Groups Geom. Dyn. 16 (2022), 493–522
DOI 10.4171/GGD/666


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

We compute, in this article, the commutator width of the *first Grigorchuk group* $G$; see Section 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 the following theorem.

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

It was already proven in [22] that the commutator width of $G$ is finite, without providing an explicit bound. Our result also answers a question of Elisabeth Fink [13, Question 3]. This is closely related to the problem of representing elements of the first Grigorchuk group by products of conjugates; see [13].

**Corollary B.** *Every element of $G'$ is a product of* 6 *conjugates of the generator $a$ and there are elements $g \in G'$ that are not products of* 4 *conjugates of $a$. Every element of $G$ is a product of at most* 8 *conjugates of the standard generators $\{a, b, c, d\}$.*

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 the following theorem.

**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.*

---

*2020 Mathematics Subject Classification.* Primary 20F65; Secondary 20F10, 20E08, 68W30.
*Keywords.* Self-similar groups, quadratic equations, commutator length.

### 1.1. Commutator width

Let $\Gamma$ be a group. It is well known that usually elements of $\Gamma'$ 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 [26].

On the other hand, some classes of groups have finite commutator width: finitely generated virtually abelian-by-nilpotent groups [29], and finitely generated solvable groups of class 3; see [27].

Finite groups are trivial examples of groups of finite commutator width. There are finite groups in which some elements of the derived subgroup are not commutators, the smallest having order 96; see [20]. On the other hand, non-abelian finite simple groups have commutator width 1, as was conjectured by Ore in 1951 (see [25]) and proven in 2010 (see [21]). The commutator width cannot be bounded among finite groups; for example,

$$\Gamma_n = \big\langle x_1, \dots, x_{2n} \mid x_1^p, \dots, x_{2n}^p, \gamma_3\big(\langle x_1, \dots, x_{2n}\rangle\big)\big\rangle$$

is a finite class-2 nilpotent group in which $\Gamma_n'$ has order $p^{\binom{2n}{2}}$ but at most $\binom{p^{2n}}{2}$ elements are commutators, so the commutator width of $\Gamma_n$ 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 [9]. It is also related to solvability of quadratic equations in groups: a group $\Gamma$ 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 [22] 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 $\Gamma$ is computable, then it may be used to compute the commutator width: Burnside shows (or, rather, hints) in [8, §238, Ex. 7] that an element $g \in \Gamma$ may be expressed as a product of $r$ commutators if and only if

$$\sum_{\chi \in \operatorname{Irr}(\Gamma)} \frac{\chi(g)}{\chi(1)^{2r-1}} > 0. \tag{1.1}$$

This may yield another proof of Theorem A, using the quite explicit description of $\operatorname{Irr}(G)$ given in [4].

Consider a group $\Gamma$ and a subgroup $\Delta$. There is in general little connection between the commutator width of $\Gamma$ and that of $\Delta$. If $\Delta$ has finite commutator width and $[\Gamma : \Delta]$ is finite, then obviously $\Gamma$ also has finite commutator width – for example, because $\Gamma/\operatorname{core}(\Delta)'$ is virtually abelian, and every commutator in $\Gamma$ can be written as a product of a commutator in $\Delta$ 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 $\Delta$ of orientation-preserving self-homeomorphisms of $\mathbb{R}$ that commute with integer translations, and let $\Gamma$ be the extension of $\Delta$ by the involution

$x \mapsto -x$. Then, by [11, Theorems 2.3 and 2.4], every element of $\Gamma' = \Delta$ is a commutator in $\Gamma$, while the commutator width of $\Delta$ is infinite.

Both $\Gamma$ and $\Delta$ 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 [18] and some of its properties. For a more detailed introduction into the topic of self-similar groups we refer to [6, 24] and to Section 3.

A *self-similar group* is a group $\Gamma$ 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 $\pi$ its *activity*. It is also convenient to identify, in a self-similar group, elements with their image under $\Psi$.

Note that, by definition, every group may be viewed as self-similar; the self-similarity is an attribute of a group, not a property. However, being regular branched with $n > 1$ imposes strong conditions on the group.

A self-similar group may be specified by giving a set $S$ of generators, some relations $\mathcal{R}$ that they satisfy, and defining $\Psi$ on $S$. There is then a maximal quotient $\Gamma$ of the group $\langle S \mid \mathcal{R} \rangle$ on which $\Psi$ 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 [2]. It is a group of intermediate word growth [16], and answered in this manner a celebrated question of Milnor. For more information about the Grigorchuk group, see the extensive survey [15].

We have decided to concentrate on the first Grigorchuk group in the computational part 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 [19].

### 1.3. Sketch of proofs

The general idea for the proof of Theorem A is the decomposition of group elements into states via $\Psi$. 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 $\sigma_i$. 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) yield 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 belong to 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 [14]. The source code for these computations can be accessed by following this article's DOI. 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 [3] and LPRES [7] 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 $\mathrm{Var}(\mathcal{E}) \to G$; by a microscopic abuse of notation we refer to such a map as a solution as well. The question of whether an equation $\mathcal{E}$ is solvable will be referred to as the *Diophantine problem* of $\mathcal{E}$.

Every homomorphism $\varphi: G \to H$ extends uniquely to an $F_{\mathcal{X}}$-homomorphism

$$\varphi_*: F_{\mathcal{X}} * G \to F_{\mathcal{X}} * H.$$

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 $\varphi$ of $F_{\mathcal{X}} * G$ that maps $\mathcal{E}$ to $\mathcal{F}$. We denote by $\mathrm{Stab}(\mathcal{E})$ the group of all $G$-automorphisms that fix $\mathcal{E}$.

**Lemma 2.3.** *Let $\mathcal{E}$ be an equation and let $\varphi$ 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 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 $m, n \geq 0$, distinct $X_i, Y_i, Z_i \in \mathcal{X}$, and $c_i \in G$, the following two kinds of equations are called in *normal form*:

$$\mathcal{O}_{n,m}: [X_1, Y_1][X_2, Y_2] \cdots [X_n, Y_n] c_1^{Z_1} \cdots c_{m-1}^{Z_{m-1}} c_m, \tag{2.1}$$

$$\mathcal{U}_{n,m}: X_1^2 X_2^2 \cdots X_n^2 c_1^{Z_1} \cdots c_{m-1}^{Z_{m-1}} c_m. \tag{2.2}$$

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** ([10]). *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 as follows.

**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 $\varphi$. 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: $v \in \mathcal{X} \cup \mathcal{X}^{-1}$.** 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 $uw_1w_2$, has fewer variables. Thus, composition with the corresponding $G$-isomorphism results in a normal form.

**Case 2: Length($v$) > 1.** In this case, $v$ is a word consisting of elements from $\mathcal{X} \cup \mathcal{X}^{-1}$ 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, $\mathcal{E} = u_1Y^{-1}u_2X^{-1}Yv_1Xw$ and 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 $\varphi$ 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

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

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

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

Note that $s \in \mathcal{O}_{n,m}$. If $n \geq 1$, then this procedure can be repeated with $Z^2s$ in place of $X^2r$. ∎

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 [22]). Given an equation $\mathcal{E} \in F_{\mathcal{X}} * G$, a group $H$ with a fixed homomorphism $\pi: G \to H$ and a homomorphism $\gamma: F_{\mathcal{X}} \to H$, the pair $(\mathcal{E}, \gamma)$ is called a *constrained* equation and $\gamma$ is called a *constraint* for the equation $\mathcal{E}$ on $H$.

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

We note that the constraint $\gamma$ needs only to be specified on $\mathrm{Var}(\mathcal{E})$, just as per our convention the solution $s$ need only be defined on $F_{\mathrm{Var}(\mathcal{E})}$. In practice, $H$ will be a quotient of $G$, and the constraints $\gamma$ specify to which coset of $\ker \pi$ the variables have to belong.

## 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 $\mathrm{Aut}(T_n)$ 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 neighbours of the root. Every $g \in \mathrm{Aut}(T_n)$ permutes the $T_{i,n}$ by a permutation $\sigma$ and simultaneously acts on each of them by isomorphisms $g_i: T_{i,n} \to T_{i^\sigma,n}$.

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 $\mathrm{Aut}(T_n)$, and obtain in this manner an isomorphism

$$\Psi: \begin{cases} \mathrm{Aut}(T_n) \xrightarrow{\sim} \mathrm{Aut}(T_n) \wr S_n, \\ \qquad g \mapsto \langle\!\langle g_1, \dots, g_n \rangle\!\rangle \sigma. \end{cases}$$

A *self-similar group* is a subgroup $G$ of $\mathrm{Aut}(T_n)$ satisfying $\Psi(G) \le G \wr S_n$. 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 $\mathrm{act}(g) := \sigma$.

### 3.1. Commutator width of 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 $\mathrm{Aut}(T_2)$, and show that every element of its derived subgroup is a commutator. This has seemingly been proven many times; the earliest reference we are aware of is [23]. In this group, we have the following useful property: for every two elements $g, h \in \mathrm{Aut}(T_n)$, 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** (Muntyan [23, Theorem 1]). *The commutator width of* $\mathrm{Aut}(T_2)$ *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 $\langle\langle X_1^{-1}Y_2^{-1}X_2Y_2(g@1), X_2^{-1}Y_1^{-1}X_1Y_1(g@2)\rangle\rangle$. Therefore, $[X, Y]g$ is solvable if the system of equations $\{X_1^{-1}Y_2^{-1}X_2Y_2(g@1), X_2^{-1}Y_1^{-1}X_1Y_1(g@2)\}$ is solvable. We apply the $\mathrm{Aut}(T_2)$-homomorphism $X_1 \mapsto X_1$, $X_2 \mapsto Y_1^{-1}X_1Y_1(g@2)$, $Y_i \mapsto Y_i$ 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:

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

and for all $i \geq 1$

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

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 [18] is a finitely generated self-similar group $G$ 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

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

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

**Lemma 4.2** ([28]). *The Grigorchuk group $G$ is regular branched with branching subgroup*

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

*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 that fix $\mathcal{E}$. Recall also that $G$-automorphisms are homomorphisms of $F_{\mathcal{X}} * G$ that fix $G$; in what follows automorphisms of $F_{2n}$ will be considered without warning as automorphisms of $F_{2n} * G$ that fix $G$.

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

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

**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 group $M(n, 0)$ of the closed orientable surface of genus $n$. It can be checked that the automorphisms $\varphi_i$ and $\psi_i$ represent the Humphries generators of $M(n, 0)$. For details on mapping class groups, see for example [12].

**Lemma 4.3** ([22]). *Given $n \in \mathbb{N}$ and a homomorphism $\gamma\colon F_{\mathcal{X}} \to Q$ with* $\mathrm{supp}(\gamma) \subset \langle X_1, \ldots, X_{2n}\rangle$*, there is an element $\varphi \in U_n <$* $\mathrm{Aut}(F_{\mathcal{X}})$ *such that*

$$\mathrm{supp}(\gamma \circ \varphi) \in \langle X_1, \ldots, X_5\rangle.$$ ∎

We may naturally identify the set $\{\gamma\colon F_{\mathcal{X}} \to Q \mid \mathrm{supp}(\gamma) \subset \langle X_1, \ldots, X_n\rangle\}$ with $Q^n$, by restricting such a $\gamma$ to $\{X_1, \ldots, X_n\}$. We then have the following lemma.

**Lemma 4.4.** $|Q^{2n}/U_n| \le 90$ *for all* $n \ge 3$.

*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 $Q^5$, so to determine $|Q^{2n}/U_n|$ it suffices to determine how many orbits of $U_n$ in $Q^{2n}$ intersect $Q^5$ nontrivially. Then since $U_n \subset U_{n+1}$, we have $|Q^{2n+2}/U_{n+1}| \le |Q^{2n}/U_n|$ for all $n \ge 3$, and direct computation gives $|Q^6/U_3| = 90$; see Section 6.3. ∎

**Remark.** It can be proved by an extra computation that indeed $|Q^{2n}/U_n| = 90$ for all $n \ge 3$.

**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 $\operatorname{supp}(\gamma) \subset \langle X_1, \dots, X_5\rangle$ that is a representative system of the orbits $Q^{2n}/U_n$ for each $n \ge 3$. Note that representatives are formally not assumed unique if $n \ge 4$ (though in fact they are unique according to the remark above). 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 $\varphi_\gamma$ the $G$-automorphism in $U_n$ such that $\gamma \circ \varphi_\gamma \in \mathfrak{R}$.

The element $\gamma \circ \varphi_\gamma$ will be called a *reduced constraint*, denoted by $\operatorname{red}(\gamma)$. We extend the function $\operatorname{red}(*)$ also to the case when $\gamma : F_{\mathcal{S}} \to Q$ is a homomorphism defined on any finite subset $\mathcal{S}$ of variables from $\mathcal{X}$: we simply extend $\gamma$ to $F_{\mathcal{X}}$ by defining $\gamma(X) = \mathbb{1}$ for $X \notin \mathcal{S}$ and then take $\operatorname{red}(\gamma)$ as already defined.

**Remark.** Considering finitely supported constraints defined on an infinite set of variables, $\mathcal{X}$ is a convenient trick that allows us to compare constraints intended for equations with different number of variables. In particular, we will assert in Section 4.2 that certain sets of constraints are independent of the number of variables of an equation.

**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 $(R_n g, \gamma \circ \varphi_\gamma)$ and vice versa. ∎

**Definition 4.7** (Branch structure [4]). 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 [4, 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 $\gamma$ 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 unsolvable for all $n$ 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 $(g, \gamma)$ is called a *good pair* if $(R_n g, \gamma)$ is solvable for some $n \in \mathbb{N}$.

**Lemma 4.9.** *For $g \in G$, let $\bar{g}$ denote the image of $G$ in $G/K'$. Then the pair $(g, \gamma)$ is a good pair if and only if $(R_3\bar{g}, \gamma)$ has a solution in $G/K'$.*

*Proof.* If $(g, \gamma)$ is a good pair and $s$ a solution for $(R_n g, \gamma)$, then $s(X_i) \in K$ for all $i \geq 6$, so $s(R_n g) = s(R_3) \cdot kg$ for some $k \in K'$. Therefore, $\overline{s(R_3)}\bar{k}\bar{g} = \overline{s(R_3)}\bar{g} = \mathbb{1}$, and $(R_3\bar{g}, \gamma)$ has a solution $\bar{s}\colon X_i \mapsto \overline{s(X_i)}$. Clearly, $\bar{s}$ satisfies the constraint $\gamma$, since $G \twoheadrightarrow Q$ factorizes through $G/K'$.

Now suppose that $(R_3\bar{g}, \gamma)$ has a solution $\bar{s}\colon F_{\mathfrak{X}} \to G/K'$, so we have $\bar{s}(R_3\bar{g}) = \bar{s}(R_3)\bar{g} = \mathbb{1}$ in $G/K'$. There are then $k \in K'$ and $g_1, \ldots, g_6 \in G$ with $R_3(g_1, \ldots, g_6)kg = \mathbb{1}$, so $(g, \gamma)$ is a good pair. ∎

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

**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 $(g, \gamma)$ 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 $G$ for some $n$ not depending on $k$; and the constraints ensure that it is solvable in $K$. Therefore, the commutator width of $K$ is at most $n$.

(a)⇒(b): if $(g, \gamma)$ 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 $n = m + 3$. ∎

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*

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

*Furthermore, these groups form a tower with indices*

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

*Proof.* The chain of indices is shown for example in [6] and the generating sets can be verified using the GAP standard methods `NormalClosure` and `Index` within FR. ■

### 4.2. Succeeding pairs

The main step in our proof is a procedure that accepts as input a good pair $(g, \gamma)$ and produces a "succeeding pair" $(g', \gamma')$ in such a manner that solvability of $(R_n g, \gamma)$ is equivalent to that of $(R_{n'} g', \gamma')$ for some $n' > n$. The procedure, while completely explicit (and actually implemented) is quite complicated, and involves the construction (via sets $\Gamma_1^{\cdots}$, $\Gamma_2^{\cdots}$, $\Gamma_3^{\cdots}$, and $\Gamma_4^{\cdots}$) of a non-empty set $\Gamma^q(\gamma)$ of admissible succeeding pairs, with $q \in G'/K'$ representing $g$, from which $\gamma'$ will be appropriately chosen. The reader is forewarned that this section is the most technical, and is encouraged to return to the general idea of the proof of Theorem A given in Section 1.3.

**Definition 4.12** ($\mathfrak{R}_{\mathrm{act}}$, active constraints). We define the activity 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 $\gamma: F_{\mathcal{X}} \to Q$. Define $\mathrm{act}(\gamma): F_{\mathcal{X}} \to C_2$ by $X \mapsto \mathrm{act}(\gamma(X))$.

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

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

*Proof.* This is a finite problem which can be checked in GAP by means of the function `verifyLemmaExistGoodConstraints`. For more details, see Section 6.1. ■

We will now give a procedure that starts with a constrained equation of class $\mathcal{O}_{n,1}$ and produces a finite family of constrained equations of class $\mathcal{O}_{2n-1,1}$. Because we need to specialize equations inside the procedure, the reduction is one-sided: solvability of any equation of the family implies solvability of the initial one. We provide a weak form of reduction in the reverse direction in Proposition 4.15; this will be enough for our purposes (and actually implies equivalence of the initial equation and the produced family).

The idea of the procedure is to replace each variable $X_\ell$ of the starting equation $(R_n g, \gamma)$ by two variables $Y_{\ell,1}$ and $Y_{\ell,2}$ representing the states of $X_\ell$, so

$$X_\ell = \langle\langle Y_{\ell,1}, Y_{\ell,2}\rangle\rangle \operatorname{act}(X_\ell),$$

and then transform the resulting system of two equations to a quadratic equation in the standard form. We denote by $\mathcal{Y} = \{Y_{\ell,i} \mid \ell \geq 1,\ i = 1,2\}$ the target set of variables and by $F_\mathcal{Y}$ the free group with basis $\mathcal{Y}$. After the transformation, we obtain a set of equations of the form $(R_{2n-1}g', \gamma')$, where $(g', \gamma')$ runs over a certain finite set and $R_{2n-1}$ is written in a subset of variables

$$\mathcal{V} = \{Y_{\ell,i} \mid 1 \leq \ell \leq 2n,\ \ell \neq 6,\ i = 1,2\}.$$

In what follows, we will define for all $q \in G'/K'$ a map $\Gamma^q$ which sends constraints to finite sets of constraints,

$$\Gamma^q\colon (\gamma\colon F_\mathcal{X} \to Q) \mapsto \big\{(\gamma'\colon F_\mathcal{Y} \to Q)\big\},$$

with the following property:

(∗) for every constraint $\gamma' \in \Gamma^q(\gamma)$, there is $x \in \{\mathbb{1}, a, b, c, d, ab, ad, ba\} \subset G$ such that if $gK' = q$, then for any $n \geq 3$ the equation $(R_n g, \gamma)$ is solvable as soon as the constrained equation $(R_{2n-1}(g@2)^x \cdot g@1, \gamma')$ is solvable.

We will define this map $\Gamma^q$ in several steps and afterwards show that for all good pairs $(q, \gamma)$ and all $g$ with $gK' = q$, there is some constraint $\gamma' \in \Gamma^q(\gamma)$ such that $((g@2)^x \cdot g@1, \gamma'|_{F_\mathcal{V}})$ is a good pair. The first step is to construct a set $\Gamma_1(\gamma)$ of constraints on the "doubled" alphabet $\mathcal{Y}$, by decomposing the constraints expressed by $\gamma$ on the variables $\mathcal{X}$ into constraints on their states $\mathcal{Y}$. The second step extracts from $\Gamma_1(\gamma)$ a subset $\Gamma_2^q(\gamma)$ of constraints compatible with the target $q$, in the sense that they are suitable to solving an equation $R_n q = \mathbb{1}$ in $G/K'$. The third step rewrites elements of $\Gamma_2^q(\gamma)$ in normal form using a letter $Y_0$, defining a set $\Gamma_3^{q,Y_0}(\gamma)$ of reduced constraints. The fourth step combines these constraints over all possible $Y_0$ into a set $\Gamma_4^q(\gamma)$, and the last step extracts from $\Gamma_4^q(\gamma)$ a set $\Gamma^q(\gamma)$ by requiring some activity to be non-trivial and lie in a specific subset of $Q$.

We assume that some $n \geq 3$ is fixed. It will be straightforward to see from the construction at each step that the corresponding set $\Gamma_i^{\cdots}$ does not depend on $n$.

For the first step, we take the branching structure $(K, Q, \pi, Q_1, \omega)$ of the Grigorchuk group. Set

$$\Gamma_1(\gamma) = \big\{\gamma' : F_\mathcal{Y} \to Q \mid \omega\big(\langle\langle \gamma'(Y_{\ell,1}), \gamma'(Y_{\ell,2})\rangle\rangle \operatorname{act}(X_\ell)\big) = \gamma(X_\ell) \text{ if } 1 \leq \ell \leq 6,\\ \gamma'(Y_{\ell,1}) = \gamma'(Y_{\ell,2}) = \mathbb{1} \text{ if } \ell > 6\big\}.$$

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

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

**Lemma 4.14.** *If $\gamma$ is a constraint with nontrivial activity, and $\Phi_\gamma(R_n \mathfrak{g}) = \langle\!\langle v, w \rangle\!\rangle$, then* $\operatorname{Var}(v) \cap \operatorname{Var}(w) \neq \emptyset$. *Moreover, for any $Y_0 \in \operatorname{Var}(v) \cap \operatorname{Var}(w)$ we have $v = v_1 Y_0^{\varepsilon} v_2$ and $w = w_1 Y_0^{-\varepsilon} w_2$, for some $\varepsilon = \pm 1$.*

*Proof.* Let $\ell \in \{1 \cdots 2n\}$ 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 $\sigma$ be the activity of $\gamma(X_k)$. Then $\Phi_\gamma(R_n \mathfrak{g})$ contains a factor

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

In both cases, $Y_{k,1}, Y_{k,2} \in \operatorname{Var}(v) \cap \operatorname{Var}(w)$. If $R_n$ has a factor $[X_k, X_{k+1}]$ with both activities of $\gamma(X_k)$ and $\gamma(X_{k+1})$ trivial, then it contributes a factor $\langle\!\langle u_1, u_2 \rangle\!\rangle$ in $\Phi_\gamma(R_n \mathfrak{g})$ with $\operatorname{Var}(u_1) \cap \operatorname{Var}(u_2) = \emptyset$. This implies the second statement. ∎

Note that the map $G' \to G \times G$ factors through $K'$ to a map $\lambda \colon G'/K' \to Q \times Q$, since $K' \leq K \times K$. Given $q \in G'/K'$, write $\lambda(q) = \langle\!\langle q_1, q_2 \rangle\!\rangle \in Q \times Q$ and define $\theta_q \colon F_{\mathcal{H}} \to Q$ by $\mathfrak{g}_i \mapsto q_i$ for $i = 1, 2$; if $\gamma' \colon F_{\mathcal{Y}} \to Q$ is a constraint, we denote by $\gamma' * \theta_q$ the natural map $F_{\mathcal{Y}} * F_{\mathcal{H}} \to Q$ agreeing with $\theta_q$ and $\gamma'$ on the respective factors, and by $(\gamma' * \theta_q)^2$ the induced map $(F_{\mathcal{Y}} * F_{\mathcal{H}}) \wr C_2 \to Q \wr C_2$. Then define the following subset of $\Gamma_1(\gamma)$:

$$\Gamma_2^q(\gamma) = \bigl\{\gamma' \in \Gamma_1(\gamma) \mid (\gamma' * \theta_q)^2\bigl(\Phi_\gamma(R_n \mathfrak{g})\bigr) = \langle\!\langle \mathbb{1}, \mathbb{1} \rangle\!\rangle\bigr\}. \tag{4.1}$$

Note that the set $\Gamma_2^q(\gamma)$ is nonempty provided that $(q, \gamma)$ is a good pair.

For $\gamma \in \mathfrak{R}_{\text{act}}$ denote by $v$ and $w$ the elements of $F_{\{Y_{1,1},\ldots,Y_{6,2}\}}$ 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, since the variables $X_7, \ldots, X_{2n}$ have trivial activity,

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

By Lemma 4.14, take any $Y_0 \in \{Y_{1,1}, \ldots, Y_{6,2}\}$ with $v = v_1 Y_0^{\varepsilon} v_2$ and $w = w_1 Y_0^{-\varepsilon} w_2$. Next, we improve the form of $\Phi_\gamma(R_n(X_*)\mathfrak{g})$, without affecting the constraint $\gamma'$, by means of the $F_{\mathcal{H}}$-homomorphism

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

that eliminates the variable $Y_0$. It maps the second coordinate of $\Phi_\gamma(R_n(X_*)\mathfrak{g})$ to $\mathbb{1}$ and the first coordinate to a quadratic $G$-equation over $F_{\mathcal{H}}$:

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

Moreover, from $\gamma' \in \Gamma_2^q(\gamma)$ we get

$$\begin{aligned}\mathbb{1} &= (\gamma' * \theta_q)\big(wR_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}\big)\\ &= (\gamma' * \theta_q)\big(w_1 Y_0^{-1} w_2 R_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}\big).\end{aligned}$$

Thus we obtain

$$(\gamma' * \theta_q)\big(w_2 R_{n-3}(Y_{7,2}, \dots, Y_{2n,2})\mathfrak{g}\, w_1\big) = (\gamma' * \theta_q)(Y_0)$$

and $\ell_{Y_0}$ does not affect the constraint $\gamma'$.

From this we conclude that $(\gamma' * \theta_q)(Y_0) = (\gamma' * \theta_q)(\ell_{Y_0}(Y_0))$. Since $\ell_{Y_0}$ fixes all $Y \neq Y_0$, we see that in fact

$$\gamma' * \theta_q = (\gamma' * \theta_q) \circ \ell_{Y_0} \quad \text{for all } \gamma' \in \Gamma_2^q(\gamma) \text{ with } \theta_q \colon \mathfrak{g}_i \mapsto q_i. \tag{4.2}$$

Consider the $F_{\mathcal{H}}$-automorphisms

$$\psi_1 \colon \begin{cases} F\mathcal{Y} * F_{\mathcal{H}} \to F\mathcal{Y} * F_{\mathcal{H}} & \\ Y_{k,1} \mapsto Y_{k,1}^{\mathfrak{g}_1^{-1}} & \text{for } k > 6,\\ Y_{k,2} \mapsto Y_{k,2}^{(\mathfrak{g}_2 w_1 v_2 \mathfrak{g}_1)^{-1}} & \text{for } k > 6,\\ Y_{k,\ell} \mapsto Y_{k,\ell} & \text{for } k \le 6,\ \ell = 1, 2,\end{cases}$$

$$\psi_2 \colon \begin{cases} F\mathcal{Y} * F_{\mathcal{H}} \to F\mathcal{Y} * F_{\mathcal{H}} & \\ Y_{k,1} \mapsto Y_{k,1}^{\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1} & \text{for } k > 6,\\ Y_{k,2} \mapsto Y_{k,2}^{\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1} & \text{for } k > 6,\\ Y_{k,\ell} \mapsto Y_{k,\ell} & \text{for } k \le 6,\ \ell = 1, 2,\end{cases}$$

and

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

For the equation $\mathcal{E}$ over $F_{\mathcal{H}}$, recall from Section 2.2 the $F_{\mathcal{H}}$-isomorphism $\mathfrak{nf}_{\mathcal{E}}$ that puts $\mathcal{E}$ in normal form. We have $\mathfrak{nf}_{\mathcal{E}} = \psi_3 \circ \psi_2 \circ \mathfrak{nf}_{v_1 w_2 \mathfrak{g}_2 w_1 v_2 \mathfrak{g}_1} \circ \psi_1$. Indeed, first $\psi_1$ groups the terms $v_1 w_2 \mathfrak{g}_2 w_1 v_2 \mathfrak{g}_1$ at the beginning; then $\mathfrak{nf}_{v_1 w_2 \mathfrak{g}_2 w_1 v_2 \mathfrak{g}_1}$ puts these terms into the form $[Y_{1,1}, Y_{1,2}] \cdots \mathfrak{g}_2^{Y_{6,1}} \mathfrak{g}_1$, and finally $\psi_2$ and $\psi_3$ reorder and renumber the variables.

Therefore,

$$\mathfrak{nf}_{\mathcal{E}}(\mathcal{E}) = R_{2n-1}\big(Y_{1,1}, Y_{1,2}, \ldots, \widehat{Y_{6,1}}, \widehat{Y_{6,2}}, \ldots, Y_{2n,2}\big)\mathfrak{g}_2^{Y_{6,1}}\mathfrak{g}_1.$$

This lets us define

$$\Gamma_3^{q,Y_0}(\gamma) = \big\{(\gamma' * \theta_q) \circ (\mathfrak{nf}_{\mathcal{E}}^{-1}|_{F_{\mathcal{Y}}}) : F\mathcal{Y} \to Q \mid \gamma' \in \Gamma_2^q(\gamma)\big\}.$$

Note that $\mathfrak{nf}_{\mathcal{E}}$ 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 $\gamma'' \in \Gamma_3^{q,Y_0}(\gamma)$ independently of $q_i$, $Y_0$, and $\gamma$. The set $\Gamma_3^{q,Y_0}$ is therefore independent of $n$ as soon as $n \ge 3$.

We will now show how to obtain a solution to the original constrained equation $(R_{2n-1}g, \gamma)$ from a solution of the obtained equation over $\mathcal{Y}$. Denote by $S$ the set $\{\mathbb{1}, a, b, c, d, ab, ad, ba\} \subset G$, and complete it to a transversal $S'$ of $K$ in $G$. For $q \in Q$, denote by $\operatorname{rep}(q) \in S'$ the coset representative of $q$. Given $g \in G'$, $g_i = g@i$ for $i = 1, 2$, an active constraint $\gamma \in \mathfrak{R}_{\mathrm{act}}$, and $\gamma'' \in \Gamma_3^{\pi(g_1),\pi(g_1),Y_0}(\gamma)$, a solution for the constrained $G$-equation

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

can be extended by the map $Y_{6,1} \mapsto \operatorname{rep}(\gamma''(Y_{6,1}))$ to a solution $s'$ of the $G$-equation $(R_{2n-1}(Y_{*,*})g_2^{Y_{6,1}}g_1, \gamma'')$. Consider the homomorphism $i_{\mathcal{H}}\colon F_{\mathcal{H}} \to G$, $\mathfrak{g}_i \mapsto g_i$ and note that since $\mathfrak{nf}_{\mathcal{E}}$ is an $F_{\mathcal{H}}$-homomorphism, the function $(\mathbb{1} * i_{\mathcal{H}}) \circ \mathfrak{nf}_{\mathcal{E}}$ maps $\mathcal{E}$ to $R_{2n-1}(Y_{*,*})g_2^{Y_{6,1}}g_1$. Moreover, by (4.2), we have $\gamma' := \gamma'' \circ (\mathbb{1} * i_{\mathcal{H}}) \circ \mathfrak{nf}_{\mathcal{E}} \in \Gamma_2^q(\gamma)$, so the map

$$s\colon Y_{i,j} \mapsto \begin{cases} w_2 g_2 w_1 & \text{if } i,j = 6,2, \\ s' \circ (\mathbb{1} * i_{\mathcal{H}}) \circ \mathfrak{nf}_{\mathcal{E}}(Y_{i,j}) & \text{otherwise} \end{cases}$$

is a solution for $((\mathbb{1} * i_{\mathcal{H}}) \circ \Phi_\gamma(R_n\mathfrak{g}), \gamma')$ and thus also for $(\widetilde{\Phi}_\gamma(R_n g), \gamma')$. By the definition of $\omega$, the element $t_i := \langle\langle s(Y_{i,1}), s(Y_{i,2})\rangle\rangle \operatorname{act}(X_i)$ belongs to $G$ for all $i$. Moreover, since $\gamma' \in \Gamma_1(\gamma)$, 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 $\Gamma_3^{q,Y_0}$ does depend on the choice of the variable $Y_0$. To remove this dependency, we observe that the set of all variables $Y_0 \in \operatorname{Var}(v) \cap \operatorname{Var}(w)$ does not depend on $n$ and define

$$\Gamma_4^q(\gamma) = \bigcup_{Y_0 \in \operatorname{Var}(v) \cap \operatorname{Var}(w)} \Gamma_3^{q,Y_0}(\gamma).$$

Filtering out those constraints that do not fulfill the requested properties, we finally define

$$\Gamma^q(\gamma) := \big\{\gamma' \in \Gamma_4^q(\gamma) \mid \operatorname{act}(\gamma')|_{\mathcal{V}} \ne \mathbb{1},\ \gamma'(Y_{6,1}) \in \pi(S)\big\}. \tag{4.3}$$

Note that $(*)$ holds automatically by construction. It is straightforward to check that the set $\Gamma^q(\gamma)$ does not depend on $n \ge 3$.

Now we track solutions of equations in the reverse direction.

**Proposition 4.15.** *For each good pair $(q,\gamma)$ with $q \in G'/K'$ and $\gamma \in \mathfrak{R}_{\mathrm{act}}$, the set $\Gamma^q(\gamma)$ contains some constraint $\gamma'$ such that for all $g \in G'$ with $gK' = q$ the pair*

$$\big((g@2)^{\mathrm{rep}(\gamma'(Y_{6,1}))} \cdot g@1, \mathrm{red}(\gamma'|_{F_\mathcal{V}})\big)$$

*is a good pair.*

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

**Lemma 4.16.** *For each $h \in G$ the map*

$$\bar{p}_h \colon \begin{cases} G'/K' \to G'/(K \times K), \\ \quad gK' \mapsto \big((g@2)^h \cdot g@1\big)(K \times K) \end{cases}$$

*is well defined.*

*Proof.* We first note that $k@i \in K \times K$ for $i = 1, 2$ and $k \in K'$, by the last line of Lemma 4.11. Define for $h \in G$ a map $p_h \colon G \to G$ by $g \mapsto (g@2)^h \cdot g@1$. This map is in general not a homomorphism, but by Lemma 3.2 and a simple commutator calculation we have $p_h(g) \in G'$ for all $g \in G'$, $h \in G$. For $k \in K'$, we then have

$$\begin{aligned} p_h(gk) &= \big((gk)@2\big)^h \cdot (gk)@1 \\ &= (g@2)^h \cdot (k@2)^h \cdot g@1 \cdot k@1 \in \big((g@2)^h \cdot g@1\big)(K \times K). \end{aligned}$$ ∎

*Proof of Proposition* 4.15. In the construction above, it is clear that the sets $\Gamma_3^{q,Y_0}$ and hence $\Gamma_4^q$ 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 $\mathrm{act}(\gamma')|_\mathcal{V} \neq \mathbb{1}$ (i.e., $\gamma' \in \Gamma^q(\gamma)$) is implemented in the procedure below.

By Lemma 4.16, we have a map $\bar{p}_h \colon G'/K' \to G'/(K \times K)$. For $g \in G'/K'$, let us denote by $\bar{g}$ the natural image of $g$ in $(G'/K')/((K \times K)/K') \simeq G'/(K \times K)$. 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 $g \mapsto \bar{g}$ form good pairs with $\mathrm{red}(\gamma'|_{F_\mathcal{V}})$. In formulas, letting $\mathcal{P}$ denote the predicate of being a good pair, what needs to be checked is

$$\begin{aligned} &\forall q \in G'/K' \ \forall \gamma \in \mathfrak{R}_{\mathrm{act}} \ \exists \gamma' \in \Gamma^q(\gamma) \ \forall r \in G'/K' \\ &\text{with } \bar{r} = \bar{p}_{\mathrm{rep}(\gamma'(Y_{6,1}))}(q) \colon \mathcal{P}(q,\gamma) \Rightarrow \mathcal{P}\big(r, \mathrm{red}(\gamma'|_{F_\mathcal{V}})\big). \end{aligned}$$

This last formula quantifies only over finite sets, and can be implemented. It appears in the GAP code as the function *verifyPropExistsSuccessor*. ∎

**Definition 4.17** (Succeeding pair). For each $q \in G'/K'$ and $\gamma \in \mathfrak{R}_{\mathrm{act}}$ such that $(q,\gamma)$ 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 $\gamma'|_{F_\mathcal{V}}$ by a reduced constraint $\gamma'_r$. Since $\mathrm{act}(\gamma')|_\mathcal{V} \neq \mathbb{1}$, we have $\gamma'_r \in \mathfrak{R}_{\mathrm{act}}$. 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* $(g_k, \gamma_k)$ of $(g,\gamma)$: $(g_0,\gamma_0) = (g,\gamma)$ and $(g_{k+1},\gamma_{k+1})$ is the succeeding pair of $(g_k,\gamma_k)$.

The following lemma illustrates the use of the construction.

**Lemma 4.18.** *Let $(g_k, \gamma_k)$ be the succeeding sequence of a good pair $(g, \gamma)$. 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 starting from index $i$ the succeeding sequence becomes periodic and hence $n'$ can be taken arbitrarily large. If $(g, \gamma)$ is a good pair, then $(g_i, \gamma_i)$ 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 three commutators

We will prove that every element $g \in G'$ is a product of three commutators by proving that all succeeding sequences $(g_k, \gamma_k)$ as defined in Definition 4.17 become periodic 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

$$\mathrm{Suc}_1^g = \{g\}, \quad \mathrm{Suc}_n^g = \left\{p_x(h) \mid h \in \mathrm{Suc}_{n-1}^g, x \in S\right\} \tag{4.4}$$

stabilizes on a finite set.

In [1], there is a choice of weights on generators of $G$ which result in a length on $G$ with good properties.

**Lemma 4.19** ([1]). *Let $\eta \approx 0.811$ be the real root of $x^3 + x^2 + x - 2$ and set the weights*

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

*Then*

$$\eta\big(\omega(b) + \omega(a)\big) = \omega(c) + \omega(a),$$
$$\eta\big(\omega(c) + \omega(a)\big) = \omega(d) + \omega(a),$$
$$\eta\big(\omega(d) + \omega(a)\big) = \omega(b).$$ ∎

The next lemma is a small variation of [1, Proposition 5].

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

**Corollary 4.21.** *For all $g \in G$, the sequence of sets $(Suc_n^g)_{n \geq 1}$ from (4.4) stabilizes at a finite step.*

*Proof of Lemma 4.20 (see [1, Proposition 5]).* Each element $g \in G$ can be written as a word of minimal length of the form $g = a^\varepsilon x_1 a x_2 a \cdots x_n a^\zeta$, 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

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

Thus the length of $p_x(g)$ grows 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 $n \ge 3$ and $(g,\gamma)$ is a good pair with active constraint $\gamma$ with* $\operatorname{supp}(\gamma) \subset \{X_1, \dots, X_{2n}\}$*, then the constrained equation $(R_n(X_1,\dots,X_{2n})g, \gamma)$ 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 two 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 $\gamma \in \mathfrak{R}^4_{\mathrm{act}}$ such that $(q,\gamma)$ is a good pair.*

*Proof.* This can be checked in GAP using `verifyLemmaExistGoodGammasForRed4`. ■

To formulate an analog of Proposition 4.15, we literally transfer the definition of the function $\Gamma^q$ to the case $n = 2$. Denote the new function $\Gamma^{q,2}$. For a constraint $\gamma\colon F_{\mathcal{X}} \to Q$ with nontrivial activity it produces a finite set $\Gamma^{q,2}(\gamma)$ of constraints $\gamma'\colon F_{\mathcal{Y}} \to Q$ for an equation $(R_3 g', \gamma')$. The role of the specialized variable $Y_{6,1}$ is now played by $Y_{4,1}$. As above, we denote by $\mathcal{V} = \{Y_{1,1}, \dots Y_{3,2}\}$ the set of variables occurring in $R_3$.

**Proposition 4.25.** *For each good pair $(q,\gamma)$ with $q \in G'/K'$ and $\gamma \in \mathfrak{R}^4_{\mathrm{act}}$, the set $\Gamma^{q,2}(\gamma)$ contains some active constraint $\gamma'$ such that for all $g$ with $gK' = qK'$ the pair*

$$\big((g@2)^{\operatorname{rep}(\gamma'(Y_{4,1}))} \cdot g@1, \operatorname{red}(\gamma'|_{\mathcal{V}})\big)$$

*is a good pair.*

*Proof.* The proof is the same as for Proposition 4.15. Recalling that for $g \in G'/K'$ we denote by $\bar{g}$ its image in $G'/(K \times K)$, the corresponding formula which needs to be checked is

$$\forall q \in G'/K' \; \forall \gamma \in \mathfrak{R}^4_{\mathrm{act}} \; \exists \gamma' \in \Gamma^{q,2}(\gamma) \; \forall r \in G'/K' \text{ with } \bar{r} = \bar{p}_{\mathrm{rep}(\gamma'(Y_{4,1}))}(q): \\ \mathcal{P}(q,\gamma) \Rightarrow \mathcal{P}(r,\gamma').$$

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 $(g,\gamma)$ is a good pair with active constraint $\gamma$ with* $\mathrm{supp}(\gamma) \subset \{X_1,\dots,X_{2n}\}$, *then the constrained equation $(R_n(X_1,\dots,X_{2n})g,\gamma)$ 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 $i = 1,\dots,4$. 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 cannot 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 $\Psi\colon \mathrm{Aut}(T_n) \to \mathrm{Aut}(T_n) \wr S_n$, 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 the $n$-th level of the tree.

Consider the epimorphism

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

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

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

```
gap> G4 := 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 $\pi$ on the generators that this element is even a member of $K$. This finishes the proof of Theorem A.

### 4.6. Bounded conjugacy width

In [13], it is proven that $G$ has finite bounded conjugacy width. Corollary B, whose three statements are all proven below in this subsection, gives 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 $\gamma$. Independently of the chosen constraint, replacement of the variable $X_i$ by $\langle\!\langle Y_i, Z_i \rangle\!\rangle \operatorname{act}(X_i)$ leads after normalization to an equivalent equation $R_2(g@2)(g@1)$. Similarly to the construction of $\Gamma^q$ in the previous section, one can find for each $q \in G'/K'$ a constraint $\gamma$ 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 $\gamma'$ 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 exists an element $g \in G'$ such that the equation*

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

*is not solvable.*

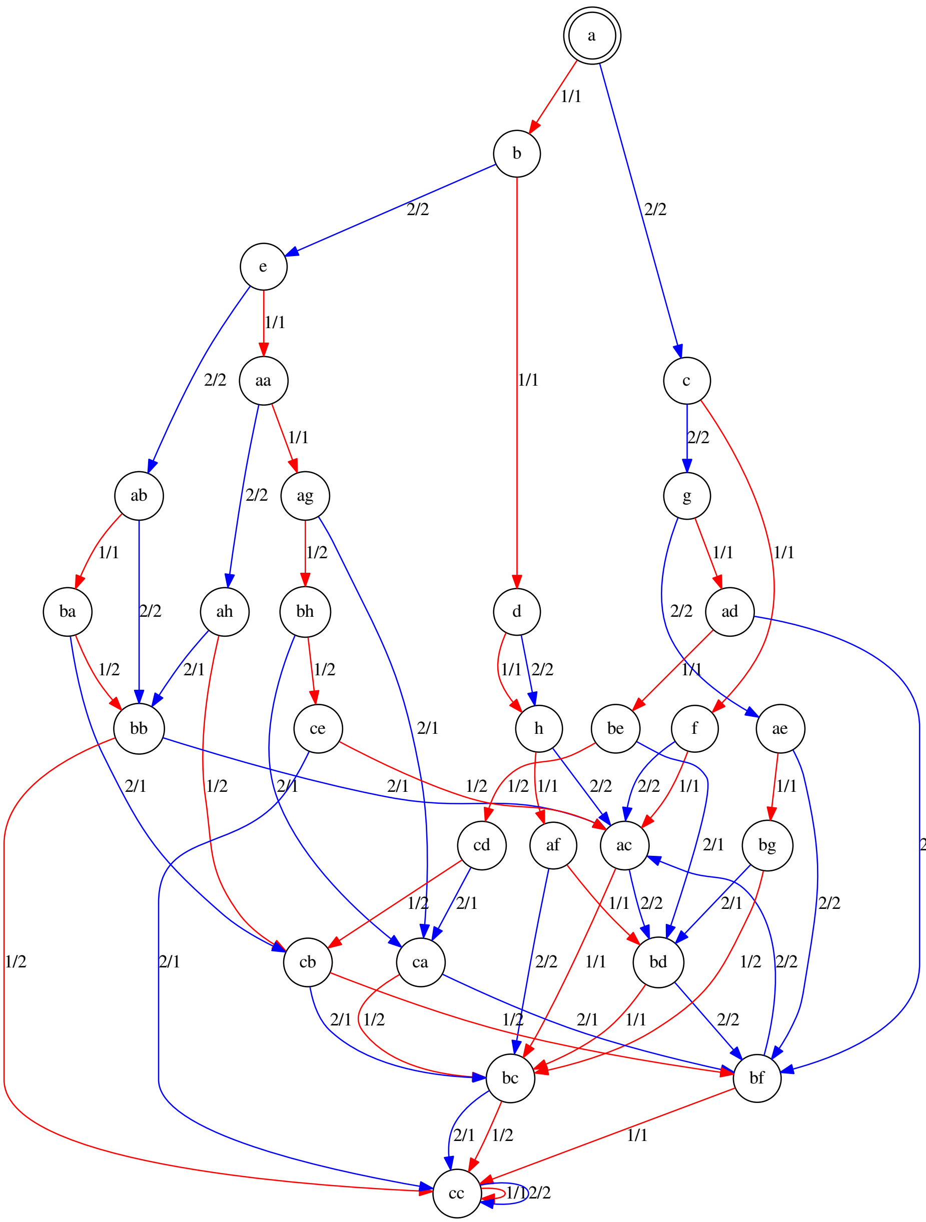


**Figure 1.** Element of the derived subgroup of the Grigorchuk group which is not a commutator. In standard generators:
$(acabacad)^3acab(ac)^2(acabacad)^2(acab)^3acadacab(ac)^2(acabacad)^2$
$(acabacadacab(ac)^3abacad(acab)^2)^5acabacadacab(ac)^2(acabacad)^2$
$(acabacadac)^2(abac)^3adacab(ac)^2(acabacad)^3acab(ac)^2(acab(ac)^3abacad)^2$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

*Proof.* As before, independently of the activities of a possible constraint $\gamma$ and of the element $g \in G'$, the normal form of $\widetilde{\Phi}_\gamma(a^{X_1}a^{X_2}a^{X_3}ag)$ 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 product 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 $\langle\langle \pi(cag), \pi(ac)\rangle\rangle \in K < G'$. ∎

**Definition 4.30** (Conjugacy width [13]). 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. ∎

This finishes the proof of Corollary B.

## 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 [5], we can find a nontrivial normal subgroup $N = \mathrm{Stab}_G(n) < H$ for some $n \in \mathbb{N}$. Since $K^{\times 2^n} < \mathrm{Stab}_G(n)$ there is an $n$ such that $K^{\times 2^n} \leq H$.

Since $K'$ has a 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 a 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 [17, Theorem 1].

This, by definition, means that every infinite finitely generated subgroup $H \le G$ contains a finite-index subgroup which is isomorphic to a finite-index subgroup of $G$. We can then repeat the argument from the proof of Proposition 5.1. ■

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 $\Gamma_n = F_n/\langle \gamma_3(F_n), x_1^2, \dots, x_n^2\rangle$. 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 $\Gamma_n/\Gamma_n'$. 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 $\binom{2^n}{2}^m \le 2^{(2n-1)m}$ products of $m$ commutators but the size of $\Gamma_n'$ is $2^{\binom{n}{2}} \ge 2^{\frac{n^2}{4}}$ and hence the commutator width of $\Gamma_{8m}$ 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 a subset of the vertices of $T_2$ defining 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 $G/\mathrm{Stab}_G(m)$. ■

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 cannot 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. ■

## 6. Implementation in GAP

### 6.1. Usage of the attached files

Running 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 $(q, \gamma)$ 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 $\mathfrak{R}^4_{\mathrm{act}}$ as claimed in Lemma 4.24.

**verifyPropExistsSuccessor.** This verifies that for each good pair $(q, \gamma) \in G'/K' \times (\mathfrak{R}_{\mathrm{act}} \cup \mathfrak{R}^4_{\mathrm{act}})$ there exists a $\gamma' \in \Gamma^q(\gamma)$ such that all preimages of $\bar{p}_{\mathrm{rep}(Y_{6,1})}(q)$ under the map $G'/K' \twoheadrightarrow G'/(K \times K)$ form good pairs with the constraint $\gamma'$. This is needed in the proof of Propositions 4.15 and 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 their 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 $\gamma$ 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 or multiple files and will override the original precomputed data. The argument is a string and can be one of the following:

**"orbits".** This computes 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 computes for each constraint $\gamma \in \mathfrak{R} \cup \mathfrak{R}^4$ the set of all $q \in G'/K'$ such that $(q,\gamma)$ is a good pair.

Then it computes for each good pair $(q,\gamma)$ one $\gamma' \in \Gamma^q(\gamma)$ with decorated $X = Y_{6,1}$ or $X = Y_{4,1} \in S$ as defined in (4.3) depending on whether $\gamma \in \mathfrak{R}_{\mathrm{act}}$ or $\gamma \in \mathfrak{R}^4_{\mathrm{act}}$ fulfills either Proposition 4.15 or Proposition 4.25. This computation takes 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 $a^{X_1}a^{X_2}a^{X_3}a^{X_4}a^{X_5}ag$. Letting $q \in G/K'$ be the image of $g$, this computes 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 $(\gamma * \pi)(\mathcal{E}_g) = \mathbb{1}$,

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

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

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

**"charactertable".** This computes the character table of the 4-th level germgroup and the set of irreducible characters. As the germgroup is quite large, this takes about 3 hours and has no 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 3,106 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 takes approximately 3 hours and is equipped with a progress bar.

**"all".** This performs all of the above one after another.

To recompute the orbits or the character table GAP should be started with the *-o* flag to provide enough memory for the computation. For example, start GAP with

```
gap -o 8G verify.g.
```

### 6.3. Implementation details

**6.3.1. Reduced constraints.** The proof of Lemma 4.3 in [22] 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 the fact that the quotient $Q = G/K$ is a polycyclic group with

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

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

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

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 than 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 to obtain 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 $\gamma$, the question whether $(g, \gamma)$ 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 constraint $\gamma$, to obtain all $q$ which form a good pair with $\gamma$, we can enumerate all possible commutators $[r_1, r_2][r_3, r_4][r_5, r_6]$ with $r_i K = \gamma(X_i)$. 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 $\Gamma_q$ in the function *GetSuccessor* in the file `gap/precomputeGoodPairs.g`. Given an element $q \in G'/K'$ and an active constraint $\gamma$, this function returns a tuple $(\gamma', X)$ with $\gamma' \in \mathfrak{R}$ and $X$ the decorated element $Y_{6,1}$ or $Y_{4,1}$ depending on whether $\gamma' \in \mathfrak{R}$ or $\gamma' \in \mathfrak{R}^4$.

Given an inactive constraint $\gamma$, it returns a pair of constraints $\gamma_1, \gamma_2$ such that both have nontrivial activity and (with $\omega$ the map from the branch structure)

$$\omega(\langle\!\langle \gamma_1(X_i), \gamma_2(X_i) \rangle\!\rangle) = \gamma(X_i).$$

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> ]
```

**Acknowledgments.** The authors are deeply grateful to Rachel Skipper and to the anonymous referee(s) for their remarks that helped us to greatly improve the presentation of this material.

## References


[1] L. Bartholdi, The growth of Grigorchuk's torsion group. *Internat. Math. Res. Notices* (1998), no. 20, 1049–1054 Zbl 0942.20027 MR 1656258

[2] L. Bartholdi, Algorithmic decidability of Engel's property for automaton groups. In *Computer Science—Theory and Applications*, pp. 29–40, Lecture Notes in Comput. Sci. 9691, Springer, Cham, 2016 Zbl 1477.20067 MR 3533843

[3] L. Bartholdi, FR, computations with functionally recursive groups, Version 2.3.6. 2016, GAP package, https://gap-packages.github.io/fr

[4] L. Bartholdi, Representation zeta functions of self-similar branched groups. *Groups Geom. Dyn.* **11** (2017), no. 1, 29–56 Zbl 1434.20005 MR 3641831

[5] L. Bartholdi and R. I. Grigorchuk, On parabolic subgroups and Hecke algebras of some fractal groups. *Serdica Math. J.* **28** (2002), no. 1, 47–90 Zbl 1011.20028 MR 1899368

[6] L. Bartholdi, R. I. Grigorchuk, and Z. Šunić, Branch groups. In *Handbook of Algebra, Vol. 3*, pp. 989–1112, Handb. Algebr. 3, Elsevier/North-Holland, Amsterdam, 2003 Zbl 1140.20306 MR 2035113

[7] L. Bartholdi and R. Hartung, Gap package LPRES, Version 0.3.0. 2016, GAP package, https://gap-packages.github.io/lpres

[8] W. Burnside, *Theory of Groups of Finite Order*. 2d edn., Dover Publications, New York, 1955 Zbl 1375.20001 MR 0069818

[9] D. Calegari, *scl*. MSJ Mem. 20, Mathematical Society of Japan, Tokyo, 2009 Zbl 1187.20035 MR 2527432

[10] L. P. Comerford Jr. and C. C. Edmunds, Quadratic equations over free groups and free products. *J. Algebra* **68** (1981), no. 2, 276–297 Zbl 0526.20024 MR 608536

[11] D. Eisenbud, U. Hirsch, and W. Neumann, Transverse foliations of Seifert bundles and self-homeomorphism of the circle. *Comment. Math. Helv.* **56** (1981), no. 4, 638–660 Zbl 0516.57015 MR 656217

[12] B. Farb and D. Margalit, *A Primer on Mapping Class Groups*. Princeton Math. Ser. 49, Princeton University Press, Princeton, NJ, 2012 Zbl 1245.57002 MR 2850125

[13] E. Fink, Conjugacy growth and width of certain branch groups. *Internat. J. Algebra Comput.* **24** (2014), no. 8, 1213–1231 Zbl 1318.20028 MR 3296364

[14] The GAP Group, GAP – Groups, algorithms, and programming, Version 4.7.5. 2014

[15] R. Grigorchuk, Solved and unsolved problems around one group. In *Infinite Groups: Geometric, Combinatorial and Dynamical Aspects*, pp. 117–218, Progr. Math. 248, Birkhäuser, Basel, 2005 Zbl 1165.20021 MR 2195454

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

[17] R. I. Grigorchuk and J. S. Wilson, A structural property concerning abstract commensurability of subgroups. *J. London Math. Soc. (2)* **68** (2003), no. 3, 671–682 Zbl 1063.20033 MR 2009443

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

[19] N. Gupta and S. Sidki, On the Burnside problem for periodic groups. *Math. Z.* **182** (1983), no. 3, 385–388 Zbl 0513.20024 MR 696534

[20] R. M. Guralnick, Expressing group elements as commutators. *Rocky Mountain J. Math.* **10** (1980), no. 3, 651–654 Zbl 0423.20035 MR 590227

[21] M. W. Liebeck, E. A. O'Brien, A. Shalev, and P. H. Tiep, The Ore conjecture. *J. Eur. Math. Soc. (JEMS)* **12** (2010), no. 4, 939–1008 Zbl 1205.20011 MR 2654085

[22] I. Lysenok, A. Miasnikov, and A. Ushakov, Quadratic equations in the Grigorchuk group. *Groups Geom. Dyn.* **10** (2016), no. 1, 201–239 Zbl 1341.20037 MR 3460336

[23] E. V. Muntyan, Verbal subgroups of automorphism groups of homogeneous rooted trees. *Mat. Stud.* **17** (2002), no. 2, 117–126 Zbl 1024.20023 MR 1950532

[24] V. Nekrashevych, *Self-Similar Groups*. Math. Surveys Monogr. 117, American Mathematical Society, Providence, RI, 2005 Zbl 1175.46048 MR 2162164

[25] Ø. Ore, Some remarks on commutators. *Proc. Amer. Math. Soc.* **2** (1951), 307–314 Zbl 0043.02402 MR 40298

[26] A. H. Rhemtulla, A problem of bounded expressibility in free products. *Proc. Cambridge Philos. Soc.* **64** (1968), 573–584 Zbl 0159.03001 MR 225889

[27] A. H. Rhemtulla, Commutators of certain finitely generated soluble groups. *Canadian J. Math.* **21** (1969), 1160–1164 Zbl 0186.03903 MR 248213

[28] A. V. Rozhkov, Centralizers of elements in a group of tree automorphisms. *Izv. Ross. Akad. Nauk Ser. Mat.* **57** (1993), no. 6, 82–105 Zbl 0829.20056 MR 1256568

[29] D. Segal, *Words: Notes on Verbal Width in Groups*. London Math. Soc. Lecture Note Ser. 361, Cambridge University Press, Cambridge, 2009 Zbl 1198.20001 MR 2547644

Received 9 June 2020.

**Laurent Bartholdi**
Fakultät für Mathematik und Informatik, Universität des Saarlandes, 66123 Saarbrücken, Germany; laurent.bartholdi@gmail.com

**Thorsten Groth**
Mathematisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen, Germany; thorsten.groth@mathematik.uni-goettingen.de

**Igor Lysenok**
Steklov Mathematical Institute, Russian Academy of Sciences, 119991 Moscow, Russia; lysionok@mi-ras.ru