ is the number corresponding to the rest of its digits, then the least significant digit of $(X - a)/3$ is y, and the rest of the digits are given by $(x' - b)/3$. It follows that if we identify a number $x \in \mathbb{Z}/2^n\mathbb{Z}$ with its binary representation in $\{0, 1\}^n$ (with the least significant digit on the left), then we have, for each $a \in \{0, 1, 2\}$,

$$\boldsymbol{\sigma}_a(x) = \frac{x - a}{3}.$$

By a symmetric argument, the dual of this automaton implements division by 2 modulo 3. Phrasing this in terms of the original automaton $\mathcal{M}$, we interpret a length-m word in $\{0, 1, 2\}$ as the representation of a number modulo 3^m written in ternary with the least significant digit on the right. Then for each $x \in \{0, 1\}$,

$$\boldsymbol{\tau}_x(a) = \frac{a - x}{2}.$$

In particular, $\boldsymbol{\tau}_0$ divides a number by 2. Since 2 generates the multiplicative group $(\mathbb{Z}/3^m\mathbb{Z})^*$, that group is an orbit of $\boldsymbol{\tau}_0$. So for every m, there is an orbit of $\boldsymbol{\tau}_0$ in Q^m of size $2 \cdot 3^{m-1}$. By Theorem 8.3, it follows that $\operatorname{diam}(\Gamma_{\mathcal{M},n}) = O(n^2)$. In fact, it can be checked explicitly that $\operatorname{diam}(\Gamma_{\mathcal{M},n}) = O(n)$. This can be seen from the observation that the sequence of applications of $\boldsymbol{\sigma}_1$, $\boldsymbol{\sigma}_2$, and $\boldsymbol{\sigma}_3$ necessary to send the binary number x to $00\ldots0$ is essentially the representation of x in base 3.

However, the group $G_{\mathcal{M}} = \langle \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2 \rangle$ is generated by $\mu = \boldsymbol{\sigma}_0^{-1}$ and $\alpha = \boldsymbol{\sigma}_1^{-1}\boldsymbol{\sigma}_0$, which are multiplication by 3 and addition of 1, respectively. E.g., $\boldsymbol{\sigma}_2 = \mu^{-1}\alpha^{-2}$. It follows that the group action factors through the group of upper-triangular 2 by 2 matrices via

$$\mu \mapsto \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \qquad \alpha \mapsto \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

This group is solvable, and therefore amenable. It follows that its Schreier graphs with respect to a fixed set of generators cannot be expanders [Lub, 3.3.7]. So, the family $\{\Gamma_{\mathcal{M},n}\}_{n=1}^{\infty}$ is not a family of expanders.

So, there are automata to which our general results apply, but whose Schreier graphs do not form a family of expanders. More work is necessary to find sufficient conditions for when an automaton gives rise to a family of expanders.

Example 10.4. It can be checked by a computation that there are no cotransitive invertible $(4, 2)$-automata. It turns out it is enough to check the actions of the τ_a on Q^4.

Example 10.5. There are seven $(5, 2)$-automata which are not cocyclic, but act transitively on Q^{10}. Of these, just one is bireversible, as the Aleshin and Bellaterra automata are. It is pictured in Figure 9. Unlike the automaton in Figure 6a, it is

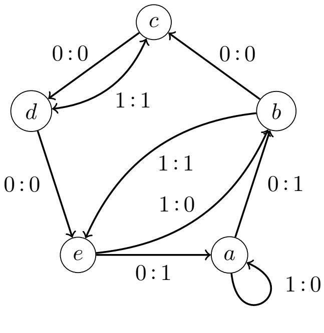


Figure 9. The only candidate to be a cotransitive bireversible $(5, 2)$-automaton.

unlikely that there is an automatic automorphism $\kappa : Q^* \to Q^*$ such that $\kappa^{-1}\tau_0\kappa$ is implemented by a cocyclic automaton. We have checked that if there is such a κ, the automaton implementing it would need to have at least 48668 states.

11. Remarks and further work

1. The results in Section 8 can be extended to non-invertible Mealy automata as well. Since we are primarily interested only in regular graphs, we prove only the simpler case. For example, to state Theorem 8.1 more generally, one needs to consider the size of balls in $\Gamma_{\mathcal{M},\infty}$ defined in terms of directed paths, but the result about the diameter of $\Gamma_{\mathcal{M},n}$ still needs diameter to be defined in terms of undirected paths.

2. As noted in the proof of Corollary 7.1, any product of two generators of the Bellaterra group $G_{\mathcal{B}} = \langle \boldsymbol{\sigma}_{\mathcal{B},a}, \boldsymbol{\sigma}_{\mathcal{B},b}, \boldsymbol{\sigma}_{\mathcal{B},c} \rangle$ belongs to the Aleshin group $G_{\mathcal{A}} = \langle \boldsymbol{\sigma}_{\mathcal{A},a}, \boldsymbol{\sigma}_{\mathcal{A},b}, \boldsymbol{\sigma}_{\mathcal{A},c} \rangle$. We used this fact to show that, since the Bellaterra graphs have small diameter, so do the Aleshin graphs. In fact, it can also be used to show that if the Bellaterra graphs form a (two-sided) expander family, so do the Aleshin graphs. In other words, Conjecture 1.3 implies Conjecture 1.2.

3. A tree automorphism $\alpha : A^* \to A^*$ is spherically transitive if and only if it is conjugate in $\mathrm{Aut}(A^*)$ to the adding machine ρ, i.e., the automorphism which interprets a word in A^n as the base-$|A|$ representation of a number modulo $|A|^n$, and adds one to that number. The adding machine is an automatic automorphism, e.g., the binary adding machine is pictured in Figure 4.

One might hope that whenever an automatic automorphism $\alpha \in \mathrm{FAut}(A^*)$ is conjugate to ρ in $\mathrm{Aut}(A^*)$, it is also conjugate to ρ in $\mathrm{FAut}(A^*)$. If so, we would have an algorithm for determining whether a given automatic automorphism is transitive. In fact, since we can enumerate the transitive cyclic automata, it would be enough if every transitive $\alpha \in \mathrm{FAut}(A^*)$ were conjugate in $\mathrm{FAut}(A^*)$ to some cyclic automorphism.

However, Example 10.5 suggests that, in the dual of the automaton in Figure 9, $\boldsymbol{\sigma}_0$ is transitive but not conjugate in $\mathrm{FAut}(A^*)$ to any cyclic automaton, in particular to ρ. However, we prove that $\boldsymbol{\sigma}_0$ is not conjugate to a cyclic automaton, nor prove that it is actually transitive.

Problem 11.1. *Exhibit a transitive $\alpha \in \mathrm{FAut}(A^*)$ which is not conjugate (in $\mathrm{FAut}(A^*)$) to a cyclic $\beta \in \mathrm{FAut}(A^*)$. (Or prove that there is no such α.)*

Problem 11.2. *Characterize the automorphisms in $\mathrm{FAut}(A^*)$ which are conjugate in $\mathrm{FAut}(A^*)$ to a cyclic automorphism.*

We can, however, exhibit a cyclic $\alpha \in \mathrm{FAut}(A^*)$ which is not conjugate in $\mathrm{FAut}(A^*)$ to the adding machine ρ:

Proposition 11.3. *Let $\mathcal{M} = (Q, A, \tau, \sigma)$ be the dual of the automaton in Figure 6b, where $Q = \{0, 1\}$ and $A = \{a, b, c\}$. Then $\boldsymbol{\sigma}_1$ acts transitively on A^*, but there is no $\kappa \in \mathrm{FAut}(A^*)$ such that $\kappa^{-1}\boldsymbol{\sigma}_1\kappa = \rho$*

Sketch of proof: Given an eventually periodic word $w \in A^*$, we let $h(w)$ denote the smallest number n such that w is periodic after the first n letters.

Note that if ρ is the adding machine, then for any eventually periodic word $v \in A^\infty$, we have

$$h(\rho^n(v)) = O(\log n).$$

Moreover, after a finite number of steps, the periodic part of $\rho^n(v)$ stabilizes. It follows that for any $\kappa \in \mathrm{FAut}(A^*)$, we have

$$h(\kappa\rho^n(v)) = O(\log n)$$

and since this applies to any v,

$$h(\kappa\rho^n\kappa^{-1}(v)) = O(\log n)$$

On the other hand, taking $\alpha = \boldsymbol{\sigma}_1$, we can check that if we read $w \in A^n$ as a ternary number modulo 3^n (with $c = 0, a = 1, b = 2$), we have

$$\alpha(w) = \frac{w+1}{-2}.$$

It follows that for $w = ccc\ldots$,

$$h(\alpha^{-n}(w)) \sim (\log_3 2)n$$

Thus α^{-1} and ρ are not conjugate in $\mathrm{FAut}(A^*)$. It is easy to check that ρ and ρ^{-1} are conjugate, so α and ρ are not conjugate in $\mathrm{FAut}(A^*)$. □

4. More generally, an open problem is the classification of conjugacy classes in $\mathrm{FAut}(A^*)$. The conjugacy classes of $\mathrm{Aut}(A^*)$ can be described in terms of orbit trees [GNS]. This tree captures the information about the orbits of an automorphism $\alpha \in \mathrm{Aut}(A^*)$, e.g. a ray with few branches in the orbit tree corresponds to a sequence of quickly growing orbits. Information about this tree can tell us whether we can apply, e.g., Theorem 8.3.

Of course, not all orbit trees arise from elements of $\mathrm{FAut}(A^*)$, since there are uncountably many. Moreover, not all automatic automorphisms with the same orbit tree are conjugate in $\mathrm{FAut}(A^*)$, as seen in Proposition 11.3.

In [BBSZ], the problem is solved for *bounded automorphisms*, and more generally automorphisms with *finite orbit-signalizer*. Such “small” automorphisms are unlikely to give expanders, so we are interested in the other end of the spectrum, automorphisms with many nontrivial sections on every level.

5. *Automaton groups*, i.e., groups of the form $G_{\mathcal{M}}$ for some Mealy automaton $\mathcal{M}$, are of independent interest in group theory.[6] A famous example is the *Grigorchuk group*, which is the first known group whose growth function is intermediate between polynomial and exponential (see [GP, G1]). For more on automaton groups, see [BGŠ, GNS, Nek].

6. The structure of $G_{\mathcal{M}}$ as an abstract group can give us information on whether or not the graphs $\Gamma_{\mathcal{M},n}$ form a family of expanders. For example, if $G_{\mathcal{M}}$ is amenable then $\{\Gamma_{\mathcal{M},n}\}_{n=1}^{\infty}$ is not a family of expanders [Lub, 3.3.7]. We already used this fact in Example 10.3 to show that the Schreier graphs of the automaton defined there are not expanders.

7. On the other hand, sometimes the structure of $G_{\mathcal{M}}$ is enough to guarantee that $\{\Gamma_{\mathcal{M},n}\}_{n=1}^{\infty}$ is a family of expanders. Notably, if $\Gamma_1, \Gamma_2, \ldots$ are Schreier graphs (with respect to a fixed generating set) of a group with *Kazhdan property* (T), and $|\Gamma_i| \to \infty$, then these graphs must form a family of expanders [Lub, 3.3.4]. This fact was used by Margulis to give the first explicit construction of expanders [Mar].

In [GM], it was shown that there are Mealy automata $\mathcal{M}$ for which $G_{\mathcal{M}}$ has property (T), so Mealy automata can be used to construct expander families. The groups $G_{\mathcal{A}}$ and $G_{\mathcal{B}}$ do not have property (T), so this approach is not sufficient to prove Conjectures 1.2 and 1.3.

8. In a recent preprint, [MSS], the ideas of [BL] were extended to construct families of bipartite Ramanujan graphs (i.e., expander graphs with optimal spectral gap) of arbitrary degree. The construction uses a new technique to pick a particular 2-lift of a graph which does not introduce any new large eigenvalues. We should note that this construction is not *very explicit*, in the sense given above.

9. In [G2, Section 10], Grigorchuk shows that in a certain formal sense, the Aleshin and Bellaterra automata are examples of *asymptotic expanders*, thus giving further evidence to Conjectures 1.2 and 1.3. He also states these conjectures as open problems, and suggests that a sequence of Schreier graphs constructed by a finite automaton cannot be Ramanujan.

Acknowledgements We are grateful to Slava Grigorchuk, Volodia Nekrashevych, Yehuda Shalom and Terry Tao for helpful discussions. The second author was partially supported by the NSF.

References

[A] V. Aleshin, A free group of finite automata (in Russian), *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* **4** (1983), 12–14.

[BGŠ] L. Bartholdi, R. Grigorchuk and Z. Šunić, Branch groups, in *Handbook of algebra*, Vol. 3, North-Holland, Amsterdam, 2003, 989–1112.

[BŠ] L. Bartholdi and Z. Šunić, Some solvable automaton groups, *Contemp. Math.* **394**, AMS, Providence, RI, 2006.

[BL] Y. Bilu and N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* **26** (2006), 495–519.

[6] Note that the term *automatic group* has a different meaning in the literature, one we do not use in this paper.

[BBSZ] I. Bondarenko, N. Bondarenko, S. Sidki and F. Zapata, On the conjugacy problem for finite-state automorphisms of regular rooted trees, *Groups Geom. Dyn.* **7** (2013), 323–355.

[B+] I. Bondarenko, R. Grigorchuk, R. Kravchenko, Y. Muntyan, V. Nekrashevych, D. Savchuk and Z. Šunić, On classification of groups generated by 3-state automata over a 2-letter alphabet, *Algebra Discrete Math.* 2008, 1–163.

[GNS] P. Gawron, V. Nekrashevych and V. Sushchansky, Conjugation in tree automorphism groups. *Internat. J. Algebra Comput.* **11** (2001), 529–547.

[GM] Y. Glasner and S. Mozes, Automata and square complexes, *Geom. Dedicata* **111** (2005), 43–64.

[G1] R. I. Grigorchuk, Solved and unsolved problems around one group, in *Infinite Groups: Geometric, Combinatorial and Dynamical Aspects*, Birkhäuser, Basel, 2005, 117–218.

[G2] R. I. Grigorchuk, Some problems of the dynamics of group actions on rooted trees, *Proc. Steklov Inst. Math.* **273** (2011), 64–175.

[GNS] R. Grigorchuk, V. Nekrashevich and V. Sushchanskiĭ, Automata, dynamical systems, and groups, *Proc. Steklov Inst. Math.* **231** (2000), 128–203.

[GP] R. I. Grigorchuk and I. Pak, Groups of intermediate growth, an introduction, *Enseign. Math.* **54** (2008), 251–272.

[HLW] S. Hoory, N. Linial and A. Wigderson, Expander graphs and their applications, *Bull. AMS* **43** (2006), 439–561.

[Lub] A. Lubotzky, *Discrete Groups, Expanding Graphs, and Invariant Measures,* Birkhäuser, Basel, 1994.

[Mar] G. Margulis, Explicit constructions of expanders, *Problemy Peredači Informacii* **9** (1973), 71–80.

[MSS] A. Marcus, D. Spielman and N. Srivastava, Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees, `arXiv:1304.4132`.

[Nek] V. Nekrashevych, *Self-similar Groups,* AMS, Providence, RI, 2005.

[St] B. Steinberg, Testing spherical transitivity in iterated wreath products of cyclic groups, `arXiv:0607563`.

[Tao] T. Tao, *Basic theory of expander graphs*, 2 December 2011 blog entry; available electronically at `http://tinyurl.com/d6hhlge`

[Vad] S. Vadhan, *Pseudorandomness*, monograph draft; available electronically at `http://tinyurl.com/o9e7qa8`

[VV] M. Vorobets and Y. Vorobets, On a free group of transformations defined by an automaton, *Geom. Dedicata* **124** (2007), 237–249.