COMMUTATOR WIDTH IN THE FIRST GRIGORCHUK GROUP

LAURENT BARTHOLDI, THORSTEN GROTH, AND IGOR LYSENOK

Abstract. Let G be the first Grigorchuk group. We show that the commutator width of G is 2: every element $g \in [G, G]$ is a product of two commutators, and also of six conjugates of a. Furthermore, we show that every finitely generated subgroup $H \leq G$ has finite commutator width, which however can be arbitrarily large, and that G contains a subgroup of infinite commutator width. The proofs were assisted by the computer algebra system GAP.

1. Introduction

Let Γ be a group and let $\Gamma' = [\Gamma, \Gamma]$ denote its derived subgroup. The *commutator width* of Γ is the least $n \in \mathbb{N} \cup \infty$ such that every element of Γ' is a product of n commutators.

We compute, in this article, the commutator width of the *first Grigorchuk group* G, see §1.2 for a brief introduction. This is a prominent example from the class of *branched groups*, and as such is a good testing ground for decision and algebraic problems in group theory. We prove:

Theorem A. *The first Grigorchuk group and its branching subgroup have commutator width* 2*.*

It was already proven in [LMU16] that the commutator width of G is finite, without providing an explicit bound. Our result also answers a question of Elisabeth Fink [Fin14, Question 3]:

Corollary B. *Every element of G' is a product of* 6 *conjugates of the generator a and there are elements $g \in G'$ which are not products of* 4 *conjugates of a.*

There are examples of groups of finite commutator width with subgroups of infinite commutator width; and even finitely presented, perfect examples in which the subgroup has finite index, see Example 1. However, we can prove:

Theorem C. *Every finitely generated subgroup of G has finite commutator width; however, their commutator width cannot be bounded, even among finite-index subgroups. Furthermore, there is a subgroup of G of infinite commutator width.*

1.1. Commutator width.

Let Γ be a group. It is well-known that usually elements of Γ' are not commutators—for example, $[X_1, X_2] \cdots [X_{2n-1}, X_{2n}]$ is not a commutator in the free group F_{2n} when $n > 1$. In fact, every non-abelian free group has infinite commutator width, see [Rhe68].

On the other hand, some classes of groups have finite commutator width: finitely generated virtually abelian-by-nilpotent groups [Seg09], and finitely generated solvable groups of class 3, see [Rhe69].

Finite groups are trivial examples of groups of finite commutator width. There are finite groups in which some elements are not commutators, the smallest having order 96, see [Gur80]. On the other hand, non-abelian finite simple groups have commutator width 1, as was conjectured by Ore in 1951, see [Ore51], and proven in 2010, see [LOST10]. The commutator width cannot be bounded among finite groups; for example, $\Gamma_n = \langle x_1, \dots, x_{2n} \mid x_1^p, \dots, x_{2n}^p, \gamma_3(\langle x_1, \dots, x_{2n}\rangle)\rangle$ is a finite class-2 nilpotent group in which Γ_n' has order $p^{\binom{2n}{2}}$ but at most $\binom{p^{2n}}{2}$ elements are commutators, so Γ_n's commutator width is at least $n/2$.

Commutator width of groups, and of elements, has proven to be an important group property, in particular via its connections with "stable commutator length" and bounded cohomology [Cal09]. It is also related to solvability of quadratic equations in groups: a group Γ has commutator width $\leq n$ if and only if the equation $[X_1, X_2] \cdots [X_{2n-1}, X_{2n}]g = \mathbb{1}$ is solvable for all $g \in \Gamma'$. Needless to say, there are groups in which solvability of equations is algorithmically undecidable. It was proven in [LMU16] that there exists an algorithm to check solvability of quadratic equations in the first Grigorchuk group.

We note that if the character table of a group Γ is computable, then it may be used to compute the commutator width: Burnside shows (or, rather, hints) in [Bur55, §238, Ex. 7] that an element $g \in \Gamma$ may be expressed as a product of r commutators if and only if

$$\sum_{\chi \in \mathrm{Irr}(\Gamma)} \frac{\chi(g)}{\chi(1)^{2r-1}} > 0.$$

This may yield another proof of Theorem A, using the quite explicit description of $\mathrm{Irr}(G)$ given in [Bar13].

Consider a group Γ and a subgroup Δ. There is in general little connection between the commutator width of Γ and that of Δ. If Δ has finite commutator width and $[\Gamma : \Delta]$ is finite, then obviously Γ also has finite commutator width—for example, because $\Gamma/\operatorname{core}(\Delta)'$ is virtually abelian, and every commutator in Γ can be written as a product of a commutator in Δ with the lift of one in $\Gamma/\operatorname{core}(\Delta)'$, but that seems to be all that can be said. Danny Calegari pointed to us the following example:

Example 1. Consider the group Δ of orientation-preserving self-homeomorphisms of $\mathbb{R}$ that commute with integer translations, and let Γ be the extension of Δ by the involution $x \mapsto -x$. Then, by [EHN81, Theorems 2.3 and 2.4], every element of $\Gamma' = \Delta$ is a commutator in Γ, while the commutator width of Δ is infinite.

Both Γ and Δ can be made perfect by replacing them respectively with $(\Gamma \wr A_5)'$ and $\Delta \wr A_5$; and can be made finitely presented by restricting to those self-homeomorphisms that are piecewise-affine with dyadic slopes and breakpoints.

1.2. **Branched groups.** We briefly introduce the first Grigorchuk group [Gri80] and some of its properties. For a more detailed introduction into the topic of self-similar groups we refer to [BGŠ03, Nek05] and to Section 3.

A *self-similar group* is a group Γ endowed with an injective homomorphism $\Psi\colon \Gamma \to \Gamma \wr S_n$ for some symmetric group S_n. It is *regular branched* if there exists a finite-index subgroup $K \leq \Gamma$ such that $\Psi(K) \geq K^n$. It is convenient to write $\langle\!\langle g_1, \dots, g_n \rangle\!\rangle \pi$ for an element $g \in \Gamma \wr S_n$. We call g_i the *states* of g and π its *activity*. It is also convenient to identify, in a self-similar group, elements with their image under Ψ.

A self-similar group may be specified by giving a set S of generators, some relations that they satisfy, and defining Ψ on S. There is then a maximal quotient Γ of the free group F_S on which Ψ induces an injective homomorphism to $\Gamma \wr S_n$.

The first Grigorchuk group G may be defined in this manner. It is the group generated by $S = \{a, b, c, d\}$, with $a^2 = b^2 = c^2 = d^2 = bcd = \mathbb{1}$, and with

$$a = \langle\!\langle \mathbb{1}, \mathbb{1} \rangle\!\rangle (1, 2), \quad b = \langle\!\langle a, c \rangle\!\rangle, \quad c = \langle\!\langle a, d \rangle\!\rangle, \quad d = \langle\!\langle \mathbb{1}, b \rangle\!\rangle.$$

Here are some remarkable properties of G: it is an infinite torsion group, and more precisely for every $g \in G$ we have $g^{2^n} = \mathbb{1}$ for some $n \in \mathbb{N}$. On the other hand, it is not an Engel group, namely it is not true that for every $g, h \in G$ we have $[g, h, \dots, h] = \mathbb{1}$ for a long-enough iterated commutator [Bar16a]. It is a group of intermediate word growth [Gri83], and answered in this manner a celebrated question of Milnor.

We have decided to concentrate on the first Grigorchuk group in the computational aspects of this text; though our code would function just as well for other examples of self-similar branched groups, such as the Gupta-Sidki groups [GS83].

1.3. **Sketch of proofs.** The general idea for the proof of Theorem A is the decomposition of group elements into states via Ψ. We show that each element $g \in G'$ is a product of two commutators by solving the equation $\mathcal{E} = [X_1, X_2] \cdots [X_{2n-1}, X_{2n}]g$ for all $n \geq 2$.

If there is a solution then the values of the variables X_i have some activities σ_i. If we fix a possible activity of the variables of $\mathcal{E}$ then by passing to the states of the X_i we are led to two new equations which (under mild assumptions and after some normalization process) yields a single equation of the same form but of higher genus.

Not all solutions for the new equations lead back to solutions of the original equation. Thus instead of pure equations we consider *constrained* equations: we require the variables to lie in specified cosets of the finite-index subgroup K. The pair composed of a constraint and an element $g \in G$ will be a *good pair* if there is some n such that the constrained equation $[X_1, X_2] \cdots [X_{2n-1}, X_{2n}]g$ is solvable. It turns out that this only depends on the image of g in the finite quotient G/K'.

Then by direct computation we show that every good pair leads to another good pair in which the genus of the equation increases. We build a graph of good pairs which turns out to be finite since the constants of the new equation are states of the old equation and we can use the strong contracting property of G.

The computations could in principle be done by hand, but one of our motivations was precisely to see to which point they could be automated. We implemented them in the computer algebra system GAP [GAP14]. The source code for these computations is distributed with this document as ancillary material. It can be validated using precomputed data on a GAP standard installation by running the command `gap verify.g` in its main directory.

To perform more advanced experimentation with the code and to recreate the precomputed data, the required version of GAP must be at least 4.7.6 and the packages FR [Bar16b] and LPRES [BH16] must be installed.

2. Equations

We fix a set $\mathcal{X}$ and call its elements *variables*. We assume that $\mathcal{X}$ is infinite countable, is well ordered, and that its family of finite subsets is also well ordered, by size and then lexicographic order. We denote by $F_{\mathcal{X}}$ the free group on the generating set $\mathcal{X}$. We use $\mathbb{1}$ for the identity element of groups, and for the identity maps, to distinguish it from the numerical 1.

Definition 2.1 (G-group, G-homomorphism)**.** Let G be a group. A *G-group* is a group with a distinguished copy of G inside it; a typical example is $H * G$ for some group H. A *G-homomorphism* between G-groups is a homomorphism that is the identity between the marked copies of G.

A *G-equation* is an element $\mathcal{E}$ of the G-group $F_{\mathcal{X}} * G$, regarded as a reduced word in $\mathcal{X} \cup \mathcal{X}^{-1} \cup G$. For $\mathcal{E}$ a G-equation, its set of *variables* $\mathrm{Var}(\mathcal{E}) \subset \mathcal{X}$ is the set of symbols in $\mathcal{X}$ that occur in it; namely, $\mathrm{Var}(\mathcal{E})$ is the minimal subset of $\mathcal{X}$ such that $\mathcal{E}$ belongs to $F_{\mathrm{Var}(\mathcal{E})} * G$.

An *evaluation* is a G-homomorphism $e\colon F_{\mathcal{X}} * G \to G$. A *solution* of an equation $\mathcal{E}$ is an evaluation s satisfying $s(\mathcal{E}) = \mathbb{1}$. If a solution exists for $\mathcal{E}$ then the equation $\mathcal{E}$ is called *solvable.* The set of elements $X \in \mathcal{X}$ with $s(X) \neq \mathbb{1}$ is called the *support* of the solution.

The support of a solution for an equation $\mathcal{E}$ may be assumed to be a subset of $F_{\mathrm{Var}(\mathcal{E})}$ and hence the data of a solution is equivalent to a map $\mathrm{Var}(\mathcal{E}) \to G$. The question of whether an equation $\mathcal{E}$ is solvable will be referred to as the *Diophantine problem* of $\mathcal{E}$.

Every homomorphism $\varphi\colon G \to H$ extends uniquely to an $F_{\mathcal{X}}$-homomorphism $\varphi_*\colon F_{\mathcal{X}} * G \to F_{\mathcal{X}} * H$. In this manner, every G-equation $\mathcal{E}$ gives rise to an H-equation $\varphi_*(\mathcal{E})$, which is solvable whenever $\mathcal{E}$ is solvable.

Definition 2.2 (Equivalence of equations)**.** Let $\mathcal{E}, \mathcal{F} \in F_{\mathcal{X}} * G$ be two G-equations. We say that $\mathcal{E}$ and $\mathcal{F}$ are *equivalent* if there is a G-automorphism φ of $F_{\mathcal{X}} * G$ that maps $\mathcal{E}$ to $\mathcal{F}$. We denote by $\mathrm{Stab}(\mathcal{E})$ the group of G-automorphisms of $\mathcal{E}$.

Lemma 2.3. *Let $\mathcal{E}$ be an equation and let φ be a G-endomorphism of $F_{\mathcal{X}} * G$. If $\varphi(\mathcal{E})$ is solvable then so is $\mathcal{E}$. In particular, the Diophantine problem is the same for equivalent equations.*

Proof. If s is a solution for $\varphi(\mathcal{E})$, then $s \circ \varphi$ is a solution for $\mathcal{E}$. □

2.1. **Quadratic equations.** A G-equation $\mathcal{E}$ is called *quadratic* if for each variable $X \in \mathrm{Var}(\mathcal{E})$ exactly two letters of $\mathcal{E}$ are X or X^{-1}, when $\mathcal{E}$ is regarded as a reduced word.

A G-equation $\mathcal{E}$ is is called *oriented* if for each variable $X \in \mathrm{Var}(\mathcal{E})$ the number of occurrences with positive and with negative sign coincide, namely if $\mathcal{E}$ maps to the identity under the natural map $F_{\mathcal{X}} * G \to F_{\mathcal{X}}/[F_{\mathcal{X}}, F_{\mathcal{X}}] * \mathbb{1}$. Otherwise $\mathcal{E}$ is called *unoriented.*

Lemma 2.4. *Being oriented or not is preserved under equivalence of equations.*

Proof. $\mathcal{E}$ is oriented if and only if it belongs to the normal closure of $[F_{\mathcal{X}}, F_{\mathcal{X}}] * G$; this subgroup is preserved by all G-endomorphisms of $F_{\mathcal{X}} * G$. □

2.2. **Normal form of quadratic equations.**

Definition 2.5 ($\mathcal{O}_{n,m}, \mathcal{U}_{n,m}$)**.** For $m, n \geq 0$, $X_i, Y_i, Z_i \in \mathcal{X}$ and $c_i \in G$ the following two kinds of equations are called in *normal form*:

$$\mathcal{O}_{n,m}: \qquad [X_1, Y_1][X_2, Y_2] \cdots [X_n, Y_n] c_1^{Z_1} \cdots c_{m-1}^{Z_{m-1}} c_m \tag{1}$$

$$\mathcal{U}_{n,m}: \qquad X_1^2 X_2^2 \cdots X_n^2 c_1^{Z_1} \cdots c_{m-1}^{Z_{m-1}} c_m\ . \tag{2}$$

The form $\mathcal{O}_{n,m}$ is called the oriented case and $\mathcal{U}_{n,m}$ for $n > 0$ the unoriented case. The parameter n is referred to as the *genus* of the normal form of an equation.

We recall the following result, and give the details of the proof in an algorithmic manner, because we will need them in practice:

Theorem 2.6 ([CE81]). *Every quadratic equation $\mathcal{E} \in F_{\mathcal{X}} * G$ is equivalent to an equation in normal form, and the G-isomorphism can be effectively computed.*

Proof. The proof proceeds by induction on the number of variables. Starting with the oriented case: if the reduced equation $\mathcal{E}$ has no variables then it is already in normal form $\mathcal{O}_{0,1}$. If there is a variable $X \in \mathcal{X}$ occurring in $\mathcal{E}$ then X^{-1} also appears. Therefore the equation has the form $\mathcal{E} = uX^{-1}vXw$ or can be brought to this form by applying the automorphism $X \mapsto X^{-1}$. Choose $X \in \mathcal{X}$ in such a way that $\mathrm{Var}(v)$ is minimal.

We distinguish between multiple cases:

Case 1.0: $v \in G$. The word uw has fewer variables than $\mathcal{E}$ and can thus be brought into normal form $r \in \mathcal{O}_{n,m}$ by a G-isomorphism φ. If r ends with a variable, we use the G-isomorphism $\varphi \circ (X \mapsto Xw^{-1})$ to map $\mathcal{E}$ to the equation $rv^X \in \mathcal{O}_{n,m+1}$. If r ends with a group constant b, say $r = sb$, we use the isomorphism $\varphi \circ (X \mapsto Xbw^{-1})$ to map $\mathcal{E}$ to the equation $sv^X b \in \mathcal{O}_{n,m+1}$.

Case 1.1: $v \in \mathcal{X} \cup X^{-1}$. For simplicity let us assume $v \in \mathcal{X}$; in the other case we can apply the G-homomorphism $v \mapsto v^{-1}$. Now there are two possibilities: either v^{-1} occurs in u or v^{-1} occurs in w. In the first case $\mathcal{E} = u_1v^{-1}u_2X^{-1}vXw$, and then the G-isomorphism $X \mapsto X^{u_1}u_2$, $v \mapsto v^{u_1}$ yields the equation $[v, X]u_1u_2w$. In the second case $\mathcal{E} = uX^{-1}vXw_1v^{-1}w_2$ is transformed to $[X, v]uw_1w_2$ by the G-isomorphism $X \mapsto X^{uw_1}w_1^{-1}$, $v \mapsto v^{-uw_1}$. In both cases u_1u_2w, respectively uw_1w_2 have fewer variables and so composition with the corresponding G-isomorphism results in a normal form.

Case 2: $\mathrm{Length}(v) > 1$. In this case v is a word consisting of elements $X \cup X^{-1}$ with each symbol occurring at most once as v was chosen with minimal variable set, and some elements of G. If v starts with a constant $b \in G$ we use the G-homomorphism $X \mapsto bX$ to achieve that v starts with a variable $Y \in \mathcal{X}$, possibly by using the G-homomorphism $Y \mapsto Y^{-1}$. As in Case 1.1 there are two possibilities: Y^{-1} is either part of u or part of w. In the first case $\mathcal{E} = u_1Y^{-1}u_2X^{-1}Yv_1Xw$ we can use the G-isomorphism $X \mapsto X^{u_1v_1}u_2$, $Y \mapsto Y^{u_1v_1}v_1^{-1}$ to obtain $[Y, X]u_1v_1u_2w$. In the second we use the G-isomorphism $X \mapsto X^{uw_1v_1}v_1^{-1}w_1^{-1}$, $Y \mapsto Y^{-uw_1v_1}v_1^{-1}$ to obtain $[X, Y]uw_1v_1w_2$. In both cases the second subword has again fewer variables and can be brought into normal form by induction.

Therefore each oriented equation can be brought to normal form by G-isomorphisms.

In the unoriented case there is a variable $X \in \mathcal{X}$ such that $\mathcal{E} = uXvXw$. Choose v to have a minimal number of variables. By induction, the shorter word $uv^{-1}w$ is equivalent by φ to a normal form r.

The G-isomorphism $\varphi \circ (X \mapsto X^u v^{-1})$ maps $\mathcal{E}$ to $X^2 r$. If $r \in \mathcal{U}_{n,m}$ for some n, m, there remains nothing to do. Otherwise $r = [Y, Z]s$, and then the G-homomorphism

$$X \mapsto XYZ, \qquad Y \mapsto Z^{-1}Y^{-1}X^{-1}YZXYZ, \qquad Z \mapsto Z^{-1}Y^{-1}X^{-1}Z$$

maps $X^2 r$ to $X^2Y^2Z^2 s$. This homomorphism is indeed an isomorphism, with inverse

$$X \mapsto X^2Y^{-1}X^{-1}, \qquad Y \mapsto XYX^{-1}Z^{-1}X^{-1}, \qquad Z \mapsto XZ.$$

Note that $s \in \mathcal{O}_{n,m}$. If $n \geq 1$ then this procedure can be repeated with Z, in place of X, r. $\square$

For a quadratic equation $\mathcal{E}$ we denote by $\mathfrak{nf}(\mathcal{E}) := \mathfrak{nf}_{\mathcal{E}}(\mathcal{E})$ the image of $\mathcal{E}$ under the G-isomorphism $\mathfrak{nf}_{\mathcal{E}}$ constructed in the proof.

From now on we will consider oriented equations $\mathcal{O}_{n,1}$. For this we will use the abbreviation

$$R_n(X_1, \dots, X_{2n}) = \prod_{i=1}^{n} [X_{2i-1}, X_{2i}]$$

and often write $R_n = R_n(X_1, \dots, X_{2n})$ if the X_i are the first generators of $F_{\mathcal{X}}$.

2.3. Constrained equations.

Definition 2.7 (Constrained equations [LMU16])**.** Given an equation $\mathcal{E} \in F_{\mathcal{X}} * G$, a group H, a homomorphism $\pi\colon G \to H$ and a homomorphism $\gamma\colon F_{\mathcal{X}} \to H$, the pair $(\mathcal{E}, \gamma)$ is called a *constrained* equation and γ is called a *constraint* for the equation $\mathcal{E}$ on H.

A *solution* for $(\mathcal{E}, \gamma)$ is a solution s for $\mathcal{E}$ with the additional property that $\pi \circ s = \gamma$.

We note that the constraint γ needs only to be specified on $\operatorname{Var}(\mathcal{E})$.

3. Self-similar groups

Let T_n be the regular rooted n-ary tree and let S_n be the symmetric group on n symbols. The group $\operatorname{Aut}(T_n)$ consists of all root-preserving graph automorphisms of the tree T_n.

Let $T_{1,n}, \dots, T_{n,n}$ be the subtrees hanging from neighbors of the root. Every $g \in \operatorname{Aut}(T_n)$ permutes the $T_{i,n}$ by a permutation σ and simultaneously acts on each of them by isomorphisms $g_i\colon T_{i,n} \to T_{i^\sigma,n}$.

Note that for all i the tree T_n is isomorphic to $T_{i,n}$; identifying each $T_{i,n}$ with T_n, we identify each g_i with an element of $\operatorname{Aut}(T_n)$, and obtain in this manner an isomorphism

$$\Psi\colon \begin{array}{c} \operatorname{Aut}(T_n) \xrightarrow{\sim} \operatorname{Aut}(T_n) \wr S_n \\ g \mapsto \langle\!\langle g_1, \dots, g_n \rangle\!\rangle \sigma. \end{array}$$

A *self-similar group* is a subgroup G of $\operatorname{Aut}(T_n)$ satisfying $G \leq \Psi(G)$. For the sake of notation we will identify elements with their image under this embedding and will write $g = \langle\!\langle g_1, \dots, g_n \rangle\!\rangle \sigma$ for elements $g \in G$. Furthermore we will call $g_i \in G$ the *states* of the element g, will write $g@i := g_i$ to address the states, will call $\sigma \in S_n$ the *activity* of the element g, and will write $\operatorname{act}(g) := \sigma$.

3.1. Commutator width of $\mathbf{Aut(T_2)}$.

To give an idea of how the commutator width of Grigorchuk's group is computed, we consider as an easier example the group $\operatorname{Aut}(T_2)$. In this group we have the following useful property: for every two elements $g, h \in \operatorname{Aut}(T_n)$ the element $\langle\!\langle g, h \rangle\!\rangle$ is also a member of the group. This is only true up to finite index in the Grigorchuk group and will produce extra complications there.

Proposition 3.1. *The commutator width of* $\operatorname{Aut}(T_2)$ *is* 1*.*

For the proof we need a small observation:

Lemma 3.2. *Let H be a self-similar group acting on a binary tree. If $g \in H'$ then $g@2 \cdot g@1 \in H'$.*

Proof. It suffices to consider a commutator $g = [g_1, g_2]$ in H'. Then $g@2 \cdot g@1$ is the product, in some order, of all eight terms $(g_i@j)^\epsilon$ for all $i, j \in \{1, 2\}$ and $\epsilon \in \{\pm 1\}$. □

Proof of Proposition 3.1. Given any element $g \in \mathrm{Aut}(T_2)'$ we consider the equation $[X,Y]g$. If in it we replace the variable X by $\langle\!\langle X_1, X_2\rangle\!\rangle$ and Y by $\langle\!\langle Y_1, Y_2\rangle\!\rangle(1,2)$ we obtain $\langle\!\langle X_1^{-1}Y_2^{-1}X_2Y_2g@1, X_2^{-1}Y_1^{-1}X_1Y_1g@2\rangle\!\rangle$. Therefore, $[X,Y]g$ is solvable if the system of equations $\{X_2^{-1}Y_2^{-1}X_2Y_2g@1,\ X_1^{-1}Y_1^{-1}X_1Y_1g@2\}$ is solvable. We apply the $\mathrm{Aut}(T_2)$-homomorphism $X_1 \mapsto X_1, X_2 \mapsto Y_1^{-1}X_1Y_1g@2, Y_i \mapsto Y_i$ to eliminate one equation and one variable.

Thus the solvability of the constrained equation $([X,Y]g,(X \mapsto \mathbb{1}, Y \mapsto (1,2)))$ follows from the solvability of $X_1^{-1}Y_2^{-1}Y_1^{-1}X_1Y_1(g@2)Y_2(g@1)$ which is under the normal form $\mathrm{Aut}(T_2)$-isomorphism $Y_1 \mapsto Y_1Y_2^{-1}$ equivalent to the solvability of $[X_1,Y_1](g@2)^{Y_2}g@1$. After choosing $Y_2 = \mathbb{1}$ we are again in the original situation since $g@2g@1 \in H'$.

This allows us to recursively define a solution s for the equation $[X,Y]g$ as follows:

$$s(X) = \langle\!\langle a_1, b_1^{-1}a_1b_1g@2\rangle\!\rangle, \qquad s(Y) = \langle\!\langle b_1, \mathbb{1}\rangle\!\rangle(1,2), \qquad c_1 = g@2 \cdot g@1,$$

and for all $i \geq 1$

$$a_i = \langle\!\langle a_{i+1}, b_{i+1}^{-1}a_{i+1}b_{i+1}c_i@2\rangle\!\rangle, \qquad b_i = \langle\!\langle b_{i+1}, \mathbb{1}\rangle\!\rangle(1,2), \quad c_{i+1} = c_i@2 \cdot c_i@1.$$

Note that the elements $a_i, b_i \in \mathrm{Aut}(T_2)$ are well-defined, although they are constructed recursively out of the a_j, b_j for *larger* j. Indeed, if one considers the recursions above for $i \in \{1,\dots,n\}$ and sets $a_{n+1} = b_{n+1} = \mathbb{1}$, one defines in this manner elements $a_1^{(n)}, b_1^{(n)} \in \mathrm{Aut}(T_2)$ which form Cauchy sequences, and therefore have well-defined limits $a_1 = \lim a_1^{(n)}$ and $b_1 = \lim b_1^{(n)}$. □

4. The first Grigorchuk Group

The first Grigorchuk group [Gri80] is a finitely generated self-similar group acting faithfully on the binary rooted tree, with generators

$$a = \langle\!\langle \mathbb{1}, \mathbb{1}\rangle\!\rangle(1,2), \quad b = \langle\!\langle a, c\rangle\!\rangle, \quad c = \langle\!\langle a, d\rangle\!\rangle, \quad d = \langle\!\langle \mathbb{1}, b\rangle\!\rangle.$$

Some useful identities are

$$\begin{gathered} a^2 = b^2 = c^2 = d^2 = bcd = \mathbb{1}, \\ b^a = \langle\!\langle c, a\rangle\!\rangle, c^a = \langle\!\langle d, a\rangle\!\rangle, d^a = \langle\!\langle b, \mathbb{1}\rangle\!\rangle, \\ (ad)^4 = (ac)^8 = (ab)^{16} = \mathbb{1}. \end{gathered}$$

Definition 4.1 (Regular branched group)**.** A self-similar group Γ is called *regular branched* if it has a finite-index subgroup $K \leq \Gamma$ such that $K^{\times n} \leq \Psi(K)$.

Lemma 4.2 ([Roz93])**.** *The Grigorchuk group is regular branched with branching subgroup*

$$K := \left\langle (ab)^2 \right\rangle^G = \left\langle (ab)^2, (bada)^2, (abad)^2 \right\rangle.$$

The quotient $Q := G/K$ has order 16. □

For an equation $\mathcal{E} \in F_{\mathcal{X}} * G$, recall that $\mathrm{Stab}(\mathcal{E})$ denotes the group of G-automorphisms of $\mathcal{E}$.

Denote by U_n the subgroup of $\mathrm{Stab}(R_n)$ generated by the following automorphisms of F_{2n}:

$$\begin{array}{lll}
\varphi_i\colon & X_i \mapsto X_{i-1}X_i, \text{ others fixed} & \text{for } i = 2,4,\dots,2n,\\
\varphi_i\colon & X_i \mapsto X_{i+1}X_i, \text{ others fixed} & \text{for } i = 1,3,\dots,2n-1,\\
 & X_i \mapsto X_{i+1}X_{i+2}^{-1}X_i, & \\
\psi_i\colon & X_{i+1} \mapsto X_{i+1}X_{i+2}^{-1}X_{i+1}X_{i+2}X_{i+1}^{-1}, & \text{for } i = 1,3,\dots,2n-3\\
 & X_{i+2} \mapsto X_{i+1}X_{i+2}^{-1}X_{i+2}X_{i+2}X_{i+1}^{-1}, & \\
 & X_{i+3} \mapsto X_{i+1}X_{i+2}^{-1}X_{i+3}, \text{ others fixed} &
\end{array}$$

Remark. In fact, we have $U_n = \mathrm{Stab}(R_n)$ though formally we do not need the equality. Due to classical results of Dehn–Nielsen, $\mathrm{Stab}(R_n)$ is isomorphic to the mapping class groups $M(n,0)$ of the closed orientable surface of genus n. It can be checked that the automorphisms φ_i and ψ_i represent the Humphries generators of $M(n,0)$. For details on mapping class groups, see for example [FM11].

Lemma 4.3 ([LMU16]). *Given $n \in \mathbb{N}$ and a homomorphism $\gamma\colon F_{\mathcal{X}} \to Q$ with* $\mathrm{supp}(\gamma) \subset \langle X_1,\dots,X_{2n}\rangle$ *there is an element $\varphi \in U_n < \mathrm{Aut}(F_{\mathcal{X}})$ such that* $\mathrm{supp}(\gamma \circ \varphi) \in \langle X_1,\dots,X_5\rangle$. □

Lemma 4.4. *Identify the set $\{\gamma\colon F_{\mathcal{X}} \to Q \mid \mathrm{supp}(\gamma) \subset \langle X_1,\dots,X_n\rangle\}$ with Q^n. Then*

$$\left|Q^{2n}/U_n\right| \leq 90 \textit{ for all } n.$$

Proof. Note that according to our identification we have $Q^m \subset Q^n$ for $m < n$. By Lemma 4.3 every orbit Q^{2n}/U_n has a representative in Q^5. Let $\mathfrak{R}_n$ denote a set of representatives of Q^{2n}/U_n in Q^5. Since $U_n \subset U_{n+1}$ we can assume that $\mathfrak{R}_{n+1} \subset \mathfrak{R}_n$ for $n \geq 3$.

Direct computation shows that $|\mathfrak{R}_3| = 90$, see Section 6.3. □

Remark. In fact we have $|Q^{2n}/U_n| = 90$ for all $n \geq 3$. To prove this one can show by direct computation that $\mathfrak{R}_3 = \mathfrak{R}_4 = \mathfrak{R}_5$ and then show for all $\theta \in U_n$, $n \geq 6$ and $\gamma, \gamma' \in \mathfrak{R}_3, \gamma \neq \gamma'$ that $\gamma \circ \theta \neq \gamma'$.

Notation 4.5 ($\mathfrak{R}$, reduced constraint). Lemmas 4.3 and 4.4 imply that there is a set of 90 homomorphisms $\gamma\colon F_{\mathcal{X}} \to Q$ with $\mathrm{supp}(\gamma) \subset \langle X_1,\dots,X_5\rangle$ that is a representative system of the orbits Q^{2n}/U_n for each $n \geq 3$. Fix such a set $\mathfrak{R}$ and for $\gamma\colon F_{\mathcal{X}} \to Q$ with finite support (say $X_1,\dots,X_{2n}$) denote by φ_γ the G-homomorphism in U_n such that $\gamma \circ \varphi_\gamma \in \mathfrak{R}$.

The element $\gamma \circ \varphi_\gamma$ will be called a *reduced constraint.*

Lemma 4.6. *The solvability of a constrained equation $(R_n g, \gamma)$ is equivalent to the solvability of $(R_n g, \gamma \circ \varphi_\gamma)$.*

Proof. If s is a solution for $(R_n g, \gamma)$ then $s \circ \varphi_\gamma$ is a solution for $(R_n g, \gamma \circ \varphi_\gamma)$. □

Definition 4.7 (Branch structure [Bar13]). A *branch structure* for a group $G \hookrightarrow G \wr S_n$ consists of

1. a branching subgroup $K \trianglelefteq G$ of finite index;
2. the corresponding quotient $Q = G/K$ and the factor homomorphism $\pi\colon G \to Q$;

(3) a group $Q_1 \subset Q \wr S_n$ such that $\langle\!\langle q_1, \dots, q_n \rangle\!\rangle \sigma \in Q_1$ if and only if $\langle\!\langle g_1, \dots, g_n \rangle\!\rangle \sigma \in G$ for all $g_i \in \pi^{-1}(q_i)$;
(4) a map $\omega\colon Q_1 \to Q$ with the following property: if $g = \langle\!\langle g_1, \dots, g_n \rangle\!\rangle \sigma \in G$ then $\omega(\langle\!\langle \pi(g_1), \dots, \pi(g_n) \rangle\!\rangle \sigma) = \pi(g)$.

All regular branched groups have a branch structure (see [Bar13, Remark after Definition 5.1]). We will from now on fix such a structure for G and take the group K defined in Lemma 4.2 as branching subgroup and denote by Q the factor group with natural homomorphism $\pi\colon G \to G/K = Q$.

Remark. The branch structure of G is included in the FR package and can be computed by the method `BranchStructure(GrigorchukGroup)`.

4.1. **Good Pairs.** It is not true that for every $g \in G'$ and every constraint γ there is an $n \in \mathbb{N}$ such that the constrained equation $(R_n g, \gamma)$ is solvable. For example

$$\Big(R_n(ab)^2, (\gamma\colon X_i \mapsto \mathbb{1}\ \forall i)\Big)$$

is not solvable for any n because $(ab)^2 \notin K'$. This motivates the following definition.

Definition 4.8 (Good pair)**.** Given $g \in G'$ and $\gamma \in \mathfrak{R}$, the tuple (g, γ) is called a *good pair* if $(R_n g, \gamma)$ is solvable for some $n \in \mathbb{N}$.

Lemma 4.9. *Denote by*

$$\tau\colon G \to G/K' \quad \text{and} \quad \rho\colon G/K' \to (G/K')/(K/K') \simeq G/K$$

the natural projections.

The pair (g, γ) is a good pair if and only if there is a solution $s\colon F_{\mathcal{X}} \to G/K'$ for $R_3\tau(g)$ with $s(X_i) \in \rho^{-1}(\gamma(X_i))$.

Proof. If (g, γ) is a good pair and s a solution for $(R_n g, \gamma)$ then $s(X_i) \in K$ for $i \geq 6$, so $s(R_n) = s(R_3) \cdot k'$ for some $k' \in K'$. Therefore there is a solution $\tau \circ s$ for $R_3\tau(g)$ with $s(X_i) = \gamma(X_i)$.

On the other hand if there is a solution $s\colon F_{\mathcal{X}} \to G/K'$ for $R_3\tau(g)$ with for each $s(X_i) \in \rho^{-1}(\gamma(X_i))$ then for $g_i \in \tau^{-1}(s(X_i))$ there is some $k' \in K'$ such that $R_3(g_1, \dots, g_6)k'g = \mathbb{1}$ and so (g, γ) is a good pair. $\square$

The previous lemma shows that the question whether (g, γ) is a good pair depends only on the image of g in G/K'. For $q \in Q$, we call (q, γ) a *good pair* if (g, γ) is a good pair for one (and hence all) preimages of q under τ.

Corollary 4.10. *The following are equivalent:*
(a) K has finite commutator width;
(b) there is an $n \in \mathbb{N}$ such that $(R_n g, \gamma)$ is solvable for all good pairs (g, γ) with $g \in G'$ and $\gamma \in \mathfrak{R}$.

Proof. (b)⇒(a): if $k \in K'$ then $(k, \mathbb{1})$ is a good pair, so $(R_n k, \mathbb{1})$ is solvable in G; and the constraints ensures that it is solvable in K. Therefore the commutator width of K is at most n.

(a)⇒(b): if (g, γ) is a good pair there is an $m' \in \mathbb{N}$ and a solution s for $(R_{m'} g, \gamma)$. As $\pi(s(X_i)) = \mathbb{1}$ for all $i \geq 6$ there is $k \in K'$ such that s is a solution for $(R_3 kg, \gamma)$. By (a) there is an m such that all k can be written as product of m commutators of elements of K and therefore there is a solution for $(R_{m+3} g, \gamma)$. We may take $n = m + 3$. $\square$

We study now more carefully the quotients G/K, G/K' and $G/(K \times K)$.

Lemma 4.11. *Let us write* $k_1 := (ab)^2, k_2 := \langle\!\langle \mathbb{1}, k_1 \rangle\!\rangle = (abad)^2$ *and* $k_3 := \langle\!\langle k_1, \mathbb{1} \rangle\!\rangle = (bada)^2$. *Then*

$$\begin{aligned}
G' &= \left\langle k_1, k_2, k_3, (ad)^2 \right\rangle, \\
K &= \langle k_1, k_2, k_3 \rangle, \\
K \times K &= \{\langle\!\langle k, k' \rangle\!\rangle \mid k, k' \in K\} \\
&= \left\langle k_2, k_3, [k_1, k_2], [k_1, k_3], [k_1^{-1}, k_2], [k_1^{-1}, k_3] \right\rangle, \\
K' &= \langle [k_1, k_2] \rangle^G \\
&= \left\langle [k_2, k_1], [k_1, k_2^{-1}], [k_2, k_1]^{k_2}, [k_1^{-1}, k_2], [k_2, k_1]^{k_1}, [k_2^{-1}, k_1^{-1}] \right\rangle^{\{\mathbb{1}, a\}}
\end{aligned}$$

Furthermore these groups form a tower with indices

$$[G : G'] = 8, \qquad [G' : K] = 2, \qquad [K : K \times K] = 4, \qquad [K \times K : K'] = 16.$$

Proof. The chain of indices is shown for example in [BGŠ03] and the generating sets can be verified using the GAP standard methods `NormalClosure` and `Index`. $\square$

4.2. **Succeeding pairs.**

Definition 4.12 ($\mathfrak{R}_{\mathrm{act}}$, active constraints)**.** We define the activity $\mathrm{act}(q)$ of an element $q \in Q$ as the activity of an arbitrary element of $\pi^{-1}(q)$. This is well defined since all elements of K have trivial activity.

Consider a constraint $\gamma\colon F_{\mathcal{X}} \to Q$. Define $\mathrm{act}(\gamma)\colon F_{\mathcal{X}} \to C_2$ by $X \mapsto \mathrm{act}(\gamma(X))$.

Denote by $\mathfrak{R}_{\mathrm{act}}$ the reduced constraints in $\mathfrak{R}$ that have a nontrivial activity.

Lemma 4.13. *For each* $q \in G'/K'$ *there is* $\gamma \in \mathfrak{R}_{\mathrm{act}}$ *such that* (q, γ) *is a good pair.*

Proof. This is a finite problem which can be checked in GAP with the function `verifyLemmaExistGoodConstraints`. For more details see Section 6.1. $\square$

We will now give a procedure to start with a constrained equation say of class $\mathcal{O}_{n,1}$ and result with an equations of class $\mathcal{O}_{2n-1,1}$ and a a set of constraints such that the solvability of any of the later constrained equations implies the solvability of the original one.

Instead of an infinitely generated free group $F_{\mathcal{X}}$ we can restrict ourselves to a finite set $\mathcal{X}$ of order $2n$ for the variables of the original equation and another set $\mathcal{Y}$ for the variables of the resulting equation. For fixed n we notate the free groups $F_{\mathcal{X}}, F_{\mathcal{Y}}$ and $F_{\mathcal{Y}'}$ on the following generating sets:

$$\mathcal{X} = \{X_\ell \mid 1 \le \ell \le 2n\}, \quad \mathcal{Y} = \{Y_{\ell,i} \mid 1 \le \ell \le 2n,\ i = 1, 2\}, \quad \mathcal{Y}' = \mathcal{Y} \setminus \{Y_{6,1}, Y_{6,2}\}.$$

Denote by S the set $\{\mathbb{1}, a, b, c, d, ab, ad, ba\} \subset G$. We will define for all $q \in G'/K'$ a map Γ^q which for any $n \ge 3$ maps a reduced constraint $\gamma \in \mathfrak{R}_{\mathrm{act}}$, say $\gamma\colon F_{\mathcal{X}} \to Q$, to a set of constraints $\gamma'\colon F_{\mathcal{Y}} \to Q$ with the following property:

(*) There is $x \in S$ with $\gamma'(Y_{6,1}) = \pi(x)$, such that for all $g \in G'$ with $\tau(g) = q$ the solvability of the constrained equation $(R_{2n-1}(g@2)^x \cdot g@1, \gamma'|_{F_{\mathcal{Y}'}})$ implies the solvability of $(R_n g, \gamma)$.

We will define this map in several steps and afterwards show that for all good pairs (q, γ) and all g such that $\tau(g) = q$ there is some constraint $\gamma' \in \Gamma^q(\gamma)$ such that $((g@2)^x \cdot g@1, \gamma'|_{F_{\mathcal{Y}'}})$ is a good pair.

For the first step we take the branching structure (K, Q, π, Q_1, ω) of the Grigorchuk group as before and complete the set S to a transversal S' of G/K. Denote by $\mathrm{rep}\colon Q \to S'$ the map such that $\pi(\mathrm{rep}(q)) = q$.

$$\Gamma_1^n(\gamma) = \left\{ \gamma'\colon F_{\mathcal{Y}} \to Q \;\middle|\; \begin{array}{l} \langle\!\langle \gamma'(Y_{\ell,1}, \gamma'(Y_{\ell,2})\rangle\!\rangle \in \omega^{-1}(\gamma(X_\ell)), \\ \gamma'(Y_{k,i}) = \mathbb{1},\ 1 \le \ell \le 6,\ k > 6,\ i = 1,2 \end{array} \right\}.$$

For some formal equalities for equations in G we will need two auxiliary free groups $F_{\mathcal{G}} = \langle \mathfrak{g} \rangle$, $F_{\mathcal{H}} = \langle \mathfrak{g}_1, \mathfrak{g}_2 \rangle$, and define homomorphisms

$$\Phi_\gamma\colon \begin{array}{rl} F_{\mathcal{X}} * F_{\mathcal{G}} & \to (F_{\mathcal{Y}} * F_{\mathcal{H}}) \wr C_2, \\ \mathfrak{g} & \mapsto \langle\!\langle \mathfrak{g}_1, \mathfrak{g}_2 \rangle\!\rangle, \\ X_i & \mapsto \langle\!\langle Y_{i,1}, Y_{i,2} \rangle\!\rangle \operatorname{act}(\gamma(X_i)), \end{array} \qquad \tilde{\Phi}_\gamma\colon \begin{array}{rl} F_{\mathcal{X}} * G & \to (F_{\mathcal{Y}} * G) \wr C_2, \\ g & \mapsto \Psi(g), \\ X_i & \mapsto \langle\!\langle Y_{i,1}, Y_{i,2} \rangle\!\rangle \operatorname{act}(\gamma(X_i)). \end{array}$$

Lemma 4.14. *If γ is a constraint with nontrivial activity, and $\Phi_\gamma(R_n\mathfrak{g}) = \langle\!\langle w_1, w_2 \rangle\!\rangle$ then* $\operatorname{Var}(w_1) \cap \operatorname{Var}(w_2) \neq \emptyset$.

Proof. Let $\ell \in 1 \dots 2n$ be such that $\gamma(X_\ell)$ has nontrivial activity. Then R_n contains either a factor $[X_\ell, X_k]$ or $[X_k, X_\ell]$ for another generator $X_k \neq X_\ell$. Assume without loss of generality the first case. Let σ be the activity of $\gamma(X_k)$ then $\Phi_\gamma(R_n\mathfrak{g})$ contains a factor

$$[\langle\!\langle Y_{\ell,1}, Y_{\ell,2} \rangle\!\rangle (1,2), \langle\!\langle Y_{k,1}, Y_{k,2} \rangle\!\rangle \sigma] = \begin{cases} \langle\!\langle Y_{\ell,2}^{-1} Y_{k,2}^{-1} Y_{\ell,2} Y_{k,1}, Y_{\ell,1}^{-1} Y_{k,1}^{-1} Y_{\ell,1} Y_{k,2} \rangle\!\rangle & \text{if } \sigma = \mathbb{1} \\ \langle\!\langle Y_{\ell,2}^{-1} Y_{k,1}^{-1} Y_{\ell,1} Y_{k,2}, Y_{\ell,1}^{-1} Y_{k,2}^{-1} Y_{\ell,2} Y_{k,1} \rangle\!\rangle & \text{if } \sigma = (1,2). \end{cases}$$

In both cases $Y_{k,1}, Y_{k,2} \in \operatorname{Var}(w_1) \cap \operatorname{Var}(w_2)$. $\square$

For $q_1, q_2 \in Q$ and $n \ge 3 \in \mathbb{N}$ define

$$\Gamma_2^{q_1,q_2,n}(\gamma) = \left\{ \gamma' \in \Gamma_1^n(\gamma) \;\middle|\; \varpi\colon \begin{smallmatrix} F_{\mathcal{H}} \to Q \\ \mathfrak{g}_1 \mapsto q_1 \\ \mathfrak{g}_2 \mapsto q_2 \end{smallmatrix} \text{ satisfies } (\gamma' * \varpi)^2(\Phi_\gamma(R_n\mathfrak{g})) = \langle\!\langle \mathbb{1}, \mathbb{1} \rangle\!\rangle \right\}.$$

For $\gamma \in \mathfrak{R}_{\mathrm{act}}$ denote by v and w the elements of $F_{\{Y_{1,1},\dots,Y_{6,2}\}}$ such that $\Phi_\gamma(R_3\mathfrak{g}) = \langle\!\langle v, w \rangle\!\rangle \langle\!\langle \mathfrak{g}_1, \mathfrak{g}_2 \rangle\!\rangle$. Then

$$\Phi_\gamma(R_n(X_*)\mathfrak{g}) = \langle\!\langle v, w \rangle\!\rangle \langle\!\langle R_{n-3}(Y_{7,1}, \dots, Y_{2n,1})\mathfrak{g}_1, R_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}_2 \rangle\!\rangle.$$

By Lemma 4.14 there is $Y_0 \in \mathcal{Y} \cup \mathcal{Y}^{-1}$ such that $v = v_1 Y_0 v_2$ and $w = w_1 Y_0^{-1} w_2$. Then the $F_{\mathcal{H}}$-homomorphism

$$\ell_{Y_0,n}\colon \begin{array}{l} F_{\mathcal{Y}} * F_{\mathcal{H}} \to F_{\mathcal{Y}} * F_{\mathcal{H}}, \\ Y \mapsto \begin{cases} Y & \text{if } Y \neq Y_0 \\ w_2 R_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}_2 w_1 & \text{if } Y = Y_0 \end{cases} \end{array}$$

maps the second coordinate of $\Phi_\gamma(R_n(X_*)\mathfrak{g})$ to $\mathbb{1}$ and the first coordinate to an equation

$$\mathcal{E} = v_1 w_2 R_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}_2 w_1 v_2 R_{n-3}(Y_{7,1}, \dots, Y_{2n,1})\mathfrak{g}_1.$$

For $\gamma' \in \Gamma_2^{q_1,q_2,n}(\gamma)$ we have $\gamma'(Y_0) = (\gamma' * \varpi)(w_2\mathfrak{g}_2 w_1) = (\gamma' * \varpi)(\ell_{Y_0,n}(Y_0))$ and hence

$$\gamma' = (\gamma' * \varpi) \circ \ell_{Y_0,n} \text{ for all } \gamma' \in \Gamma_2^{q_1,q_2,n}(\gamma), \varpi\colon \mathfrak{g}_i \mapsto q_i, Y_0. \tag{1}$$

Consider the automorphisms

$$\psi_1\colon \begin{array}{l} F_{\mathcal{Y}} * F_{\mathcal{H}} \to F_{\mathcal{Y}} * F_{\mathcal{H}} \\ Y_{k,1} \mapsto Y_{k,1}^{\mathfrak{g}_1^{-1}} \\ Y_{k,2} \mapsto Y_{k,2}^{(\mathfrak{g}_2 w_1 v_1 \mathfrak{g}_1)^{-1}} \\ Y_{k,j} \mapsto Y_{k,j}, \end{array} \qquad \psi_2\colon \begin{array}{ll} F_{\mathcal{Y}} * F_{\mathcal{H}} \to F_{\mathcal{Y}} * F_{\mathcal{H}} & \\ Y_{k,1} \mapsto Y_{k,1}^{\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1} & \text{for } k > 6 \\ Y_{k,2} \mapsto Y_{k,2}^{\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1} & \text{for } k > 6 \\ Y_{k,j} \mapsto Y_{k,j} & \text{for } k \le 6,\ j = 1,2 \end{array}$$

$$\psi_3\colon \begin{array}{rll} F_{\mathcal{Y}} * F_{\mathcal{H}} \to F_{\mathcal{Y}} * F_{\mathcal{H}} & & \\ Y_{2k,1} \mapsto Y_{n+k,2} & & \text{for } k > 3 \\ Y_{2k-1,1} \mapsto Y_{n+k,1} & & \text{for } k > 3 \\ Y_{2k,2} \mapsto Y_{3+k,2} & & \text{for } k > 3 \\ Y_{2k-1,2} \mapsto Y_{3+k,1} & & \text{for } k > 3 \\ Y_{k,\ell} \mapsto Y_{k,\ell} & & \text{for } k \le 6,\ \ell = 1,2 \end{array}$$

and note that for $\mathfrak{nf}_{\gamma,n,Y_0} := \psi_3 \circ \psi_2 \circ \mathfrak{nf}_{v_1 w_2 \mathfrak{g}_2 w_1 v_2 \mathfrak{g}_1} \circ \psi_1$ we have

$$\mathfrak{nf}_{\gamma,n,Y_0}(\mathcal{E}) = R_{2n-1}(Y_{1,1}, Y_{1,2}, \dots, \cancel{Y_{6,1}}, \cancel{Y_{6,2}}, \dots, Y_{2n,2})\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1.$$

This leads to the following definition.

$$\Gamma_3^{q_1,q_2,n,Y_0}(\gamma) = \left\{\gamma' \circ \mathfrak{nf}_{\gamma,n,Y_0}^{-1}\colon F_{\mathcal{Y}} \to Q \;\middle|\gamma' \in \Gamma_2^{q_1,q_2,n}\right\}.$$

Note that $\mathfrak{nf}_{\gamma,n,Y_0}$ fixes the sets $\{Y_{k,\ell} \mid k > 6, \ell = 1,2\}$ and $\{Y_{k,\ell} \mid k \le 6, \ell = 1,2\}$ and hence for $k > 6$ we have $\gamma''(Y_{k,\ell}) = \mathbb{1}$ for all $\gamma'' \in \Gamma_3^{q_1,q_2,n,Y_0}(\gamma)$ independently of q_i, n, Y_0 and γ and therefore we can naturally identify the mappings $\Gamma_3^{q_1,q_2,n,Y_0}$ and $\Gamma_3^{q_1,q_2,3,Y_0}$ for every $n \ge 3$. Note further that $\mathfrak{nf}_{\gamma,n,Y_0}(Y_0) = Y_{6,2}$.

Given $g \in G'$, $g_i = g@i$ for $i = 1,2$, an active constraint $\gamma \in \mathfrak{R}_{\mathrm{act}}$ and $\gamma'' \in \Gamma_3^{\pi(g_1),\pi(g_1),n,Y_0}(\gamma)$ then a solution for the the constrained equation

$$\mathcal{E}' = (R_{2n-1}(Y_{*,*})g_2^{\mathrm{rep}(\gamma''(Y_{6,1}))}g_1, \gamma'')$$

can be extended by the map $Y_{6,1} \mapsto \mathrm{rep}(\gamma''(Y_{6,1}))$ to a solution s' of the equation $(R_{2n-1}(Y_{*,*})g_2^{Y_{6,1}}g_1, \gamma'')$. Notate the epimorphism $i_{\mathcal{H}}\colon F_{\mathcal{H}} \to G, \mathfrak{g}_k \mapsto g_k$ and note that since $\mathfrak{nf}_{\gamma,n,Y_0}$ is an $F_{\mathcal{H}}$-homomorphism $\mathfrak{n} := (\mathbb{1} * i_{\mathcal{H}}) \circ \mathfrak{nf}_{\gamma,n,Y_0}$ maps $\mathcal{E}$ to $R_{2n-1}(Y_{*,*})g_2^{Y_{6,1}}g_1$. Moreover by (1) we have that $\gamma' := \gamma'' \circ \mathfrak{n} \in \Gamma_2^{q_1,q_2,n}(\gamma)$.

Hence the map

$$s\colon Y_{i,j} \mapsto \begin{cases} w_2 g_2 w_1 & \text{if } i,j = 6,2 \\ s' \circ \mathfrak{n}(Y_{i,j}) & \text{otherwise} \end{cases}$$

is a solution for $\big((\mathbb{1} * i_{\mathcal{H}}) \circ \Phi_\gamma(R_n\mathfrak{g}), \gamma'\big)$ and thus also for $\big(\tilde{\Phi}_\gamma(R_n g), \gamma'\big)$. By the definition of ω the element $t_i := \langle\!\langle s(Y_{i,1}), s(Y_{i,2})\rangle\!\rangle\,\mathrm{act}(\gamma(X_i))$ belongs to G for all i. Moreover since $\gamma' \in \Gamma_1^n(\gamma)$ we have $\pi(t_i) = \gamma(X_i)$. Thus the mapping $X_i \mapsto t_i$ is a solution for $(R_n g, \gamma)$.

The map $\Gamma_3^{q_1,q_2,n,Y_0}$ does depend on the choice of the variable Y_0. To remove this dependency we observe that the set of all variables $Y_0 \in \mathrm{Var}(v) \cap \mathrm{Var}(w)$ does not depend on n and define

$$\Gamma_4^{q_1,q_2}(\gamma) = \bigcup_{Y_0 \in \mathrm{Var}(v)\cap\mathrm{Var}(w)} \Gamma_3^{q_1,q_2,3,Y_0}(\gamma).$$

Note that $q_1, q_2 \in Q$ are determined by $q \in G'/K'$ in the sense that there is a map $\bar{@}i\colon G'/K' \to Q$ such that if $\tau(g) = q$ and $g_i = g@i$ then $q_i = q\bar{@}i$. This map $\bar{@}i$ is well defined since $k'@i \in K$ for all $k' \in K'$. Thus we can write $\Gamma_4^{q_1,q_2}(\gamma)$ as $\Gamma_4^q(\gamma)$, and filter out those constraints that do not fulfill the requested properties; we finally define

$$\Gamma^q(\gamma) := \left\{\gamma' \in \Gamma_4^q(\gamma) \;\middle|\mathrm{act}(\gamma') \ne \mathbb{1}, \gamma'(Y_{6,1}) \in \pi(S)\right\}. \tag{2}$$

Note that (*) holds automatically by construction.

Proposition 4.15. *For each good pair (q,γ) with $q \in G'/K'$ and $\gamma \in \mathfrak{R}_{\mathrm{act}}$ the set $\Gamma^q(\gamma)$ contains some constraint γ' such that for all g with $\tau(g) = q$ the pair $\left((g@2)^{\mathrm{rep}(\gamma'(Y_{6,1}))} \cdot g@1, \gamma'|_{F_{\mathcal{Y}'}}\right)$ is a good pair.*

For the proof of this proposition we need an auxiliary lemma:

Lemma 4.16. *The map*

$$\bar{p}_h \colon \begin{array}{ll} G'/K' & \to G'/K \times K \\ gK' & \mapsto \left((g@2)^h \cdot g@1\right) K \times K \end{array}$$

is well defined.

Proof. We need to show that $k@i \in K \times K$ for $i = 1,2$ and $k \in K'$. Remember the generators $k_1 = (ab)^2$, $k_2 = (abad)^2$. Then

$$[k_1,k_2] = bb^a(db^a)^2 b^a b (b^a d)^2 = \langle\!\langle \mathbb{1}, cabab \rangle\!\rangle = \langle\!\langle \mathbb{1}, \langle\!\langle \mathbb{1}, dabac \rangle\!\rangle \rangle\!\rangle = \langle\!\langle \mathbb{1}, \langle\!\langle \mathbb{1}, k_2^{-1} k_1 \rangle\!\rangle \rangle\!\rangle.$$

So both states of $[k_1,k_2]$ are in $K \times K$. Now take an arbitrary element $k \in K'$. There is $n \in \mathbb{N}$, $\varepsilon \in \{1,-1\}$ and $g_i \in G$ such that $k = \prod_{j=1}^n [k_1,k_2]^{\varepsilon g_j}$ and therefore

$$k@i = \prod_{j=1}^n \left(([k_1,k_2]^{\varepsilon g_j})\, @i\right) = \prod_{j=1}^n \left(([k_1,k_2])@i^{g_j^{-1}}\right)^{\varepsilon g_j @ i^{g_j^{-1}}} \in K \times K$$

Then for $k \in K'$ we have

$$p_h(gk) = ((gk)@2)^h \cdot (gk)@1 = (g@2)^h \cdot (k@2)^h \cdot g@1 \cdot k@1 \in ((g@2)^h \cdot g@1)K \times K.$$

$\square$

Proof of Proposition 4.15. In the construction above it is clear that the sets Γ_3^{q,Y_0} and hence Γ_4^{q,Y_0} are nonempty. For the finitely many $\gamma \in \mathfrak{R}_{\mathrm{act}}$ checking whether some of the finitely many $\gamma' \in \Gamma_4^q(\gamma)$ fulfill $\gamma'(Y_{6,1}) \in \pi(S)$ and $\mathrm{act}(\gamma') \neq \mathbb{1}$ (i.e. $\gamma' \in \Gamma^q(\gamma)$) is implemented in the procedure below.

Define for $h \in G$ maps $p_h \colon G \to G$ by $g \mapsto (g@2)^h \cdot g@1$. These maps are in general not homomorphisms but by Lemma 3.2 for $g \in G'$ we have $p_h(g) \in G'$ for all $h \in G$.

By Lemma 4.16 we can define the map $\bar{p}_h \colon G'/K' \to G'/(K \times K)$ and the natural homomorphism

$$\rho' \colon G'/K' \to \left(G'/K'\right) / \left(K \times K/K'\right) \simeq G'/(K \times K)$$

and now we only need to show that there is a $\gamma' \in \Gamma^q(\gamma)$ such that all preimages of $\bar{p}_{\mathrm{rep}(\gamma'(Y_{6,1}))}(q)$ under ρ' form good pairs with $\gamma'|_{F_{\mathcal{Y}'}}$. In formulas with $\mathcal{P}$ the predicate of being a good pair what needs to be checked is:

$$\forall q \in G'/K'\ \forall \gamma \in \mathfrak{R}_{\mathrm{act}}\ \exists \gamma' \in \Gamma^q(\gamma)\ \forall r \in \rho'^{-1}(\bar{p}_{\mathrm{rep}(\gamma'(Y_{6,1}))}(q)) \colon \mathcal{P}(q,\gamma) \Rightarrow \mathcal{P}(r, \gamma'|_{F_{\mathcal{Y}'}}).$$

This last formula quantifies only over finite sets, and could be implemented. It can be checked in GAP with the function `verifyPropExistsSuccessor`. $\square$

Definition 4.17 (Succeding pair)**.** For each $q \in G'/K'$ and $\gamma \in \mathfrak{R}_{\mathrm{act}}$ such that (q,γ) is a good pair fix a constraint $\gamma' \in \Gamma^q(\gamma)$ and an element $x = \mathrm{rep}(\gamma'(Y_{6,1})) \in S$ with the property of Proposition 4.15.

By Lemma 4.6 we can replace $\gamma'|_{F_{\mathcal{Y}'}}$ by a reduced constraint γ'_r. For a good pair $(g,\gamma) \in G' \times \mathfrak{R}_{\mathrm{act}}$ the *succeeding pair* is defined as $((g@2)^x g@1, \gamma'_r)$. Moreover by applying this iteratively we get the *succeeding sequence* (g_k,γ_k) of (g,γ): $(g_0,\gamma_0) = (g,\gamma)$ and (g_{k+1},γ_{k+1}) is the succeding pair of (g_k,γ_k).

The following lemma illustrates the use of the construction.

Lemma 4.18. *Let (g_k, γ_k) be the succeeding sequence of a good pair (g, γ). If $(g_i, \gamma_i) = (g_j, \gamma_j)$ for some distinct i, j then the equation $(R_n g, \gamma)$ is solvable for all $n \geq 3$.*

Proof. By (*) for any i, j with $i < j$ and any $n \geq 3$ there exists $n' > n$ such that solvability of $(R_{n'} g_j, \gamma_j)$ implies solvability of $(R_n g_i, \gamma_i)$. If $(g_i, \gamma_i) = (g_j, \gamma_j)$ then n' can be taken arbitrarily large. If (g, γ) is a good pair then (g_i, γ_i) is also a good pair by construction. We deduce the solvability of $(R_n g_i, \gamma_i)$ and hence the solvability of $(R_n g, \gamma)$. □

4.3. **Product of 3 commutators.** We will prove that every element $g \in G'$ is a product of three commutators by proving that all succeeding sequences (g_k, γ_k) as defined after Proposition 4.15 loop after finitely many steps. For this purpose remember the map $p_x \colon g \mapsto (g@2)^x g@1$ from the proof of Proposition 4.15. We will show that for each $g \in G'$ the sequence of sets

$$\mathrm{Suc}_1^g = \{g\}, \ \mathrm{Suc}_n^g = \{p_x(h) \mid h \in \mathrm{Suc}_{n-1}^g, x \in S\}$$

stabilizes in a finite set.

In [Bar98] there is a choice of weights on generators which result in a length on G with good properties.

Lemma 4.19 ([Bar98])**.** *Let $\eta \approx 0.811$ be the real root of $x^3 + x^2 + x - 2$ and set the weights*

$$\omega(a) = 1 - \eta^3 \qquad \omega(c) = 1 - \eta^2$$
$$\omega(b) = \eta^3 \qquad \omega(d) = 1 - \eta$$

then

$$\begin{aligned} \eta(\omega(b) + \omega(a)) &= \omega(c) + \omega(a) \\ \eta(\omega(c) + \omega(a)) &= \omega(d) + \omega(a) \\ \eta(\omega(d) + \omega(a)) &= \omega(b). \end{aligned}$$

□

The next lemma is a small variation of a lemma in [Bar98].

Lemma 4.20. *Denote by ∂_ω the length on G induced by the weight ω. Then there are constants $C \in \mathbb{N}$, $\delta < 1$ such that for all $x \in S$, $g \in G$ with $\partial_\omega(g) > C$ it holds $\partial_\omega(p_x(g)) \leq \delta \partial_\omega(g)$.*

Corollary 4.21. *The sequences of sets*

$$\mathrm{Suc}_1^g = \{g\}, \ \mathrm{Suc}_n^g = \{p_x(h) \mid h \in \mathrm{Suc}_{n-1}^g, x \in S\}$$

stabilizes at a finite step for all $g \in G$.

Proof of Lemma (see [Bar98, Proposition 5]*).* Each element $g \in G$ can be written in a word of minimal length of the form $g = a^\varepsilon x_1 a x_2 a \dots x_n a^\zeta$ where $x_i \in \{b, c, d\}$ and $\varepsilon, \zeta \in \{0, 1\}$. Denote by n_b, n_c, n_d the number of occurrences of b, c, d accordingly.

Then

$$\begin{aligned}
\partial_\omega(g) &= (n-1+\varepsilon+\zeta)\omega(a) + n_b\omega(b) + n_c\omega(c) + n_d\omega(d) \\
\partial_\omega(p_x(g)) &\leq (n_b+n_c)\omega(a) + n_b\omega(c) + n_c\omega(d) + n_d\omega(b) + 2\partial_\omega(x) \\
&= \eta\left((n_b+n_c+n_d)\omega(a) + n_b\omega(b) + n_c\omega(c) + n_d\omega(d)\right) + 2\partial_\omega(x) \\
&= \eta(\partial_\omega(g) + (1-\varepsilon-\zeta)\omega(a)) + 2\partial_\omega(x) \\
&\leq \eta(\partial_\omega(g) + \omega(a)) + 2(\omega(a)+\omega(b)) \\
&= \eta(\partial_\omega(g) + \omega(a)) + 2.
\end{aligned}$$

Thus the length of $p_x(g)$ growths with a linear factor smaller than 1 in terms of the length of g. Therefore the claim holds. For instance one could take $\delta = 0.86$ and $C = 50$ or $\delta = 0.96$ and $C = 16$. □

This completes the proof of the following proposition:

Proposition 4.22. *If $n \geq 3$ and (g, γ) is a good pair with active constraint γ with* $\operatorname{supp}(\gamma) \subset \{X_1, \ldots, X_{2n}\}$ *then the constrained equation $(R_n(X_1, \ldots, X_{2n})g, \gamma)$ is solvable.* □

Corollary 4.23. *The Grigorchuk group G has commutator width at most* 3.

Proof. This is a direct consequence of the proposition and Lemma 4.13. □

4.4. **Product of 2 commutators.** The case of products of two commutators can be reduced to the case of three commutators by using the same method as before.

We can compute the orbits of Q^4/U_2 and take a representative system denoted by $\mathfrak{R}^4$. It turns out that there are 86 orbits and we can check that there are again enough active constraints:

Lemma 4.24. *For each $q \in G'/K'$ there is $\gamma \in \mathfrak{R}^4_{\text{act}}$ such that (q, γ) is a good pair.*

Proof. This can be checked in GAP with the function
`verifyLemmaExistGoodGammasForRed4`. □

To formulate an analog of Proposition 4.15 we literally transfer the definition of the function Γ^q to the case $n = 2$. Denote the new function $\Gamma^{q,2}$. For a constraint $\gamma\colon F_{\{X_1,\ldots,X_4\}} \to Q$ with nontrivial activity it produces a set $\Gamma^{q,2}(\gamma)$ of constraints $\gamma'\colon F_{\{Y_{1,1},\ldots,Y_{4,2}\}} \to Q$.

Proposition 4.25. *For each good pair (q, γ) with $q \in G'/K'$ and $\gamma \in \mathfrak{R}^4_{\text{act}}$ the set $\Gamma^{q,2}(\gamma)$ contains some active constraint γ' such that for all g with $\tau(g) = q$ the pair $\left((g@2)^{\operatorname{rep}(\gamma'(Y_{4,1}))} \cdot g@1, \gamma'|_{\mathcal{F}_{\{Y_{1,1},\ldots Y_{3,2}\}}}\right)$ is a good pair.*

Proof. The proof is the same as for Proposition 4.15. The corresponding formula which needs to be checked is

$$\forall q \in G'/K'\ \forall \gamma \in \mathfrak{R}^4_{\text{act}}\ \exists \gamma' \in \Gamma^q(\gamma)\ \forall r \in \rho'^{-1}(\bar{p}_{\operatorname{rep}(\gamma'(Y_{4,1}))}(q))\colon \mathcal{P}(q,\gamma) \Rightarrow \mathcal{P}(r,\gamma').$$

This can be checked in GAP with the function `verifyPropExistsSuccessor`. □

The resulting succeeding pairs are now equations of genus 3 with an active constraint. Those are already shown to be solvable by Proposition 4.22. Hence we have the following corollary which improves Proposition 4.22:

Corollary 4.26. *If $n \geq 2$ and (g, γ) is a good pair with active constraint γ with* $\operatorname{supp}(\gamma) \subset \{X_1, \ldots, X_n\}$ *then the constrained equation $(R_n(X_1, \ldots, X_{2n})g, \gamma)$ is solvable.*

Together with Lemma 4.24 this proves the first part of Theorem A.

Corollary 4.27. *K has commutator width at most* 2.

Proof. To show that K has commutator width at most 2 it is sufficient to show that the constrained equations $(R_2 g, \mathbb{1})$ have solutions for all $g \in K'$. Since $\mathbb{1}$ has trivial activity one cannot directly apply Proposition 4.22. However one can check that all pairs $(h, \gamma_1), (f, \gamma_2)$ such that $g = \langle\!\langle h, f \rangle\!\rangle$ and $\gamma_1 = (\mathbb{1}, \mathbb{1}, \pi(bad), \mathbb{1})$, $\gamma_2 = (\mathbb{1}, \mathbb{1}, \mathbb{1}, \pi(ca))$ are good pairs with active constraints and hence admit solutions $s_1, s_2 \colon F_4 \to G$.

We can then define the map $s \colon F_4 \to G, X_i \mapsto \langle\!\langle s_1(X_i), s_2(X_i) \rangle\!\rangle$; it is a solution for $R_2 g$ and $s(X_i) \in K$ for all $i = 1, \ldots, 4$. Therefore the commutator width of K is at most 2.

This can be checked in GAP with the function `verifyCorollaryFiniteCWK`. □

4.5. **Not every element is a commutator.** The procedure used to prove that every element is a product of two commutators can not be used to prove that every element is a commutator since for equations of genus 1 the genus does not increase by passing to a succeeding pair.

In fact not every element $g \in G'$ is a commutator. This can be seen by considering finite quotients. A commutator in the group would be also a commutator in the quotient group.

We will define an epimorphism to a finite group with commutator width 2.

Analogously to the construction of $\Psi\colon \operatorname{Aut}(T_n) \to \operatorname{Aut}(T_n) \wr S_n$ we can define a homomorphism $\Psi_n \colon G \to G \wr_{2^n} (G/\operatorname{Stab}_G(n))$ by mapping an element g to its actions on the subtrees with root in level n and the activity on th n-th level of the tree.

Consider the following epimorphism:

$$\text{germ}\colon \quad \begin{aligned} G &\to \langle b, c, d \rangle \simeq C_2 \times C_2, \\ a &\mapsto \mathbb{1}, \\ b, c, d &\mapsto b, c, d. \end{aligned}$$

It extends to an epimorphism $\text{germ}_n \colon G \wr_{2^n} G/\operatorname{Stab}_G(n) \to \text{germ}(G) \wr_{2^n} G/\operatorname{Stab}_G(n)$. We will call the image $\text{germ}(G) =: G_0$ the 0-th *germgroup* and furthermore $G_n := \text{germ}_n \circ \Psi_n(G)$ the n-th *germgroup*.

The 4-th germgroup of the Grigorchuk group has order 2^{26} and has commutator width 2. If the FR package is present this group can be constructed in GAP with the following command.

```
gap> Range(EpimorphismGermGroup(GrigorchukGroup,4))
```

There is an element in the commutator subgroup of this germgroup which is not a commutator. This element is part of the precomputed data and can be accessed in GAP as `PCD.nonCommutatorGermGroup4`. For the computation of this element we used the character table of G_4. For more details see Section 6.2.

A corresponding preimage in G with a minimal number of states is the automaton shown in Figure 1. The construction of the element can be found in the file `gap/precomputeNonCommutator.g`. With the representation in standard generators it is easy to show using the homomorphism π on the generators that this element is even a member of K. This finishes the proof of Theorem A.

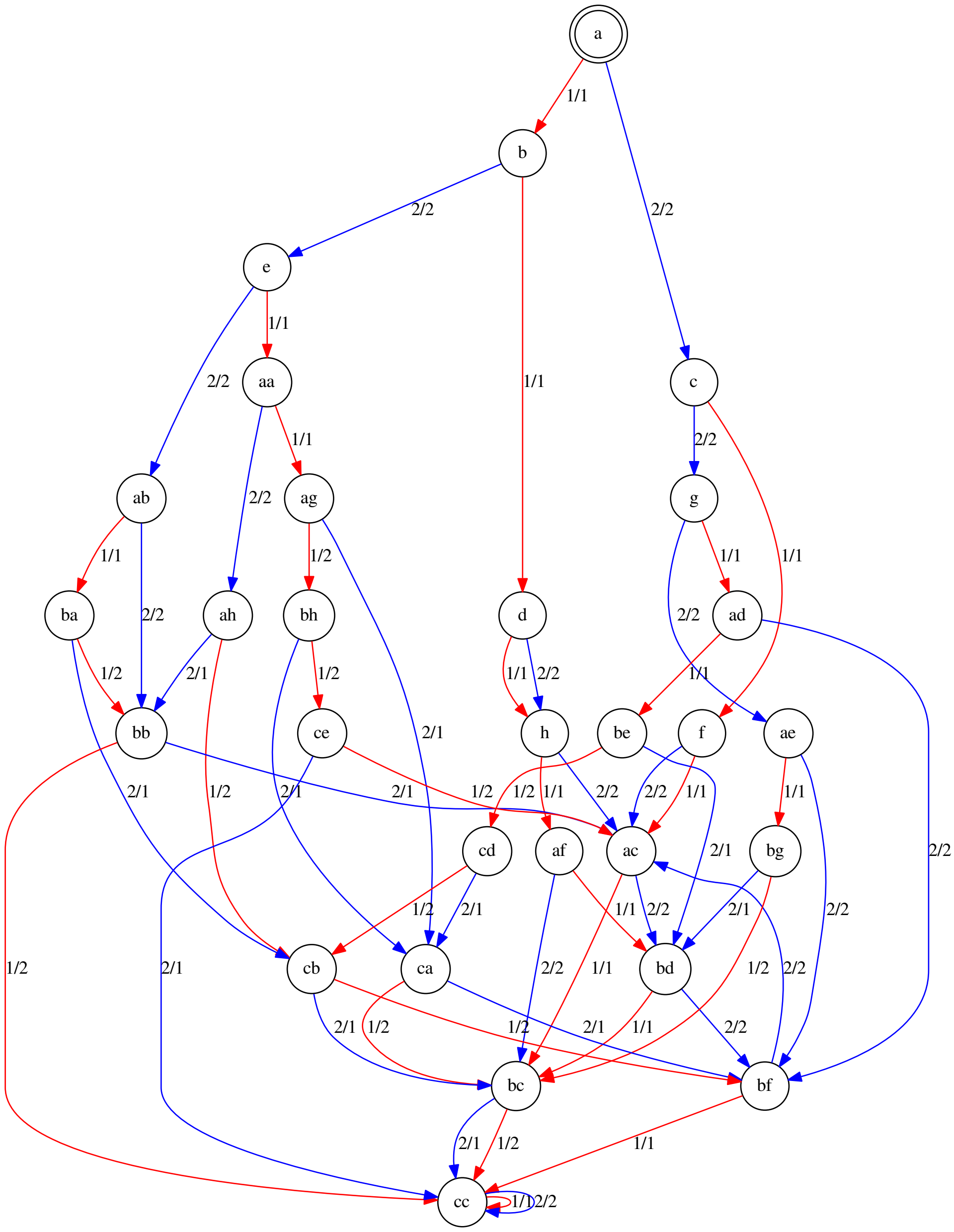

FIGURE 1. Element of the derived subgroup of the Grigorchuk group which is not a commutator. In standard generators:
$(acabacad)^3acab(ac)^2(acabacad)^2(acab)^3acadacab(ac)^2$
$(acabacad)^2(acabacadacab(ac)^3abacad(acab)^2)^5acabacadacab(ac)^2$
$(acabacad)^2(acabacadac)^2(abac)^3adacab(ac)^2(acabacad)^3$
$acab(ac)^2(acab(ac)^3abacad)^2acabacad((acabacadacab(ac)^2)^2$
$acabacad(acab)^3acadacab(ac)^2)^2((acabacad)^3acab)^2$
$acab(acabacad)^2acab(ac)^2(acabacad)^3acab(ac)^3aba$

4.6. **Bounded conjugacy width.** In [Fin14] it is proven that G has finite bounded conjugacy width. Here we give an explicit bound on this width.

Proposition 4.28. *Let g be in G'. Then the equation*

$$a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}ag = \mathbb{1}$$

is solvable in G.

Proof. We need to solve the constrained equation $(\mathcal{E} = a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}ag, \gamma)$ for some constraint γ. Independently of the chosen constraint, replacement of the variable X_i by $\langle\!\langle Y_i, Z_i\rangle\!\rangle\,\mathrm{act}(\gamma(X_i))$ leads after normalization to an equivalent equation $R_2(g@2)(g@1)$. Similarly to the construction of Γ^q in the previous section, one can find for each $q \in G'/K'$ a constraint γ such that $\gamma(\mathcal{E}^{\mathbb{1}*\pi}) = \mathbb{1}$ and $\gamma' \in \Gamma_1(\gamma)$ such that for all $g \in \pi^{-1}(q)$ the pairs $(g@2g@1, \gamma')$ are good pairs and γ' is an active constraint. Therefore the constrained equation $(R_2(g@2)(g@1), \gamma')$ is solvable by Corollary 4.26 for each $g \in G'$ and hence the equation $a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}ag$. This can be checked in GAP with the function `verifyExistGoodConjugacyConstraints`. □

Lemma 4.29. *There exits an element $g \in G'$ such that the equation*

$$a^{X_1}a^{X_2}a^{X_3}ag = \mathbb{1}$$

is not solvable.

Proof. As before independently of the activities of a possible constraint γ and of the element $g \in G'$ the normalform of $\tilde{\Phi}_\gamma(a^{X_1}a^{X_2}a^{X_3}ag)$ turns out to be $R_1(g@2)g@1$. So all there is to prove is that there is an element $h \in K$ where the products of states $h@2 \cdot h@1$ is not a commutator.

The element g displayed in Figure 1 provides such an element. It can easily be verified that $\langle\!\langle \pi(cag), \pi(ac)\rangle\!\rangle \in Q_1$ and $\omega(\langle\!\langle \pi(cag), \pi(ac)\rangle\!\rangle) = \mathbb{1}$. Thus by the properties of the branch structure we have $\langle\!\langle \pi(cag), \pi(ac)\rangle\!\rangle \in K < G'$. □

This finishes the proof of Corollary B.

Definition 4.30 (Conjugacy width [Fin14])**.** The *conjugacy width* of a group G with respect to a generating set S is the smallest number $N \in \mathbb{N}$ such that every element $g \in G$ is a product of at most N conjugates of generators $s \in S$.

Corollary 4.31. *The Grigorchuk group G with generating set $\{a, b, c, d\}$ has conjugacy width at most* 8.

Proof. The following set T is a transversal of G/G':

$$T = \{\mathbb{1}, a, d^a a, d^a, b, ab^a, ca^d, bd^a\}.$$

Therefore, every element $g \in G$ can be written as $g = th$ with $t \in T$ and $h \in G'$. As every element of G' is a product of at most 6 conjugates of a this proves the claim. □

5. Proof of Theorem C

We will prove the statement first for finite-index subgroups.

Proposition 5.1. *All finite-index subgroups $H \leq G$ have finite commutator width.*

Proof. Note that from Corollary 4.27 it follows that $K \times K$ and furthermore $K^{\times n}$ have commutator width 2.

Let H be a subgroup of finite index. Since G has the congruence subgroup property ([BG02]) we can find a nontrivial normal subgroup $N = \mathrm{Stab}_G(m) < H$ for some $m \in \mathbb{N}$. Furthermore since K is inactive we have

$$K < \mathrm{Stab}_G(1), \qquad K \times K < \mathrm{Stab}_G(2), \qquad K^{\times 4} < \mathrm{Stab}_G(3).$$

Furthermore we have $\mathrm{Stab}_G(n) = \mathrm{Stab}_G(3)^{\times 2^{n-3}}$ for $n \geq 4$ and hence for every subgroup H of finite index there is an n such that $K^{\times 2^n} \leq H$.

Since K' has finite index in K by Lemma 4.11, the index in $[H,H]$ of $[K^{\times 2^n}, K^{\times 2^n}]$ is finite. Taking a transversal T of $[H,H]/[K^{\times 2^n}, K^{\times 2^n}]$ we can find $m \in \mathbb{N}$ such that every element in T is a product of at most m commutators in H. We can thus write each element $h \in [H,H]$ as product kt with $k \in K^{\times 2^n}$, $t \in T$ and thus as a product of at most $2+m$ commutators. □

Proposition 5.2. *All finitely generated subgroups $H \leq G$ are of finite commutator width.*

Proof. Every infinite finitely generated subgroup of G is abstractly commensurable to G, see [GW03, Theorem 1].

This, by definition, means that every infinite finitely generated subgroup $H \leq G$ contains a finite-index subgroup which is isomorphic to a finite-index subgroup of G and hence by Proposition 5.1 has finite commutator width. □

To show that there cannot be a bound on the commutator width of subgroups we need some auxiliary results. They are well-known, but since we could not find an original reference we will sketch their proofs here.

Proposition 5.3.

(1) *For all $n \in \mathbb{N}$ there is a finite 2-group of commutator width at least n.*
(2) *K contains every finite 2-group as a subgroup.*
(3) *Every finite 2-group is a quotient of two finite-index subgroups of G.*

Proof.

(1) Consider the groups $\Gamma_n = F_n/\langle \gamma_3(F_n), x_1^2, \ldots, x_n^2\rangle$. These are extensions of C_2^n by $C_2^{\binom{n}{2}}$ and are class 2-nilpotent 2-groups. The derived subgroup is hence of order $2^{\binom{n}{2}}$. Let T be a transversal of Γ_n/Γ_n'. Thus T is of order 2^n and for $x, y \in \Gamma_n$ there are $t, s \in T$ and $x', y' \in \Gamma'$ such that every commutator $[x,y] = [tx', sy'] = [t,s]$. Therefore there are at most $\binom{2^n}{2}$ commutators.

   This means there are at most $\binom{2^n}{2}^m \leq 2^{(2n-1)m}$ products of m commutators but the size of Γ_n' is $2^{\binom{n}{2}} \geq 2^{\frac{n^2}{4}}$ and hence the commutator width of Γ_{8m} is at least m.
(2) K contains for each n the n-fold iterated wreath product $W_n(C_2) = C_2 \wr \cdots \wr C_2$. This can be shown by finding finitely many vertices of the tree T_2 which define a (spaced out) copy of the finite binary rooted tree with n levels T_2^n, and finding elements $k_i \in K$ such that $\langle k_i \rangle$ acts on T_2^n like the full group of automorphisms $\mathrm{Aut}(T_2^n) \simeq W_n(C_2)$.

   Then since $W_n(C_2)$ is a Sylow 2-subgroup of S_{2^n} every finite 2-group is a subgroup of $W_n(C_2)$ for some n, and hence a subgroup of K.

(3) Consider again some the vertices of T_2 which define a copy of the finite tree T_2^n on which a subgroup of K acts like $W_n(C_2)$. If we take m large enough such that all these vertices are above the m-th level we can find a copy of $W_n(C_2)$ inside $G/\operatorname{Stab}_G(m)$. $\square$

In the following theorem we summarize our results for the commutator width of the Grigorchuk group.

Theorem 5.4.

(1) G and its branching subgroup K have commutator width 2.
(2) All finitely generated subgroups $H \le G$ have finite commutator width.
(3) The commutator width of subgroups is unbounded even among finite-index subgroups.
(4) There is a subgroup of G with infinite commutator width.

Proof. Statements (1) and (2) are proven in Theorem A and Proposition 5.2. For every $n \in \mathbb{N}$ we can find two groups H_1, H_2 of finite index in G such that H_1/H_2 has commutator width at least n. Then H_1 has commutator width at least n as well and thus the commutator width of finite-index subgroups can not be bounded.

For the last claim, consider a sequence (H_i) of subgroups of K such that H_i has commutator width at least i. Let $\psi_0\colon K \to K \times K \le K$ be the map $k \mapsto \langle\!\langle k, \mathbb{1}\rangle\!\rangle$ and for $i \ge 1$ let $\psi_i\colon K \to K \times K \le K$ be the map $k \mapsto \langle\!\langle \mathbb{1}, \psi_{i-1}(k)\rangle\!\rangle$. Then $H := \langle \psi_i(H_i) : i \in \mathbb{N}\rangle$ is a subgroup of K and hence of G and is isomorphic to the restricted direct product of the H_i, so it has infinite width. $\square$

6. Implementation in GAP

6.1. **Usage of the attached files.** Typing the command `gap verify.g` in the main directory of the archive will produce as output a list of functions with their return value. All these functions should return `true`.

This approach uses precomputed data which are also in the archive, and is very fast.

Furthermore, these data can be recomputed if a sufficiently new version of GAP and some packages are present. For details see Section 6.2.

This is what the functions check:

`verifyLemma90orbits`: This function verifies that there are indeed 90 orbits of U_3 on Q^6 as claimed in Lemma 4.4.

`verifyLemma86orbits`: Analogously to the previous function this one verifies that there are 86 orbits of U_2 on Q^4.

`verifyLemmaExistGoodConstraints`: This verifies that for each $q \in G'/K'$ there is some $\gamma \in \mathfrak{R}_{\mathrm{act}}$ such that (q, γ) forms a good pair. This is claimed in Lemma 4.13.

`verifyLemmaExistGoodConstraints4`: This is a sharper version of the previous function. It checks that the above statement is already true if one replaces $\mathfrak{R}_{\mathrm{act}}$ by $\mathfrak{R}_{\mathrm{act}}^4$ as claimed in Lemma 4.24.

`verifyPropExistsSuccessor`: This verifies that for each good pair $(q, \gamma) \in G'/K' \times (\mathfrak{R}_{\mathrm{act}} \cup \mathfrak{R}_{\mathrm{act}}^4)$ there exists a $\gamma' \in \Gamma^q(\gamma)$ such that all preimages of $\bar{p}_{\mathrm{rep}(Y_{6,1})}(q)$ under the map ρ' form good pairs with the constraint γ'. This is needed in the proof of Proposition 4.15 and Proposition 4.25.

`verifyCorollaryFiniteCWK`: Corollary 4.27 needs the existence of succeeding good pairs of the pair $(\mathbb{1}, \mathbb{1}) \in K'/K' \times \mathfrak{R}^4$. This function verifies this existence.

`verifyExistGoodConjugacyConstraints`: This verifies that for the equation $a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}a$ there are constraints γ that admit good succeeding pairs. This is needed in the proof of Proposition 4.28.

`verifyGermGroup4hasCW`: This function verifies the existence of an element in the derived subgroup of the 4-th level germgroup that is not a commutator.

6.2. **Precomputed data.** In the interactive gap shell started by `gap verify.g` the precomputed data is read from some files in `gap/PCD/` and stored in a record `PCD`.

One can use the function `RedoPrecomputation` with one argument. In each case the result is written to one ore multiple files and will override the original precomputed data. The argument is a string and can be one of the following:

`"orbits"`: This will compute the 90 orbits of $\mathrm{Aut}(F_6)/U_3$ and the 86 orbits of $\mathrm{Aut}(F_4)/U_2$. This computation will take about 12 hours on an ordinary machine and has no progress bar.

`"goodpairs"`: First this will compute for each constraint $\gamma \in \mathfrak{R} \cup \mathfrak{R}^4$ the set of all $q \in G'/K'$ such that (q, γ) is a good pair.

Then it computes for each good pair (q, γ) one $\gamma' \in \Gamma_q(\gamma)$ with decorated $X = Y_{6,1}$ or $X = Y_{4,1} \in S$ as defined in equation (2) which fulfills depending whether $\gamma \in \mathfrak{R}^4_{\mathrm{act}}$ or $\gamma \in \mathfrak{R}_{\mathrm{act}}$ either Proposition 4.15 or Proposition 4.25. This computation will take about half an hour on ordinary machines and is equipped with a progress bar.

Afterwards the succeeding pairs of $(\mathbb{1}, \mathbb{1})$ which are needed for Corollary 4.27 are computed.

`"conjugacywidth"`: Denote by $\mathcal{E}_g$ the equation $a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}ag$. For each $\tau(g) = q \in G'/K'$ this will compute a constraint $\gamma\colon F_5 \to Q$ for the equations $\mathcal{E}_g$ and a constraint $\gamma'\colon F_4 \to Q$ such that $(\gamma * \pi)(\mathcal{E}_g) = \mathbb{1}$,

$$\mathcal{E}'_g := \mathfrak{nf}(\tilde{\Phi}_\gamma(\mathcal{E}_g)) = [X_1, X_2][X_3, X_4](g@2)(g@1),$$

and $(\mathcal{E}'_g, \gamma')$ is a good pair for all g with $\tau(g) = q$.

The computation will take about one hour and is equipped with a progress bar.

`"charactertable"`: This will compute the character table of the 4-th level germgroup and the set of irreducible characters. As the germgroup is quite large, this will take about 3 hours. There is no kind of progress bar.

`"noncommutator"`: Inside the 4-th level germgroup there is an element which is not a commutator but in the commutator subgroup. Since this group is finite we could in principle search by brute force for a commutator. Luckily there are only 3106 irreducible characters in this group and therefore we can use Burnside's formula (1.1). The search will almost immediately give a result. Most of the computation time is used to assert that the found element is indeed not a commutator.

The element is then lifted to its preimage in G with a minimal number of states.

Checking the assertion will take approximately 3 hours and is equipped with a progress bar.

`"all"`: This will do all of the above one after another.

To recompute the orbits or the charactertable GAP should be started with the `-o` flag to provide enough memory for the computation. For example start GAP by `gap -o 8G verify.g`

6.3. Implementation details.

6.3.1. *Reduced Constraints.* The proof of Lemma 4.3 in [LMU16] provides a constructive method to reduce any constraint to one with support only in the first five variables. We have implemented this in the function `ReducedConstraint` in the file `gap/functionsFR.g`.

It uses that the quotient $Q = G/K$ is a polycyclic group with

$$C_0 = Q = \langle \pi(a), \pi(b), \pi(d) \rangle\,, \qquad C_1 = \langle \pi(a), \pi(d) \rangle\,, \qquad C_2 = \langle \pi(ad) \rangle\,.$$

We take the generators of U_n as given in the proof of Lemma 4.4 plus additional ones which switch two neighboring pairs:

$$s_i\colon \begin{array}{rl} X_i \mapsto & X_{i+2} \\ X_{i+1} \mapsto & X_{i+3} \\ X_{i+2} \mapsto & X_i^{[X_{i+2},X_{i+3}]} \\ X_{i+3} \mapsto & X_{i+1}^{[X_{i+2},X_{i+3}]} \end{array} \quad \text{for } i = 1, 3, \ldots, 2n-3.$$

It can easily be checked, that these are also contained in U_n. These elements are used to reduce a given constraint in a form of a list with entries in Q to a list where all entries with index larger then 5 are trivial. This constraint can then be further reduced by a lookup table for the orbits of $\mathrm{Aut}(F_6)/U_3$.

If the file `verify.g` is loaded in a GAP environment with the FR package available the function `ReducedConstraint` can be used as an alias to get reduced constraints. For example:

```
    gap> f1 := Q.3;
    gap> gamma:= [f1,f1,f1,f1,f1,f1];
    gap> constr := ReducedConstraint(gamma);;
    gap> Print(constr.constraint);

[ <id>, <id>, <id>, <id> , f1, <id>]
```

6.3.2. *Good pairs.* For $g \in G$ and a constraint γ the question whether (g, γ) is a good pair depends only on the image of g in G/K' and the representative of $\gamma \in \mathfrak{R}$. (See Section 4.1.) So this is already a finite problem.

Given a given constraint γ, to obtain all q which form a good pair we can enumerate all possible commutators $[r_1, r_2][r_3, r_4][r_5, r_6]$ with $r_i \in \rho^{-1}(\gamma(X_i))$. Since $|K/K'| = 64$, it would take too much time to consider all combinations at once; thus the possible values for $[r_1, r_2]$ are computed and in a second step triple products of those elements are enumerated. This is implemented in the function `goodPairs` in the file `gap/functions.g`.

6.3.3. *Successors.* The key ingredient for the proof of Theorem A is Proposition 4.15. The main computational effort there is to compute the sets $\Gamma_q(\gamma)$ and find good pairs inside them.

This is implemented exactly as explained in the construction of the map Γ_q in the function `GetSuccessor` in the file `gap/precomputeGoodPairs.g`. Given an element $q \in G'/K'$ and an active constraint γ this function returns a tuple (γ', X) with $\gamma \in \mathfrak{R}$ and X the decorated element $Y_{6,1}$ or $> Y_{4,1}$ depending if $\gamma \in \mathfrak{R}^4$ or $\gamma \in \mathfrak{R}$.

Given an inactive constraint γ it returns a pair of constraints γ_1, γ_2 such that both have nontrivial activity and with ω the map from the branch structure it holds: $\omega(\langle\!\langle \gamma_1(X_i), \gamma_2(X_i) \rangle\!\rangle) = \gamma(X_i)$.

If the FR package is available the function `GetSuccessorLookup` can be used to explore the successors of elements. It returns the succeeding pair. For example

```
gap> f4 := Q.1;
gap> gamma:= [f4,f4,f4,f4,f4,f4];;
gap> g := (a*b)^8;;
gap> IsGoodPair(g,gamma);
```

```
true
```

```
gap> suc := GetSuccessorLookup(g,gamma);;
gap> suc[1];
```

```
<Trivial Mealy element on alphabet [ 1 .. 2 ]>
```

```
gap> suc[2].constraint;
```

```
[ <id>, <id>, <id>, <id>, f1*f3, <id> ]
```

References

[Bar98] Laurent Bartholdi, *The growth of Grigorchuk's torsion group*, Internat. Math. Res. Notices (1998), no. 20, 1049–1054. MR 1656258

[Bar13] ———, *Representation zeta functions of self-similar branched groups*, ArXiv e-prints (2013).

[Bar16a] ———, *Algorithmic decidability of Engel's property for automaton groups*, Computer science—theory and applications, Lecture Notes in Comput. Sci., vol. 9691, Springer, [Cham], 2016, pp. 29–40. MR 3533843

[Bar16b] ———, *FR, computations with functionally recursive groups, Version 2.3.6*, `https://www.gap-system.org/Packages/fr.html`, Apr 2016, GAP package.

[BG02] Laurent Bartholdi and Rostislav I. Grigorchuk, *On parabolic subgroups and Hecke algebras of some fractal groups*, Serdica Math. J. **28** (2002), no. 1, 47–90. MR 1899368

[BGŠ03] Laurent Bartholdi, Rostislav I. Grigorchuk, and Zoran Šunić, *Branch groups*, Handbook of algebra, Vol. 3, Handb. Algebr., vol. 3, Elsevier/North-Holland, Amsterdam, 2003, pp. 989–1112. MR 2035113

[BH16] Laurent Bartholdi and René Hartung, *Gap package LPRES, Version 0.3.0*, `https://laurentbartholdi.github.io/lpres`, Mai 2016, GAP package.

[Bur55] William Burnside, *Theory of groups of finite order*, Dover Publications, Inc., New York, 1955, 2d ed. MR 0069818

[Cal09] Danny Calegari, *scl*, MSJ Memoirs, vol. 20, Mathematical Society of Japan, Tokyo, 2009. MR 2527432

[CE81] Leo P. Comerford, Jr. and Charles C. Edmunds, *Quadratic equations over free groups and free products*, J. Algebra **68** (1981), no. 2, 276–297. MR 608536

[EHN81] David Eisenbud, Ulrich Hirsch, and Walter Neumann, *Transverse foliations of Seifert bundles and self-homeomorphism of the circle*, Comment. Math. Helv. **56** (1981), no. 4, 638–660. MR 656217

[Fin14] Elisabeth Fink, *Conjugacy growth and width of certain branch groups*, Internat. J. Algebra Comput. **24** (2014), no. 8, 1213–1231. MR 3296364

[FM11] Benson Farb and Dan Margalit, *A primer on mapping class groups*, Princeton University Press, 2011.

[GAP14] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.7.5*, 2014.

[Gri80] Rostislav I. Grigorchuk, *On Burnside's problem on periodic groups*, Funktsional. Anal. i Prilozhen. **14** (1980), no. 1, 53–54. MR 565099

[Gri83] ———, *On the Milnor problem of group growth*, Dokl. Akad. Nauk SSSR **271** (1983), no. 1, 30–33. MR 712546

[GS83] Narain Gupta and Saï d Sidki, *On the Burnside problem for periodic groups*, Math. Z. **182** (1983), no. 3, 385–388. MR 696534

[Gur80] Robert M. Guralnick, *Expressing group elements as commutators*, Rocky Mountain J. Math. **10** (1980), no. 3, 651–654. MR 590227

[GW03] Rostislav I. Grigorchuk and John S. Wilson, *A structural property concerning abstract commensurability of subgroups*, J. London Math. Soc. (2) **68** (2003), no. 3, 671–682. MR 2009443

[LMU16] Igor Lysenok, Alexei Miasnikov, and Alexander Ushakov, *Quadratic equations in the Grigorchuk group*, Groups Geom. Dyn. **10** (2016), no. 1, 201–239. MR 3460336

[LOST10] Martin W. Liebeck, Eamonn A. O'Brien, Aner Shalev, and Pham Huu Tiep, *The Ore conjecture*, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 4, 939–1008. MR 2654085

[Nek05] Volodymyr Nekrashevych, *Self-similar groups*, Mathematical Surveys and Monographs, vol. 117, American Mathematical Society, Providence, RI, 2005. MR 2162164

[Ore51] Oystein Ore, *Some remarks on commutators*, Proc. Amer. Math. Soc. **2** (1951), 307–314. MR 0040298

[Rhe68] A. H. Rhemtulla, *A problem of bounded expressibility in free products*, Mathematical Proceedings of the Cambridge Philosophical Society **64** (1968), no. 3, 573–584.

[Rhe69] ———, *Commutators of certain finitely generated soluble groups*, Can. J. Math. **21** (1969), no. 5, 1160–1164.

[Roz93] Alexander V. Rozhkov, *Centralizers of elements in a group of tree automorphisms*, Izv. Ross. Akad. Nauk Ser. Mat. **57** (1993), no. no. 6, 82–105. MR 1256568

[Seg09] Dan Segal, *Words. notes on verbal width in groups*, London Mathematical Society Lecture Note Series 361, Cambridge University Press, 2009.